UNIVERSITA' DEGLI STUDI DI PARMA

Dottorato di ricerca in Scienze Chimiche

Ciclo XXVIII (2013-2015)

Modified PNA design and synthesis: a novel approach using Molecular Dynamics and Metadynamics

Coordinatore:
Chiar.mo Prof. Roberto Cammi

Tutor:
Chiar.mo Prof. Roberto Corradini
Chiar.mo Dr. Vincenzo Verdolino

Dottorando:
Massimiliano Donato Verona

2016

Contents

1. Introduction

1.1. Modern challenges in drug discovery

Every year, millions of people all over the world come into contact with hard reality of a genetic disease or a tumor. Treatment of this pathologies is a great challenge for science, comparable to that against bacteria at the beginning of last century.[1] Penicillin, discovered by Fleming in 1928, started revolution of antibiotics that led, in a few time, to cure most of bacteriological infections. Treatment of genetic diseases is a complex problem that requires a multidisciplinary approach involving many different fields, ranging from medicine, biology, biotechnology, which provide pre-clinical and clinical studies, to chemistry, which allows the synthesis of new compounds, and to physics and engineering, which are necessary for the development of new materials. During the last decades, huge efforts have been done to make therapies more effective, to find new ones or, in general, to improve the quality of life of patients and their life expectation. Many steps forward have been done in particular for tumors, for which 5 years surviving is nowadays averagely of about 60%. Alzheimer, Cystic Fibrosis, Glioblastoma, Pancreatic cancer are just some, among many, that unfortunately are nowadays still almost incurable. Prevention is the first mean to sensitize population about bad habits and thus reducing incidence of some pathologies. Studies on mechanism of action of diseases allowed in many cases to better understand causes and to isolate new targets for the therapies. [2,3] Advancements in Science have indeed also provided new tools for early diagnosis and targeted therapy. The major approach for the development of new bioactive compounds have traditionally concerned small molecules. Drug discovery process was simplified using combinatorial approach or screening of molecules derived from natural products; prediction of the behavior of these could be done computationally using docking techniques.[4] Small molecules are relatively easy to synthesize and cheaper to scale up. On the other hand, often they could bind multiple targets causing many side effects difficult to predict and thus failure of many drugs at clinical phase tests can occur. Examples of commonly used drugs for treatment of cancer are reported in Figure 1.1.

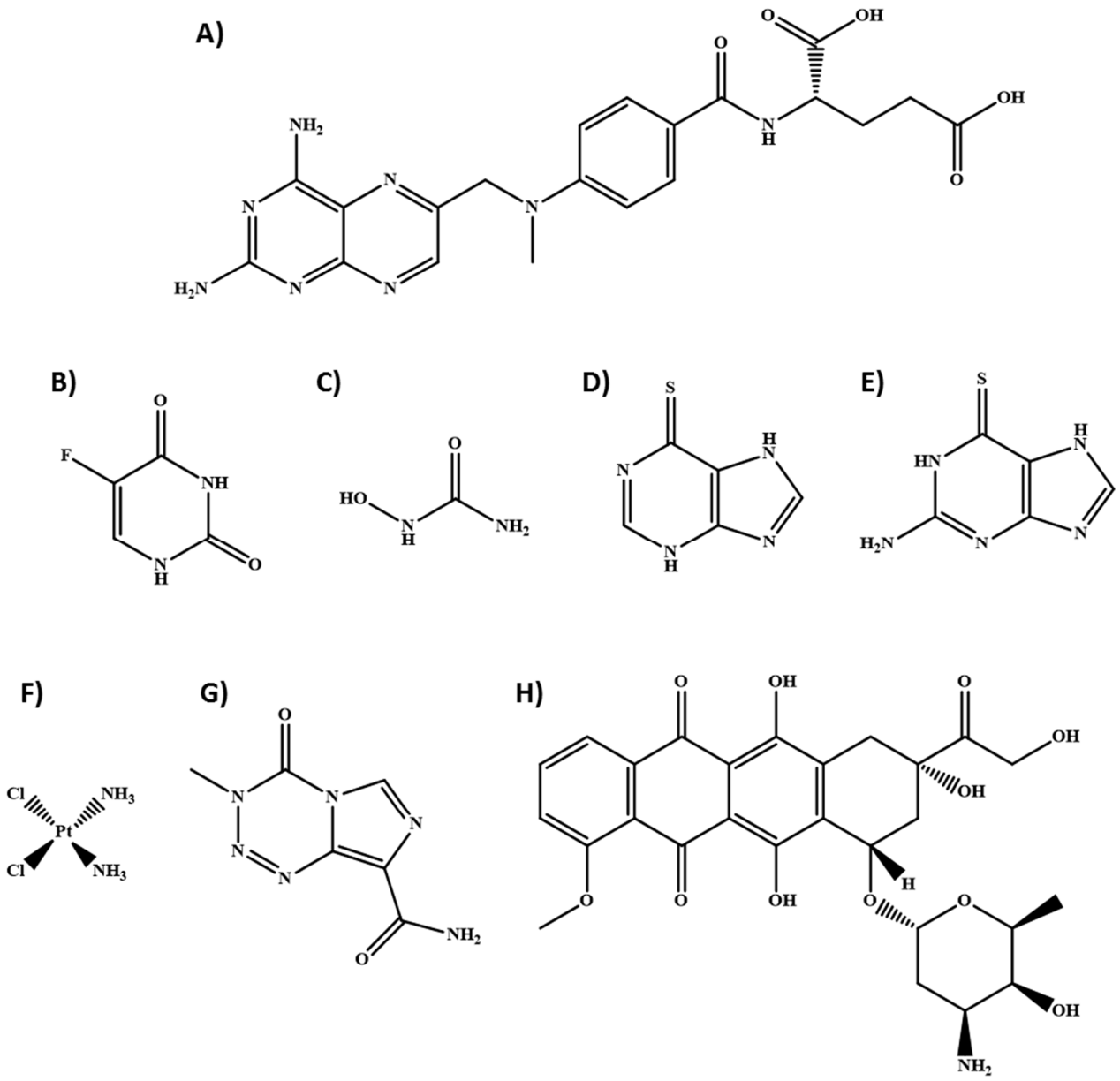


Figure 1.1: examples of antitumor agents used for therapy. **A**) Methotrexate; **B**) 5-Fluoro Uracil; **C**) hydroxyurea; **D**) Mercaptopurine; **E**) Thioguanine; **F**) Cisplatin; **G**) Temozolomide; **H**) Doxorubicin.

Generally, they follow two mechanisms of action: stop of DNA building blocks synthesis or DNA and RNA alkylation. Compounds from **A** to **E** belong to the first class. Methotrexate inhibits folic acid reductase that is responsible for conversion of folic acid to tetrahydrofolic acid.[5] This compound is a common intermediate necessary for the production of a cofactor used for nucleoside synthesis. Compounds **B**, **C**, **D** and **E** inhibit selectively one nucleotide formation and respectively thymine,[6] cytosine,[7] adenine[8] and guanine.[9] Compounds from **F** to **H** belong instead to the second class of chemotherapic agents. Cisplatin is one of the first and simplest drugs used in tumor treatment, but is still used together other platinum based compounds.[10] Temozolomide is used as adjuvant against

some brain cancers,[11] and Doxorubicin and its derivatives are used as therapy for leukemias or lymphomas.[12]

Alongside this small-molecule approach for the treatment of cancers, new innovative strategies are being developed in the recent years in order to maximize effectiveness and selectivity. For example, the cell therapy, which consists in diseases treatment by injection of cellular material, has recently reached a translational level. Many different types of cells are used and, among others, Stem Cells and T Cells have showed interesting therapeutic properties. Stem Cells have been proposed for their ability to differentiate in different specialized cells. This particular behavior finds then application in regenerative medicine, allowing to restore damaged tissues. Nowadays these cells are widely studied for treatment of neurodegenerative diseases,[13,14,15] with the aim to recover damaged neurons and restore normal cerebral activity. T cells are a type of lymphocytes and have been used for immunotherapic approach.[16] These cells are able to kill cancer cells and, in particular, their resulted effective also in treatment of metastasis.[17,18] For example, Rosenberg and coworkers reported use of engineered T lymphocytes for effective reduction of metastatic melanoma,[19] inducing rapid reduction of the tumor and absence of tumor recurrence one year after the treatment. However, cure of diseases with cells presents some drawbacks, in particular costs, since culture results in extremely low yields, and difficult delivery.

Use of therapeutic antibodies is another approach diffused in recent years. In particular several monoclonal antibodies (mAbs) have been studied and nowadays are applied in diagnostic[20,21] and treatment of different pathologies.[22] Antibody engineering allowed to generate new artificial compounds able to selectively bind targets, but at the same time to avoid rejection and clearance by patient's immune system.[23] Chimeric antibodies represent the first generation of compound in which human constant domains are fused with murine variable ones (human portion ~70%). Humanized antibodies represent instead the second generation and present only a little fraction of murine domains (human portion 85-90%). Mechanisms of action of these compounds is not always clear or can derive from multiple effects. Most common pathways consist in a direct action of mAbs that can interact with specific receptors on the surface of cells or in stimulating immune-mediated tumor cell killing. Monoclonal antibodies can be used also in conjugation with drugs or toxins that can be released in cells.[24] Rituximab is one of the best examples of commercial antibody. It is a chimeric compound against protein CD20, which is mainly present on the surface of B cells. Therefore, is used in case of excess or over reactivity or dysfunctionality of this type of cells. Nowadays Rituximab is commonly administered in treatment of different diseases like lymphomas,

leukemias and other autoimmune disorders.[25,26] Despite their unique properties of activity and selectivity, monoclonal antibodies are large proteins (~ 150 kDa) and their synthesis is still very expensive. Moreover, they must be administered in large amounts to achieve clinical results due mainly to scarce permeation toward target cells.

Nanomedicine is another field of research well explored in recent years. Advance of chemistry in the field of materials allowed to easily produce a large variety of compounds. Nanomaterials apply to both diagnostics,[27,28] and therapeutics,[29,30] often with the goal of achieving both effects (theranostic).[31] Due to their properties of penetration in cells, nanoparticles are good candidates for drug delivery.[32] A first generation of nanocarriers consisted in passive delivery systems (Figure 1.2A). The most known compounds of this type, that are in clinical use, are liposomes,[33] in which drugs are trapped in vesicles constituted by lipid bilayers, and mixed structures such as nab-paclitaxel,[34] in which toxic paclitaxel is bound to albumin nanoparticles. This kind of systems often are nonspecific toward tumors, causing undesired side effects. This led to a second generation of nanomaterials in which nanocarriers are functionalized with antibodies,[35] peptides,[36,37] ligand[38,39] and other biomolecules[40–42] in order to discriminate target cells. Many examples are available in the literature, but the use of these compounds in clinic is still limited. Particular attention has been given to biocompatibility of these compounds, since some types of nanoparticles may result very toxic.[43]

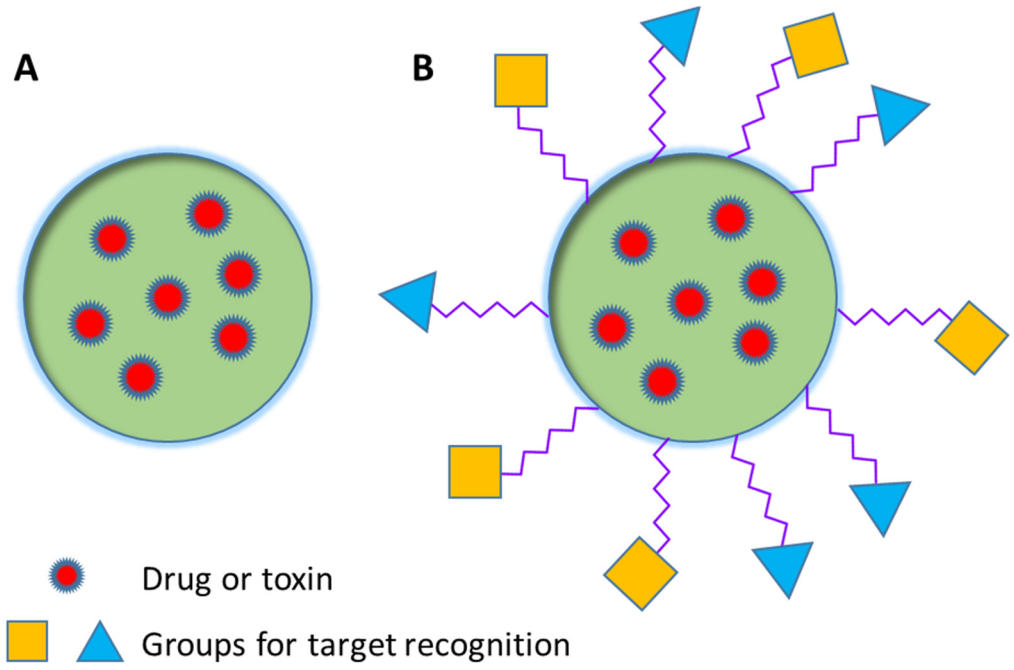


Figure 1.2: Schematic representation of first (**A**) and second (**B**) generation nanostructures for drug delivery.

Finally, the treatment of diseases by rationally targeting of nucleic acids by bio-macromolecules is a long-standing, and still very active, field of research, which is the subject of this thesis and will be discussed in detail in the following paragraphs.

1.2. Nucleic Acids Targeting

Since many diseases are linked to DNA expression and mutation,[44,45] a part of the research focuses on genes or on transcription, transport and translation of their information. Main processes involved in gene expression are represented in Figure 1.3

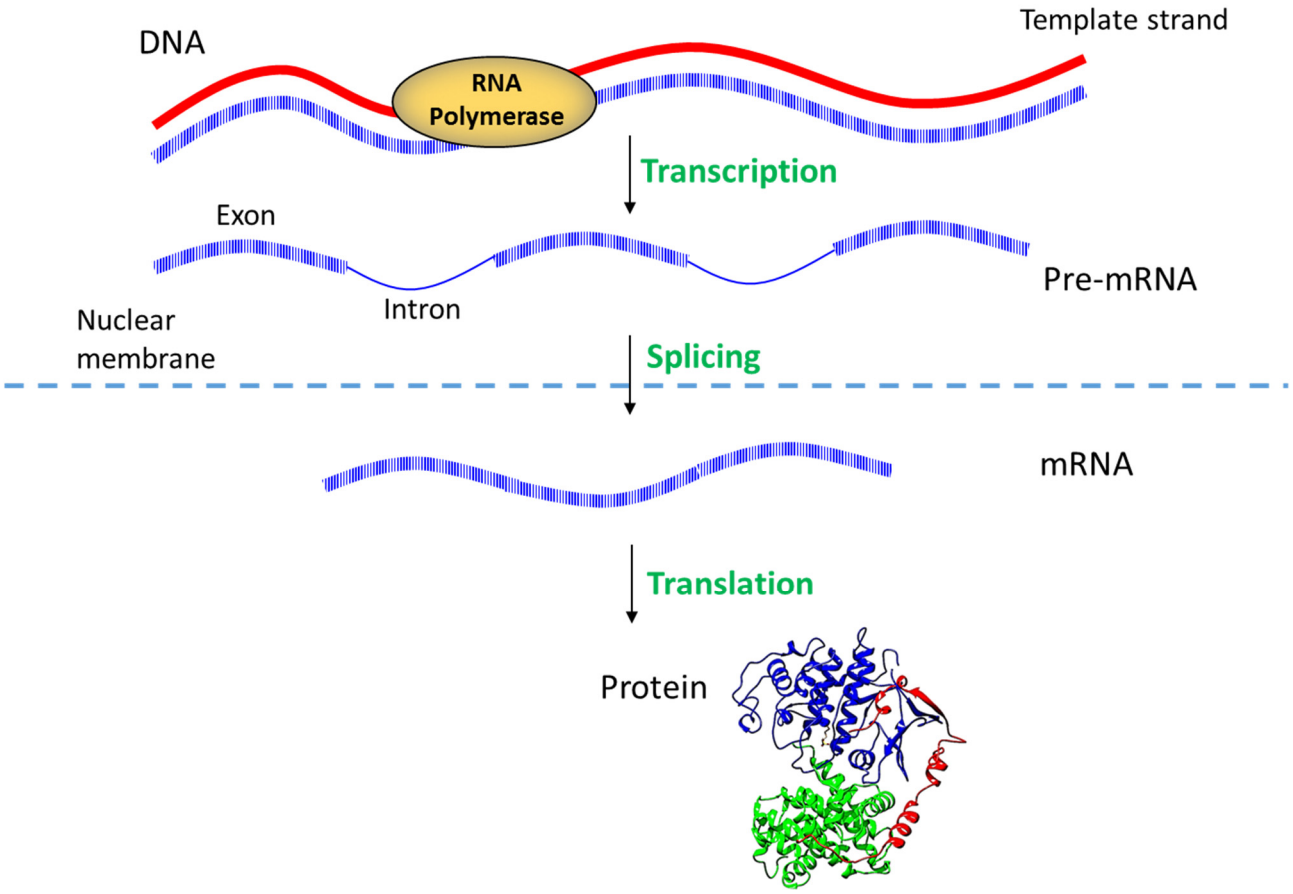


Figure 1.3: Schematic representation of gene expression processes.

The antigene approach consists in directly targeting DNA.[46] Indeed, it is possible to bind double stranded DNA with oligonucleotide fragments or artificial analogues, forming triplex structures or invading double helix and thus interfering with transcription. However, direct interaction with DNA is difficult because the targeting molecules must enter into cells and then into nucleus.[47] Another

possibility for affecting transcription is the use of decoy compounds[48] for transcription factors. These proteins are able to recognize specific gene sequences and then promote or block RNA polymerase activity. Decoys bind to these factors preventing or reducing interaction with target DNA and thus controlling transcription process.[49]

Pre-mRNA is a precursor of messenger RNA in which exons and introns are both present. To obtain mature mRNA is necessary to remove introns and correctly join exons (splicing); this process is usually performed by spliceosomes and often is possible to obtain different mRNAs from the same pre-mRNA, by changing exons recombination. Physiological conditions can then influence production of different proteins, but it is also possible to have errors in this phase leading to non-functional or partially functional proteins (aberrant splicing). Oligonucleotides or analogues are used for directing splicing process[50] or to block aberrant slicing favoring normal one.[51]

The antisense approach[52] consists in targeting messenger RNA. Formation of a duplex with mRNA prevents translation and RNase-H mediated cleavage of RNA strand.[53] mRNAs are found in cytoplasm and are then easier to target. Similarly, it is also possible to bind micro RNA using an anti-miR approach (this methodology will be discussed further in section 1.3).[54,55]

DNA or RNA oligonucleotides can directly target specific complementary sequences by the Watson-Crick base pairing scheme and can thus provide a rational design of new drugs; the use of this approach using naturally occurring nucleic acids in pharmaceutical applications is difficult because DNA, and even more RNA, are quite unstable in biological media due to the presence of nucleases. Moreover, the cellular uptake of naturally occurring nucleic acids was found to be extremely poor due to the phosphate negative charge. To overcome these problems, it has been necessary to move to modified artificial DNAs. These systems must retain, and hopefully increase, ability to bind to the complementary sequence of DNA and at the same time present better resistance, cellular uptake and, possibly, new functionalities. Among the first proposed compounds (Figure 1.4), phosforothioates[56]or alkyl phosphonate[57] presented modifications on the phosphate group; in the first case an atom of oxygen was substituted with sulfur, while in the second with an alkyl group. In both cases modifications induced chirality in the system leading to two different diasteroisomer with different behaviors.[58]

2'-O-Alkyl[59] (and in particular 2'-O-methyl)[60] or 2'-fluoro-oligonucleotides[61] were the first examples of sugar modified nucleotides proposed to substitute RNA. Indeed, since the OH in 2' position greatly enhance hydrolysis rate, compared to DNA, due to possible intramolecular attack on phosphate, substitution at this site enhances the stability of these derivatives.

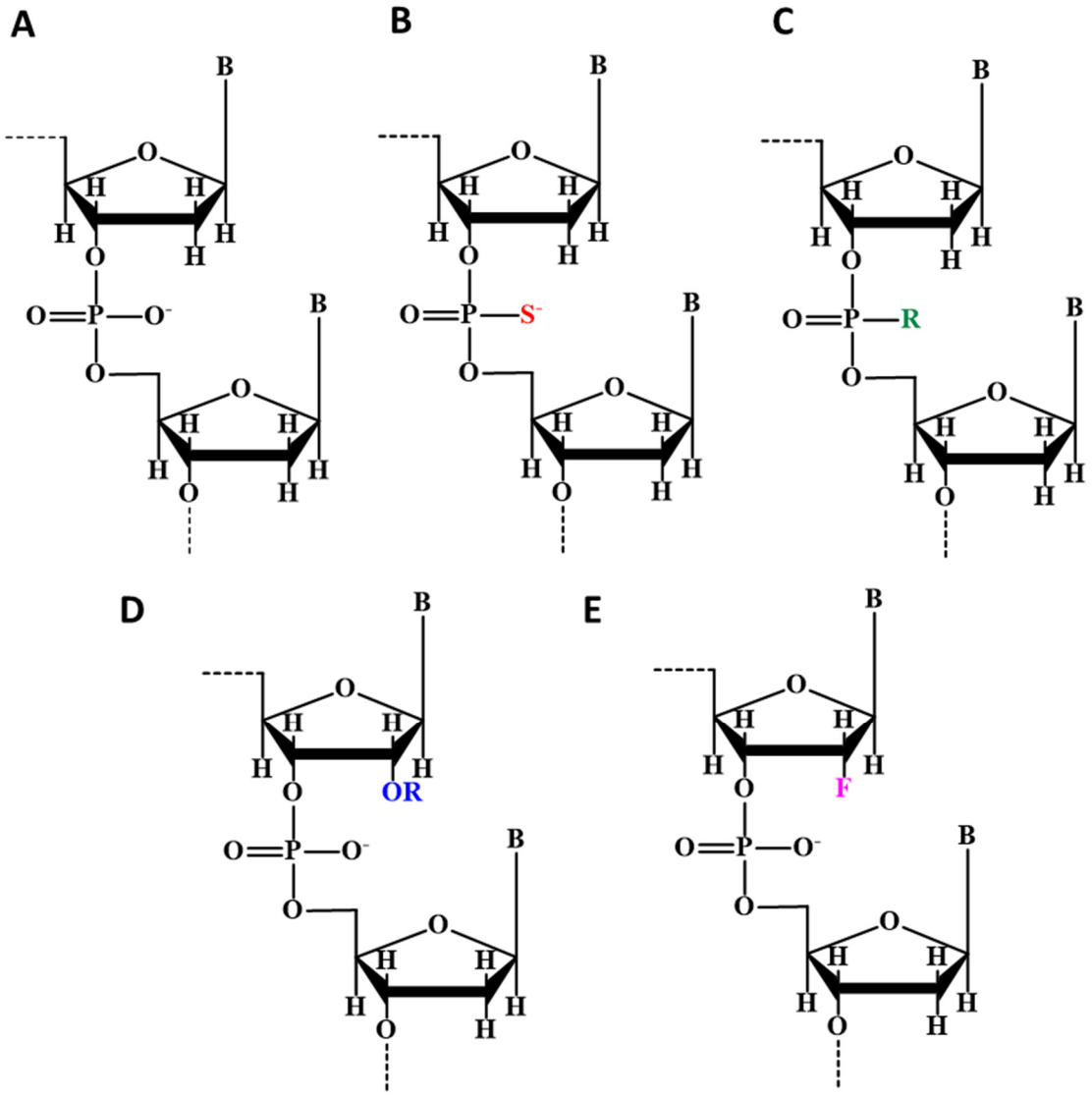


Figure 1.4: **A**) standard DNA; **B**) Phosphorothioate; **C**) alkyl phosphonate; **D**) or 2'-O-Alkyl RNA; **E**) 2'-Fluoro RNA modifications.

Tricyclo DNA[62] (Figure 1.5) are examples of conformationally constrained DNA that are not recognized by nucleases and recognize RNA with high affinity.[63] Threose Nucleic Acids[64] are analogues in which ribose is substituted with threose sugar ring. Although TNAs are artificial compounds they are thought as possible precursor of DNA or RNA.[65,66] Among sugar modified systems, Locked Nucleic Acid (LNA) are very effective ones and present high affinity and selectivity toward target sequence.[67] These type of compounds present oxygen in 2' position linked to 4' carbon with a methylene bridge, thus reducing hydrolysis. Moreover, this linkage blocks the

nucleoside unit in a C3' endo (N-type) conformation, which is typical of RNA. Many variations of this system have been proposed during the years.[68]

It is also possible to replace sugar rings and phosphate with different units. The class of morpholino[69,70,71] oligonucelotides is an example in which the nucleic acid backbone is substituted with methylenemorpholine rings and phosphorodiamidate linkages; triazole DNAs[72,73] present a imidazole ring covalently bound to ribose.

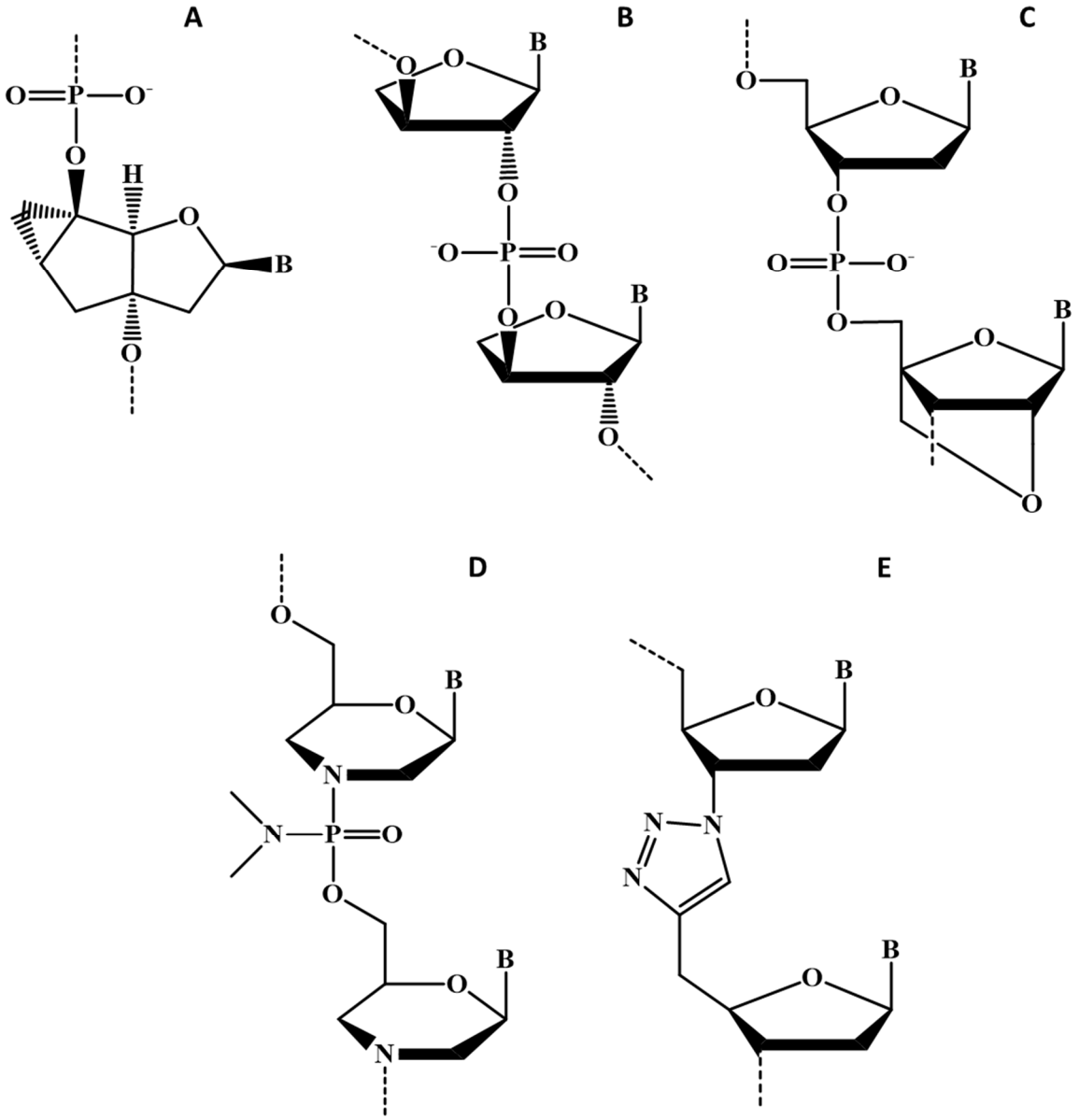


Figure 1.5: **A**) Tricyclo DNA; **B**) TNA; **C**) LNA; **D**) Morpholino phosphodiamides; **E**) Triazole DNA modifications.

1.3. Micro-RNA as therapeutic targets

MicroRNAs (miRNAs) are small RNAs, 21–25-nucleotide long, that regulate gene expression at the post-transcriptional level.[74] miR up- or down-regulation has been correlated with many different diseases like hearing loss,[75] hearth failures[76,77,78] and, most notably, different types of cancer,[79,80] including leukemia.[81,82,83] Gene expression is regulated by guiding RNA-induced silencing complex (RISC) that is able to bind target mRNA and to degrade it (Figure 1.6).

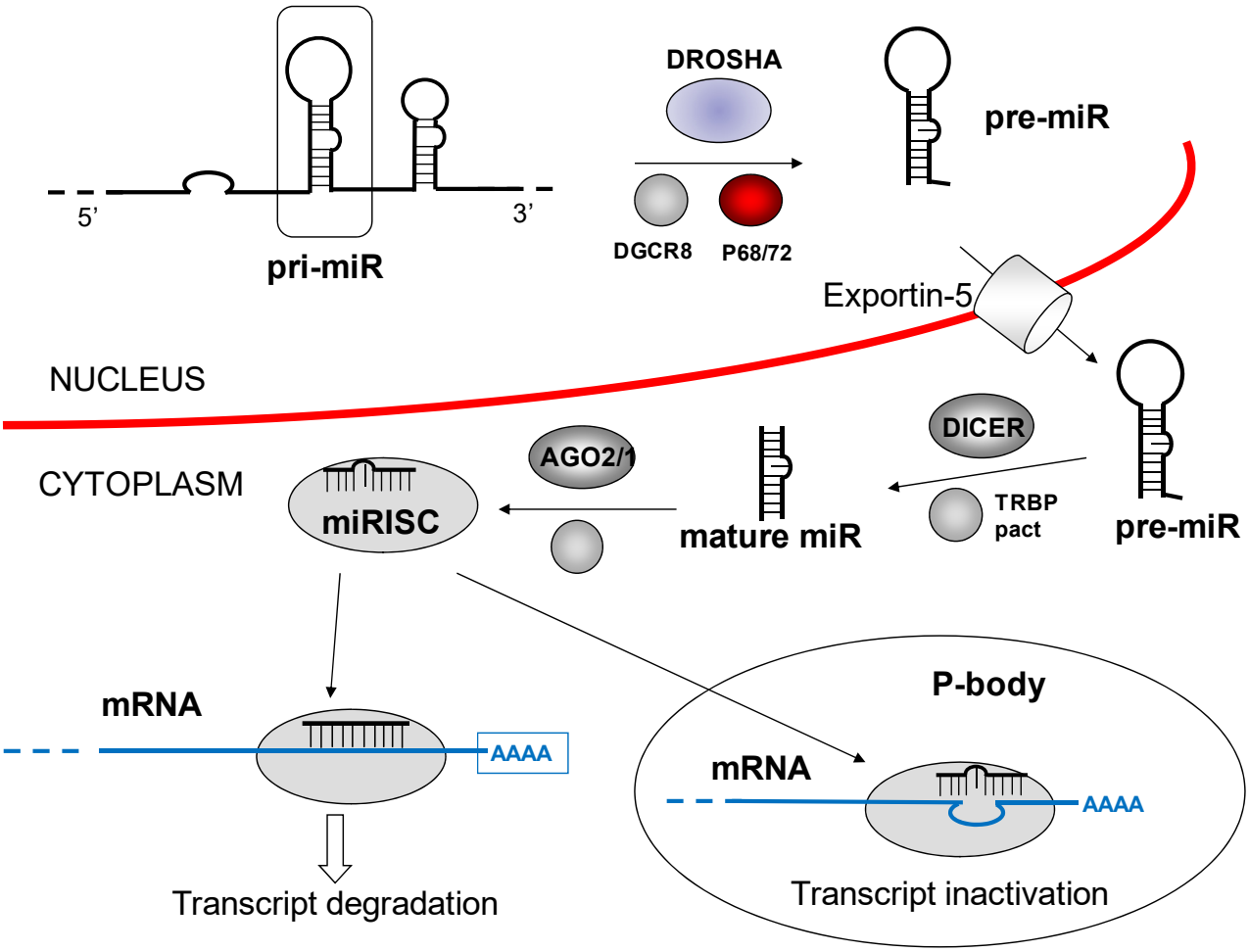


Figure 1.6: Mechanism of maturation of microRNAs and their activity through RISC complex.

As an example of the potentiality of miR-based therapeutics, the work by Giacca and coworkers,[84] who studied miR connected to cardiac growth, is worth citing. In embryos enlargement of the hearth is principally due to proliferation of cardiomyocytes; this process ends at birth and growth is then connected to hypertrophic enlargement of myocytes. miR 590 and miR 199a were found to stimulate cardiomyocyte proliferation and were used to promote cardiac regeneration. Infarcted

mice treated with this two miR showed a regeneration of cardiac tissue and an almost complete recovery of functional parameters (Figure 1.7). Thus, this miR based approach is the only available example of possible cure of infarctuated heart by administration of drugs.

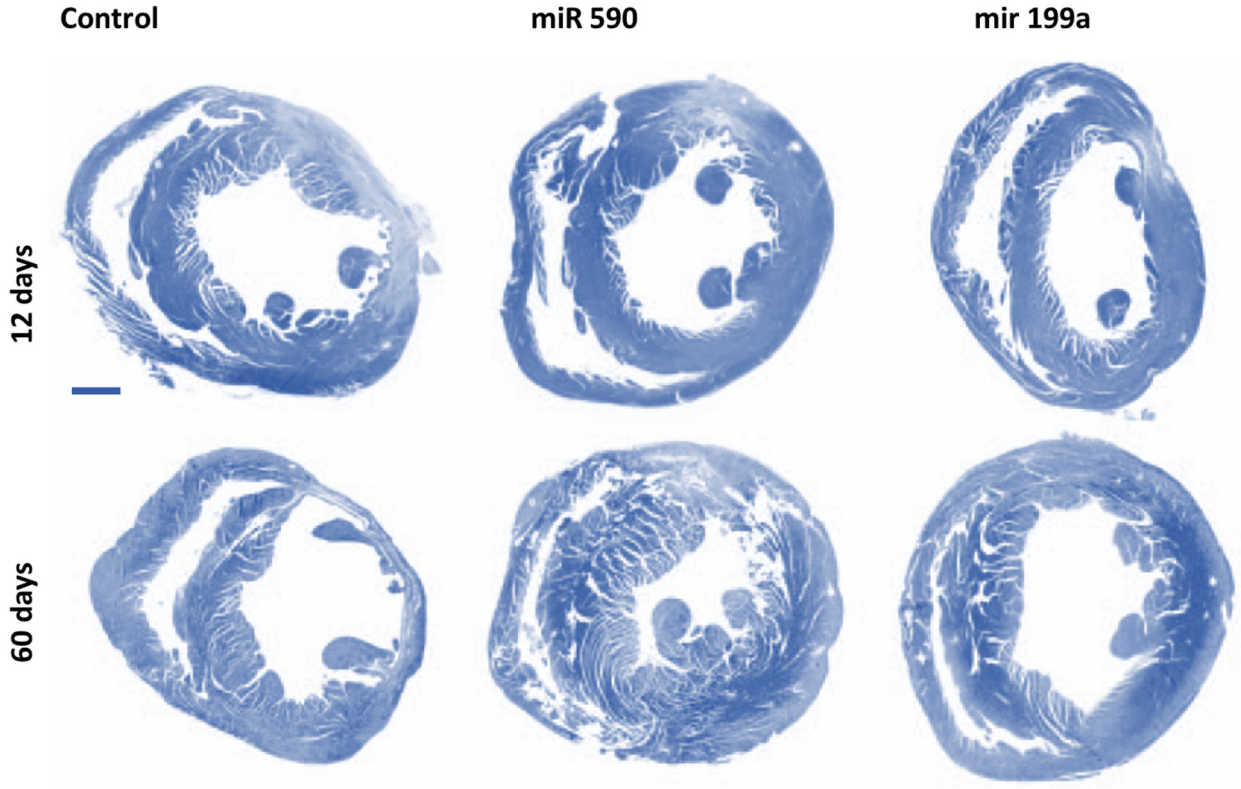


Figure 1.7: Heart cross section of infarcted mice after 12 days and after 60 days. From left, control (untreated mouse), mouse treated with miR 590 and mouse treated with miR 199a. Reprinted from ref. 84 with the permission of Nature Publishing Group.

Due to problems explained before, usually use of modified systems is preferred. AntagomiR are synthetic 2'-O-methyl RNA (Figure 1.4D) that are able to strongly bind micro RNA.[85] In order to further increase resistance to nucleases usually phosphorothioate modifications are inserted on the first two and last four bases and cholesterol is linked at 3' end to enhance distribution and cell permeability.[86] These type of compounds is commercially available and is actually tested for treatment of several pathologies.[87,88,89]

LNA have also been successfully used to target microRNA, for example in treatment of Hepatitis C Virus.[90] Miravirsen is a LNA/phosphorothioate hybrid that is able to efficiently bind miR-122 (Figure 1.8) that is overexpressed in affected liver cells, and activates virus replication.[91] This drug has demonstrated broad antiviral activity and it is actually in phase 2 clinical trial.

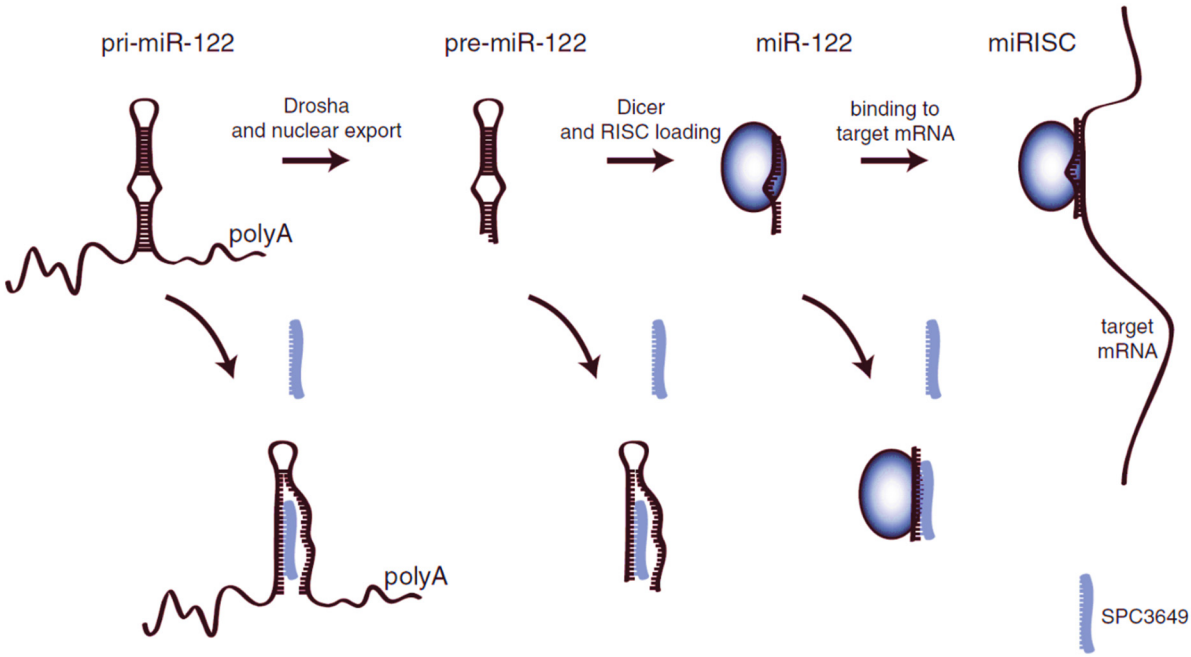


Figure 1.8: Proposed mechanism of action of anti-miR drugs; Miravirsen (SPC3649) is taken as an example. Reprinted from ref. 90 with the permission of Oxford University Press.

An emerging challenging application is use of peptide nucleic acids to inhibit action of micro-RNAs by targeting their sequences.

1.4. Peptide Nucleic Acids

1.4.1. General properties

Peptide nucleic acids (PNAs)[92] are particularly interesting DNA mimicking compounds, in which the deoxyribose phosphate backbone is substituted by *N*-(2- aminoethyl)glycine units and bases are linked to the backbone with methylenecarboxylic moieties (Figure 1.9).

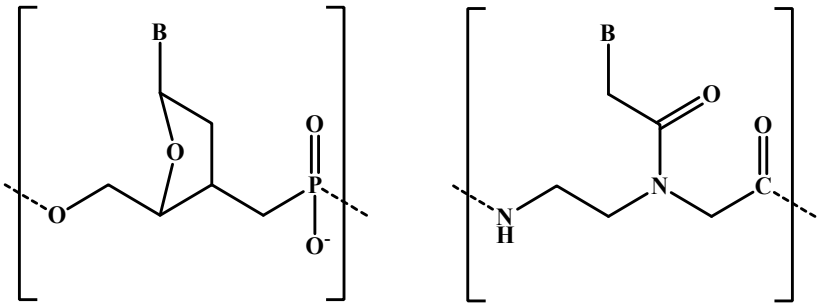


Figure 1.9: left, structure of DNA; right, general structure unmodified Peptide Nucleic Acids.

Like DNA, generally PNAs form a double helix with the target DNA sequence through Watson and Crick bonds;[93] in the case of homopyrimidine sequences PNA can also form triplex structures with homopurine DNA strands by forming both Watson-crick and Hoogsteen bonds.[94] Depending on the sequence, pairing in PNA:DNA duplexes can occur both in parallel and antiparallel orientations (as defined in Figure 1.10), but generally latter is preferred.[95]

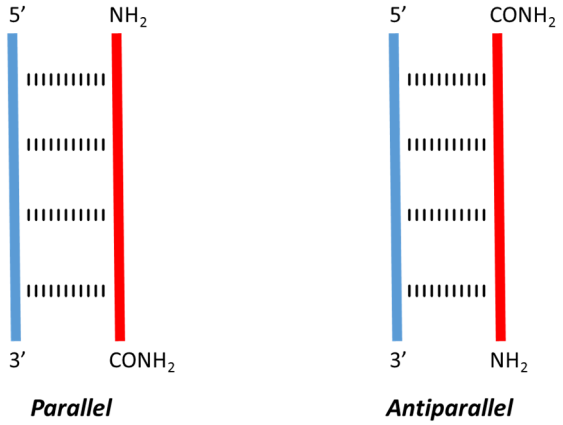


Figure 1.10 : Possible orientations for PNA:DNA (RNA) duplexes. Blue strand is DNA or RNA, while red strand is PNA.

The amidic backbone of PNA is neutral, therefore repulsive electrostatic effects, present in DNA or RNA duplexes, are avoided, and this induces high stability of the PNA duplexes with complementary DNA or RNA. Furthermore, compared to oligonucleotides or their analogs, PNA present higher specificity in sequence recognition, even for a single base mismatch.[95] Moreover previous studies revealed that PNAs display high chemical and biological stability,[96] making them particularly suitable for *in vivo* use.

PNA can be obtained with protocols similar to those used for peptides: synthesis is routinely performed on a resin and monomers require presence of a temporary protecting group, that permit chain elongation, and of semi-permanent groups, that protect nucleobases, avoiding side reactions or multiple chains grow (Figure 1.11). Many strategies have been proposed, but Fmoc/Bhoc[97] and Boc/Cbz[98] ones are the most used and thus the corresponding protected monomers are commercially available.

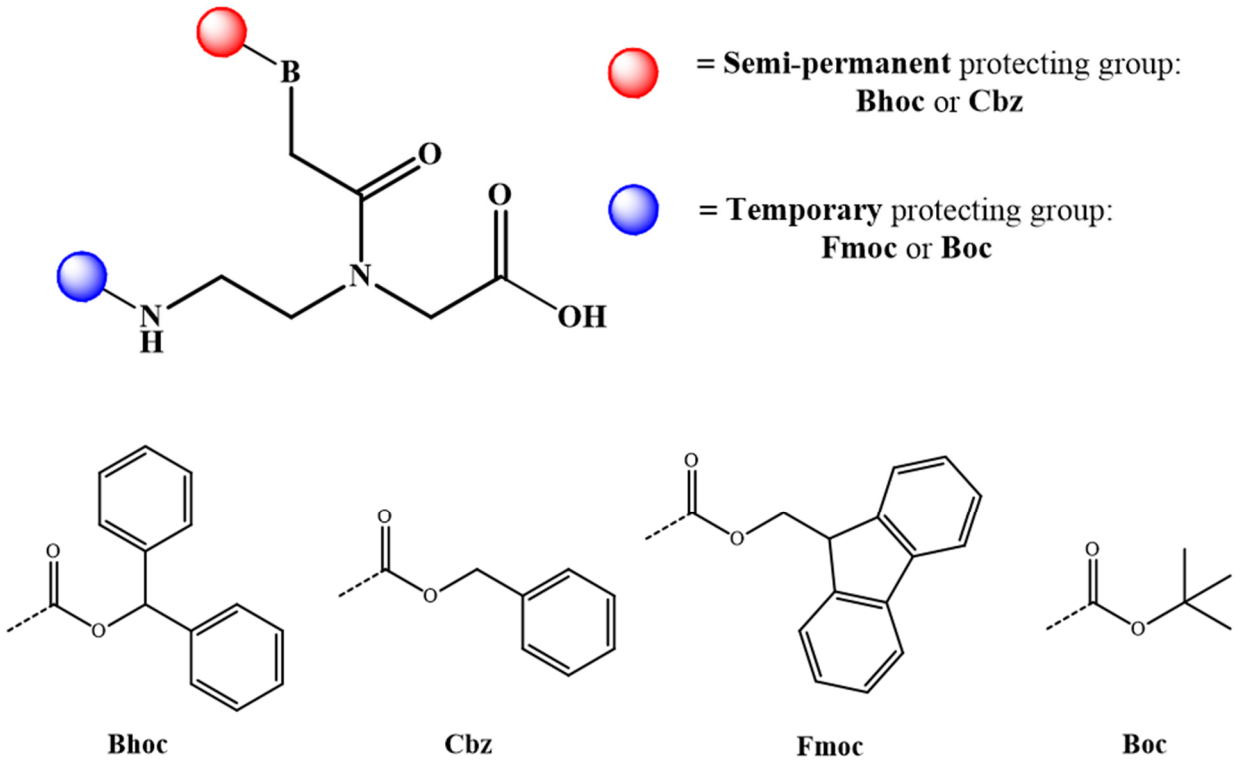


Figure 1.11: Representative monomer structure for solid phase synthesis of PNA.

PNAs have been successfully used in different field ranging from analytical methods, in which they are used as probes in sensors for DNA recognition,[99,100,101] to diagnostics, in which they are coupled with reporter group for imaging,[102,103] and, recently, to therapeutics, as potential new drugs.[104,105]

1.4.2. Modifications

Since their introduction in 1991, many modifications have been proposed to improve PNAs properties or to introduce new functionalities. Modifications can be introduced on the backbone,[106,107] or on the nucleobases.[108,109] Those on the backbone usually were designed to affect the conformational freedom of the system and, in general, to reduce entropy loss during duplex formation.[110] Insertion of chiral substituents in C2 or C5 position of the backbone (Figure 1.12) can modify handedness of the helix,[111] favoring or disfavoring the binding with DNA.

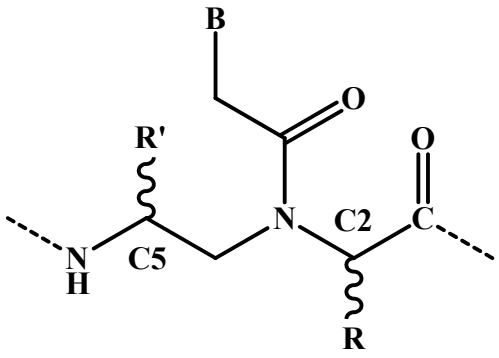


Figure 1.12: Schematic representation of PNA bearing modifications in C2 and C5 carbon atoms.

Substituent groups in C2 position, derived from D-amino acid synthons, induce preference to form a right handed helix, while L groups induce preferred left-handedness; for C5 position the stereochemical preference is inverted, since predisposition to form right-handed helices was observed for the derivatives obtained from L-amino acids. Maximum positive effect on PNA:DNA duplex stability was found with a combination of both modification: a D-amino acid derived group in C2 position and a L-amino acid derived group in C5, leading to a stabilization of ~ 7°C compared to unmodified PNA. Opposite case, with a constructive left-handed induction, caused a drop of T_m of more than 30 °C (Table 1.1).[111]

Table 1.1 : Effect of chiral substituents in C2 and C5 positions. Table is taken from ref (111).

PNA	Helical induction C2	Helical induction C5	Overall helical preference	Chiral accordance/conflict	PNA–DNA Thermal stability [°C]
2D,5L	right-handed	right-handed	right-handed	accordance	57
5L	–	right-handed	right-handed	–	56
2L,5L	left-handed	right-handed	right-handed	conflict	52
2D	right-handed	–	right-handed	–	52
Achiral	–	–	–	–	50
(2L	left-handed	–	left-handed	–	47
2D,5D	right-handed	left-handed	left-handed	conflict	33
5D	–	left-handed	left-handed	–	32
2L,5D	left-handed	left-handed	left-handed	accordance	20

Insertion of groups in C2 position modify also the selectivity of PNA probe. “Chiral box” PNA (Figure 1.13) with three consecutive D-lysine monomers in the middle of the sequence led to a high drop of stability in presence of a single mismatch.[112,113] D-Lysine containing chiral box PNA was effective in

inducing preference for the antiparallel orientation (with high direction control) while L-Lysine containing PNA favored the parallel one.

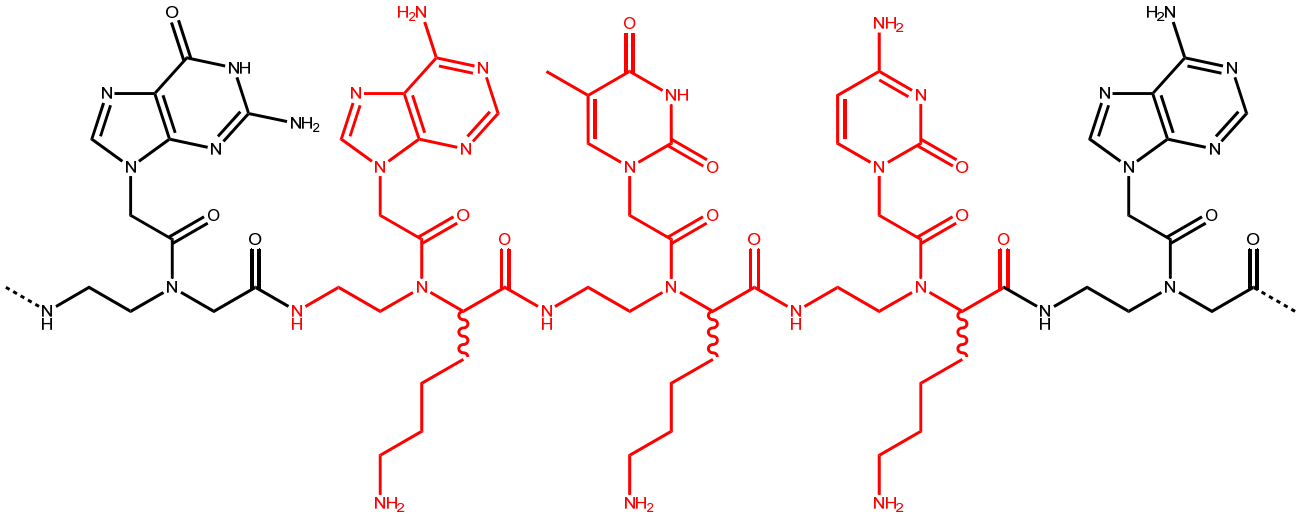


Figure 1.13: "Chiral Box PNA" bearing Lysine group in C2 position.

On the contrary, C5 modification influence principally the preorganization of the system, proved by strong circular dichroism (CD) signals for single strand PNA. [114] Insertion of positively charged groups in the same position favored cellular uptake of the PNA;[115,116] it can also be improved by conjugation with carrier peptides,[117] like octa-arginine[118] or NLS.[119] However, incorporation (embedding) of the amino acid side chains into the PNA structure was shown to increase resistance to peptidases.[118] Modifications of bases allow to improve PNA properties or to introduce new functionalities. In order to improve stability of the complex is possible, for example, to use bases with larger π-surface area, keeping H-bonding pattern.[120,121]

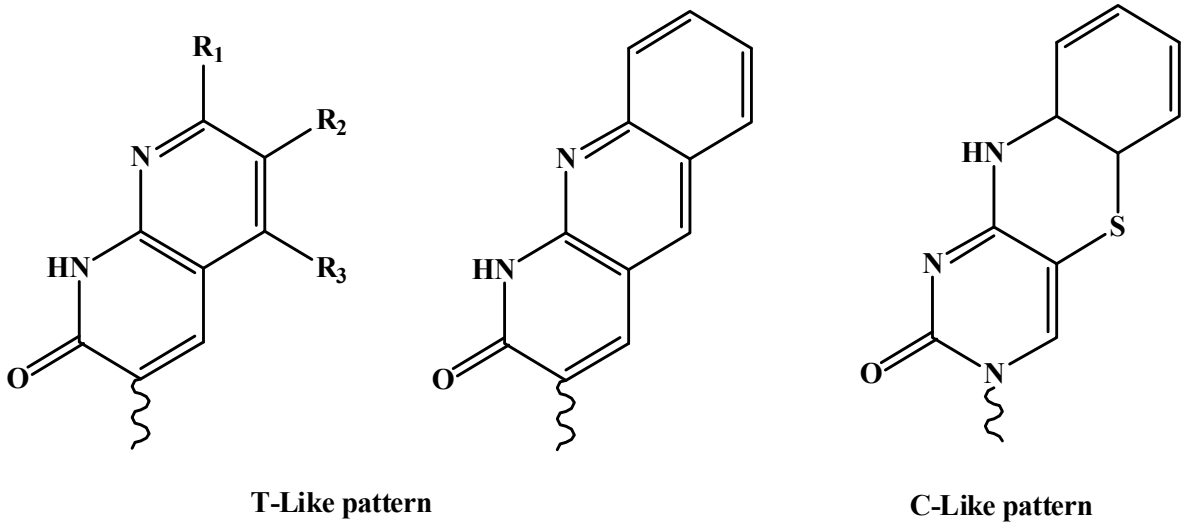


Figure 1.14: Examples of extended π-surface bases with thymine or cytosine bonding pattern.

Modified bases can be used also to insert reporter groups,[122,123] allowing their use for analytical or diagnostic purposes. Particularly interesting are modifications that take advantage of Hoogsteen sites of nucleobases in addition to Watson-Crick interactions, thus increasing the number of H-bonds and therefore stability of the duplex. The so-called "G-clamp" bases (Figure 1.15), introduced by Matteucci and co-workers,[124] belong to this type and in particular are cytosine analogues with extra moieties able to "chelate" target guanine. Many systems have been reported in literature,[125] some examples are shown in Figure 1.15. It has been demonstrated that system **B** is able to form one additional H-bonds with guanine, while **C** is able to form two extra H-bonds.[126] A very effective system is **D**, which, if introduced in a PNA sequence, led to a stabilization of 12 °C compared to unmodified PNA.[127]

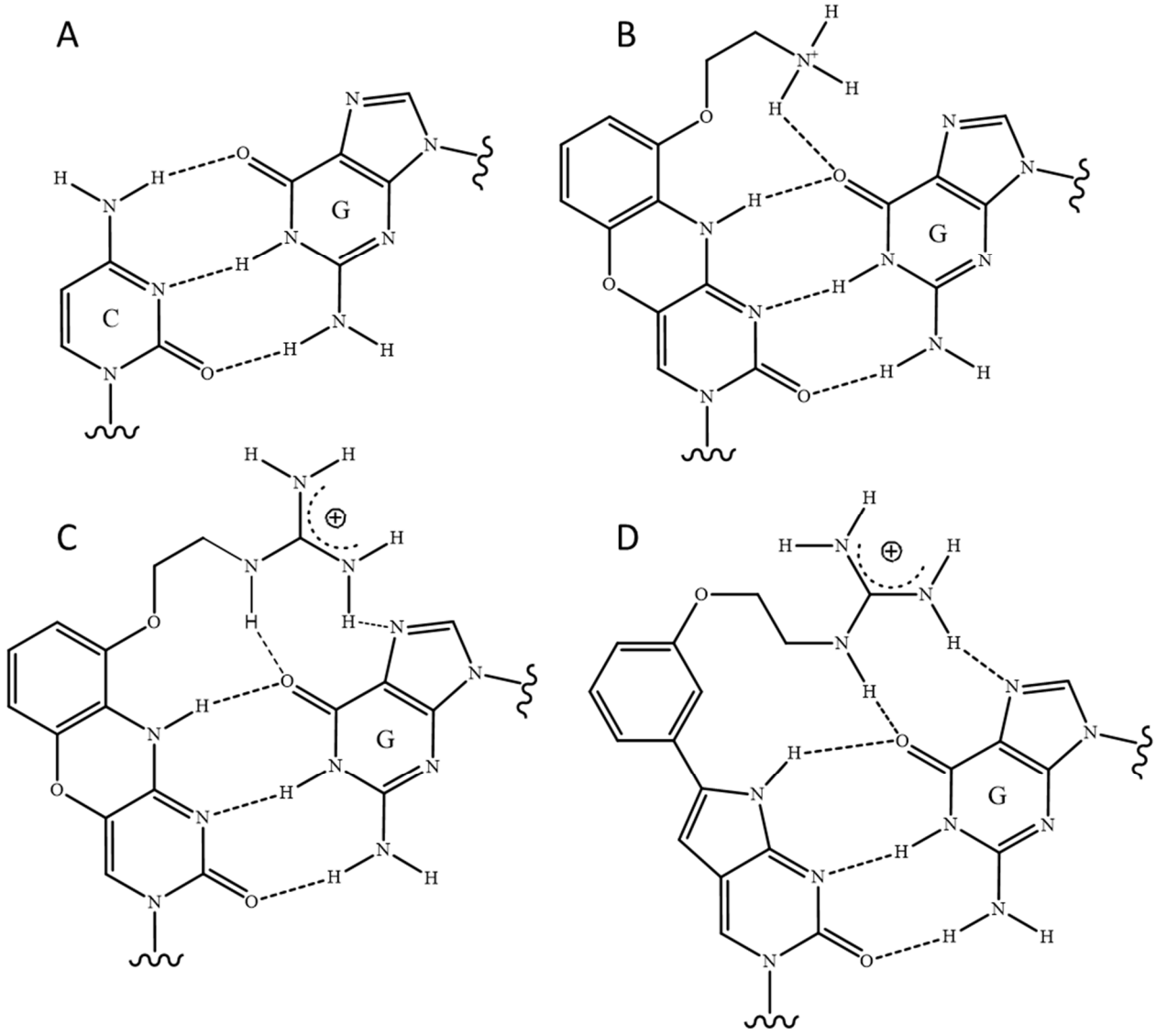


Figure 1.15: **A**) Standard cytosine-guanine base pairing; **B**) modified base presents an extended π-surface and an amino ethyl group able to interact with guanine; **C**) same base of B, but with guanidinium instead of amino group; this base is able to interact with Hoogsteen site; **D**) this base presents guanidinium unit like in case of C, but different polycyclic aromatic part.

Combination of modified backbone and base can lead to very effective systems and, for example, use of **C** with a γ-modified PNA allowed B-DNA double strand invasion.[128]

Uracil is a well-studied system because substituents in 5- position don't alter duplex stability, since the side chains are placed in major grove of duplexes. Proof of this are many works, reported in the literature, in which polymerase enzymes incorporate modified nucleotides.[129,130] For these reasons many modifications of uracil moiety have been proposed,[131] bearing in position 5 reporter groups, metallic centers, reactive groups, etc.. Some examples are reported in Figure 1.16.

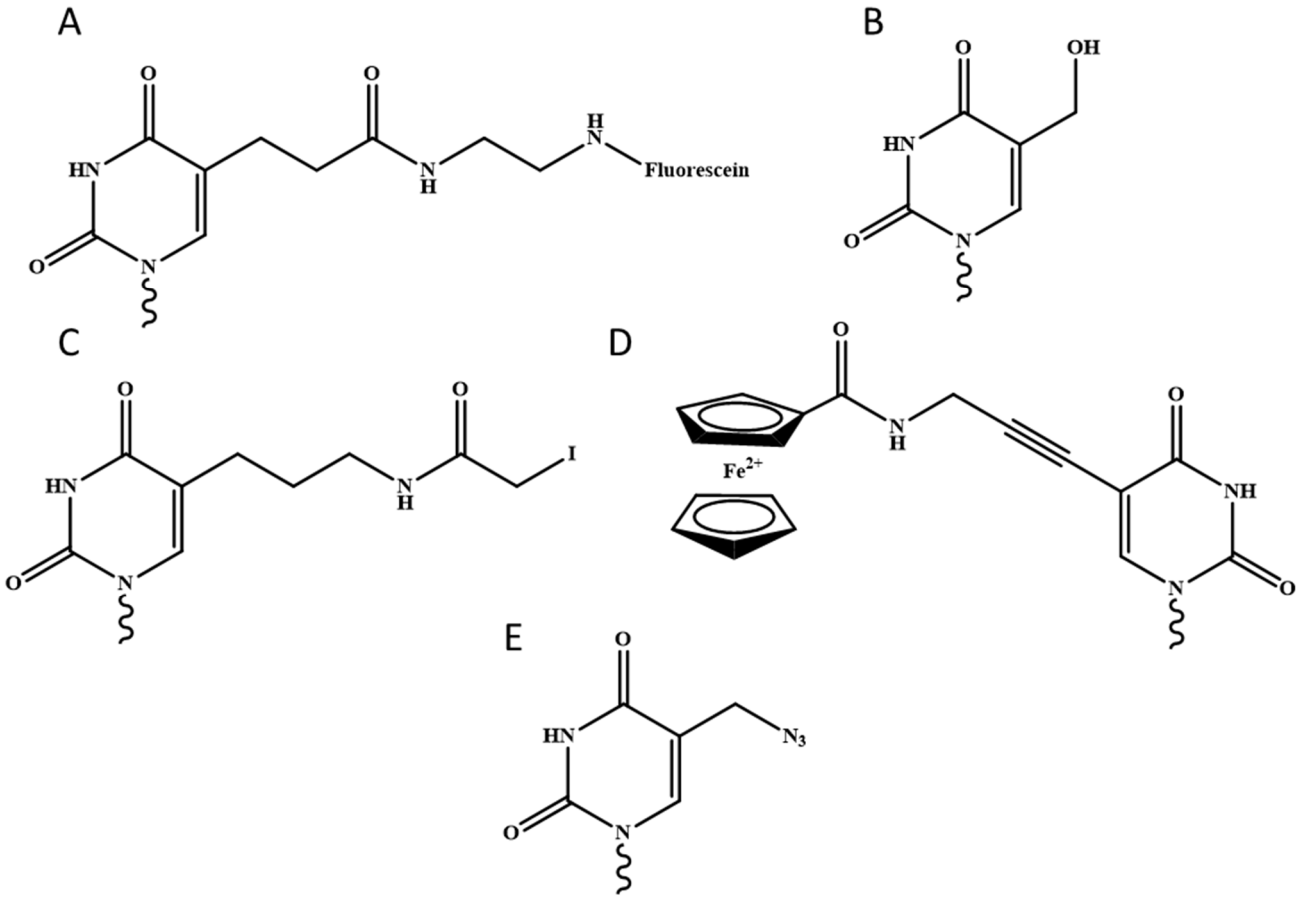


Figure 1.16: Examples of 5 modified uracil. **A**) uracil modified with reporter group; **B**) 5-hydroxymethyluracil **C**) uracil-based cross-linking agent; **D**) uracil coupled with ferrocene unit; **E**) 5-azidomethyluracil.

A is an example of system bearing a reporter group,[132] coupled through a linker, useful for diagnostic or analytical purposes. **B** presents a modification that increases uracil hydrophilicity, favoring PNA solubility.[133] **C** is an example of base able to make cross linking with DNA and in particular it makes a covalent bond only with guanine.[134] Sonogashira's reaction has been used to create a wide range of modified monomers starting from 5-iodouracil and this type of reactivity has been used also for the introduction of metal centers (**D**).[135] Finally **E** is a versatile monomer that allows to introduce a wide range of groups by means of click reactions or coupling after azide reduction.[136]

1.5. Design of new modified PNA

A widely used approach for design of new modified PNA is that based on rational design combined with "trial and error" method. Availability of crystal structures containing PNA[93,137–143] allowed to better focus on the conformational properties and to design modified PNA monomers more accurately. Several review papers report a complete overview of the backbone[110,144] and of the nucleobase-modified PNA structures[108]proposed in the literature. Examples, relevant for the thesis work will be discussed in detail.

An interesting case of modifications derived from crystal analysis was reported by Ganesh and co workers.[145,146] Standard aminoethylglycine (aeg)-PNA is conformationally flexible and can adapt to target DNA or RNA preferred helical conformation. A drawback of this adaptability is the high negative entropic change in duplex formation. To avoid this problem, many constrained systems have been proposed. From the observation of characteristics angles of PNA (Figure 1.17) in PNA:DNA and PNA:RNA duplex structures, changes in the β angle going from DNA to RNA heteroduplex and from duplex to triplex were identified. In particular in PNA:DNA duplex β is about 140° while in PNA:RNA or PNA_2DNA range from 65 to 70°.

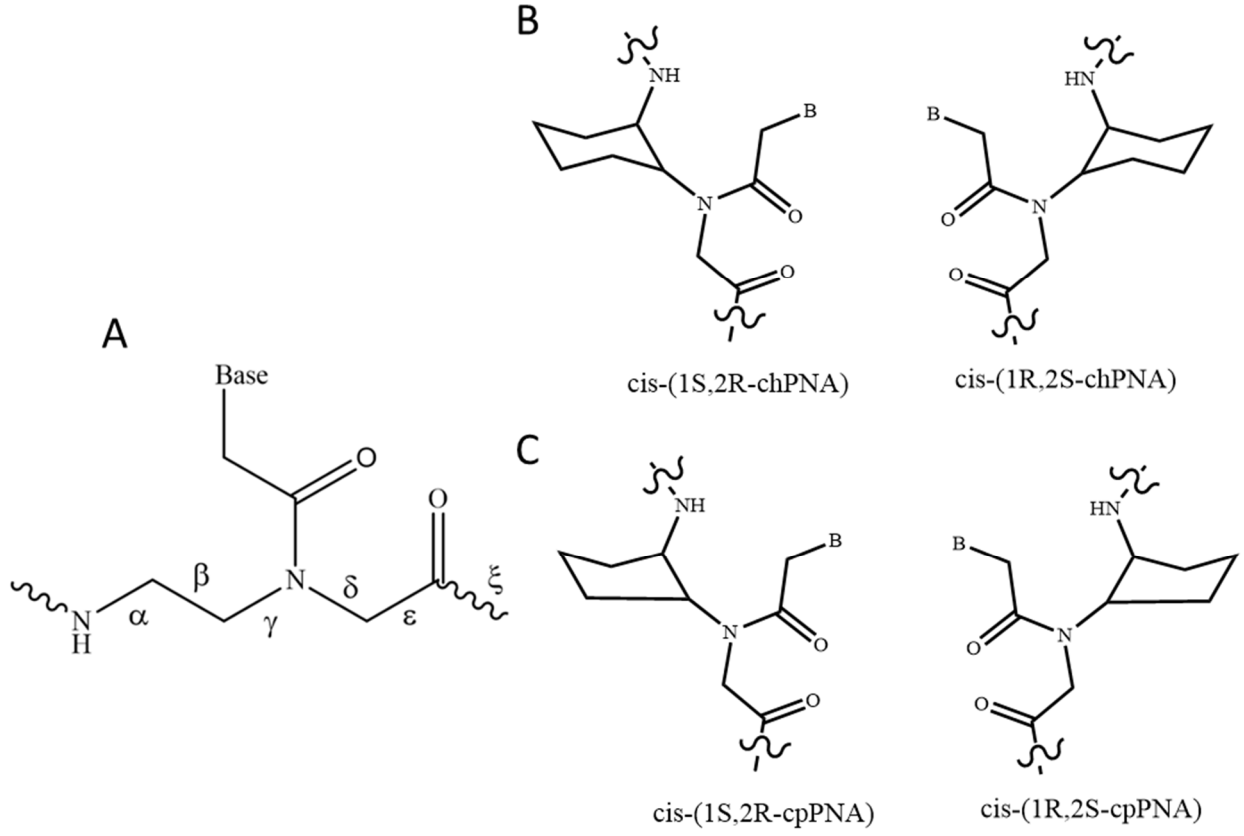


Figure 1.17: **A**) PNA characteristic angles; **B**) ch-PNAs structures; **C**) cp-PNA structures.

From this observation cyclohexanyl PNAs (chPNAs) were designed, by introducing an aminocylohexyl unit instead of the standard aminoethyl one. This modification reduced backbone mobility and thus conformational freedom of the duplex. Two different systems, cis-(a,e)(1S,2R/1R,2S) were considered; the former presented a beta angle equal to -63°, the latter equal to 66°. Study of stability revealed that ch-PNAs destabilized the formation of DNA:PNA duplex, while the PNA:RNA stability was not affected. In particular complexes with RS monomers resulted in a higher melting temperature compared to SR analogues, confirming the importance of the structural design based on β angle. Thus, although these systems were not particularly effective in binding, they allowed to discriminate between DNA and RNA, showing differences in melting temperatures up to 50 °C. Poor binding was probably a consequence of the rigidity of cylohexyl ring that remained trapped in either of the two chair conformations, limiting pairing during duplex formation. To resolve this drawback, Ganesh and coworkers proposed a cyclopentyl-modified backbone (cp-PNAs), in which endo-exo puckering allowed better adaptability. Effectively this led to an increased duplex stability for both DNA and RNA, but to the detriment of the selectivity shown in case of ch-PNAs.

Although this was an example of successful rational design of new backbone-modified PNA structures, this approach presents the drawback to use information obtained from crystal structures that are rigid models, which is only a partial set of all possible conformations present in solution, where most processes of interest occur. A new interesting approach, based on dynamic study of systems using Molecular Dynamics, will be shown in following chapters.

1.6. Design of anti-miR PNA: a challenge

Anti-miR PNAs have been used for modulating differentiation,[118] arrest cellular proliferation[147,148] and inducing apoptosis.[149] Detection of miRs in living cells has also been achieved by means of PNA probes.[150]

In miR nomenclature, “seed region” refers to the part of miR that goes from 2nd to 7th nucleotide of sequence. It has been demonstrated that targeting this region is sufficient to inhibit the activity of the miR without the necessity to have full sequence complementarity.[151] Binding short sequences, like micro-RNA, requires very effective systems and even more efficient ones if the objective is to target the seed region, leading to demand of new modified oligonucleotides or PNAs. The present

PhD thesis was developed in the context of a regional project, PNA NOVA, with the purpose of developing new effective PNAs for miR targeting.

PNAs used in anti-miR approach essentially act by reducing micro RNA availability in the cell through duplex formation, but thus requiring equimolecular quantity or an excess of PNA to observe an effect (usually micro molar scale). A possibility to overcome this problem is to insert catalytic moieties able to cleave RNA. This would permit PNA to bind RNA and, once duplex is formed, to cleave phosphodiester bond. Duplex with two RNA fragments is less stable than duplex with whole miR. In this way it would be possible for PNA to dissociate from the complex and then bind again RNA strand and to act in catalytic way (Figure 1.18) like a sequence selective artificial nuclease.[152,153]

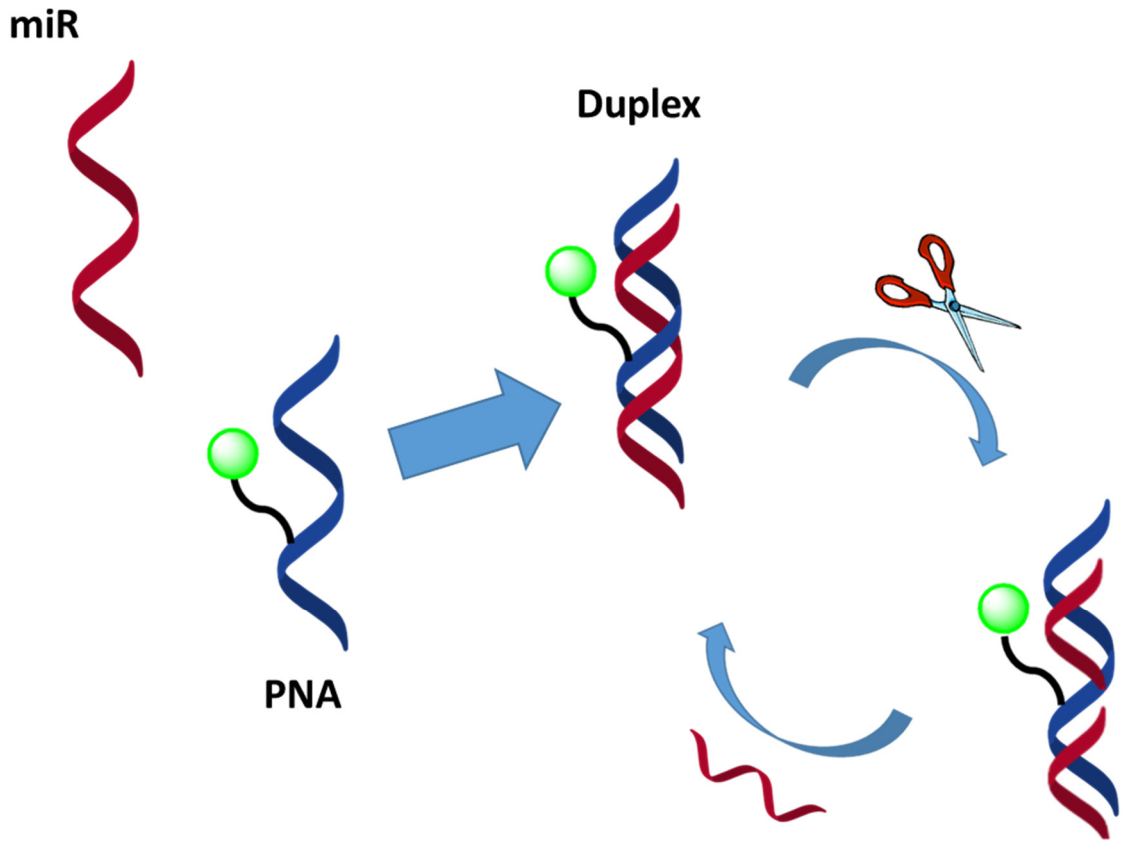


Figure 1.18: Schematic representation of activity of a PNA bearing a hydrolytic moiety.

A useful chemical toolbox for achieving this goal is provided by the extensive studies carried out in the field of artificial nucleases, and in particular of metallo-nucleases.

1.7. Artificial Metallonucleases

DNA must preserve genetic information and to ensure that its phosphodiester bond is quite resistant to hydrolysis. Actually, estimated half-life of DNA at 25 °C and pH 6.8 is of 30 millions of years.[154] RNA, which presents the role of transporting information, is more unstable, due to the possibility of intramolecular catalyzed hydrolysis and under the same conditions exhibits a half-life of "only" four years.[155] However, in biological media DNA and RNA can be rapidly degraded due to the presence of nucleases. These enzymes are able to greatly accelerate the hydrolysis reaction. Many types of enzyme can be found, but generally, their activity is ensured by presence of a metal center.

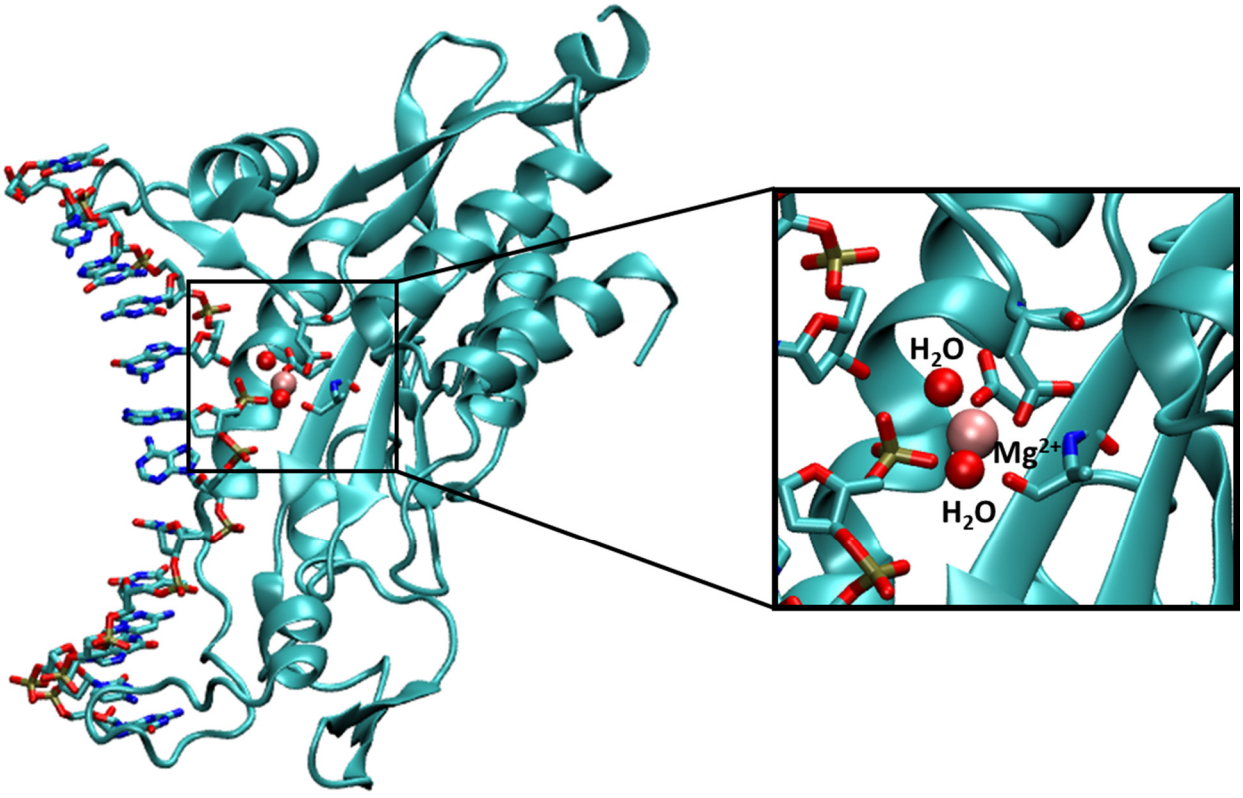


Figure 1.19: Crystal structure of EcoRI complex with DNA taken from PDB (1QPS)[156] and insight of catalytic center showing the role of the magnesium ion.

In Figure 1.19 an example of structure of a nuclease isolated from *Escherichia coli* is shown. Enzyme is able to bind DNA strand and to hydrolyze it through water molecules coordinated to Mg^{2+}. Many other examples of nucleases can be found in literature[157–159] and Mg, Mn, Fe and Zn are commonly

found in this type of enzymes. Several efforts have been spent trying to mimicking nucleases activity. There are two principal cleavage methods: oxidative and nucleophilic.

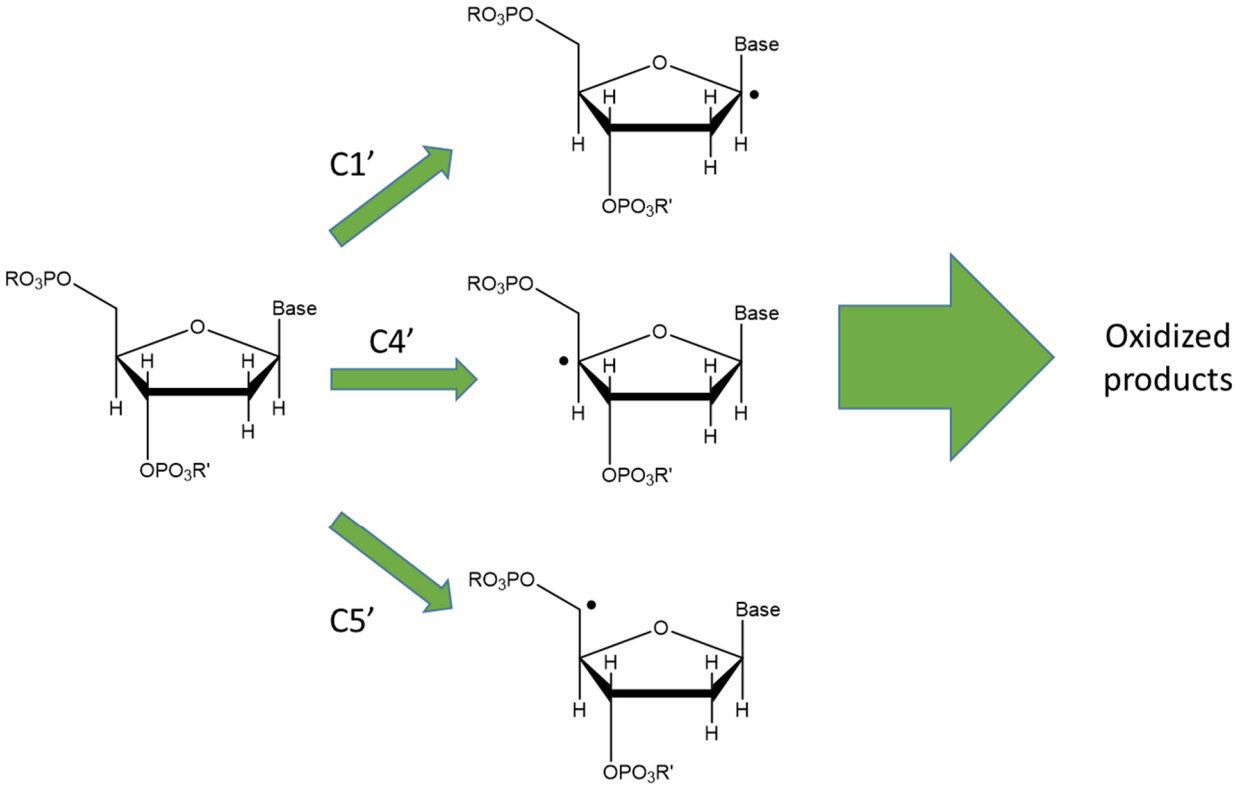


Figure 1.20: Schematic representation of oxidative cleavage pathway by hydrogen abstraction.

Oxidative cleavage (Figure 1.20) proceeds via different pathways, of which, the most effective is through hydrogen abstraction, that can occur in different positions, mainly in C1', C4' and C5'. After radical formation, DNA can be oxidized to obtain different products.[160]

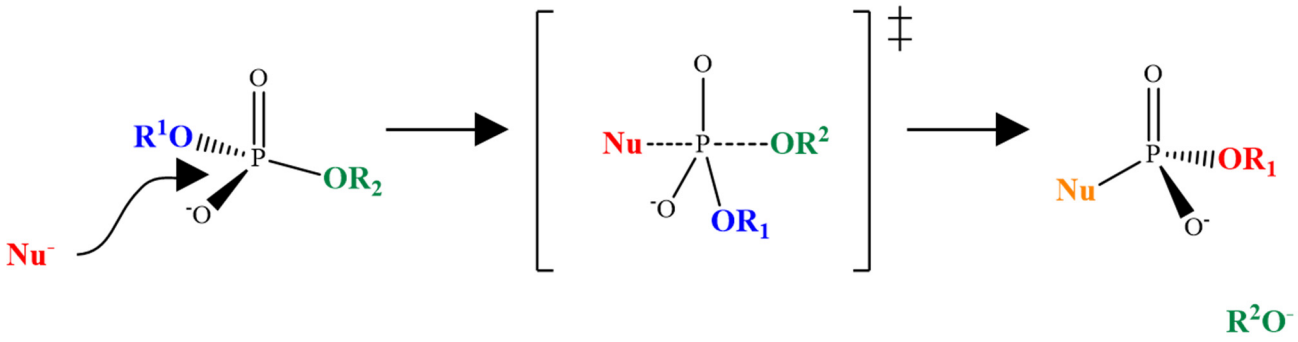


Figure 1.21: Schematic representation of hydrolytic cleavage pathway.

At contrary, nucleophilic cleavage proceeds through attack of a nucleophile on phosphorus, followed by scission of phosphodiester via pentacoordinate transition state (Figure 1.21). Nucleases acts following this pathway and nucleophile is usually a coordinated water molecule or a side chain of an amino acid residue. Many examples of metallic complexes, exhibiting hydrolytic effects toward DNA or RNA, have been reported in the literature.[161] This type of compounds are called artificial metallonucleases. Most effective complexes contained cobalt,[162] nickel[163] and cerium. In particular complex of Ce(IV)/EDTA[164,165] resulted very effective in DNA single strand hydrolysis and have also been used in combination with PNAs:[166] PNA conjugated with NLS peptide was used to perform DNA double strand invasion and at that point was possible hydrolysis by Ce(IV)-EDTA complex. Drawback of using these metals is their cytotoxicity if released from ligands. Iron would be a perfect candidate, since is not toxic, however complexes with Fe(II) or Fe(III) resulted scarcely active or needed basic conditions to perform cleavage.[167] Interest of scientific community has been focused toward two alternatives: copper and zinc. Toxicity is higher compared to iron, but limited metal release is tolerated by biological systems, in particular in case of zinc. Complexes with these metals showed a sufficient activity due to coordinated water molecules with pKa within the range of 7-9. Many examples of complexes with this metals have been tested as nucleases.[168] Artificial nucleases can be mononuclear complexes,[169–171] with only one metal center, or multinuclear complexes (Figure 1.22).[172,173] Generally complexes with copper were more active than those with zinc, and di- or multi-nuclear complexes exhibited better properties than mononuclear counterparts. Molecules bearing pendant nucleophiles have been synthetized (Figure 1.22): this groups, usually alcohols, react in place of coordinated water molecule with a higher hydrolysis rate, but intermediate phosphate is stable and it is not possible to achieve turnover.[174,175]

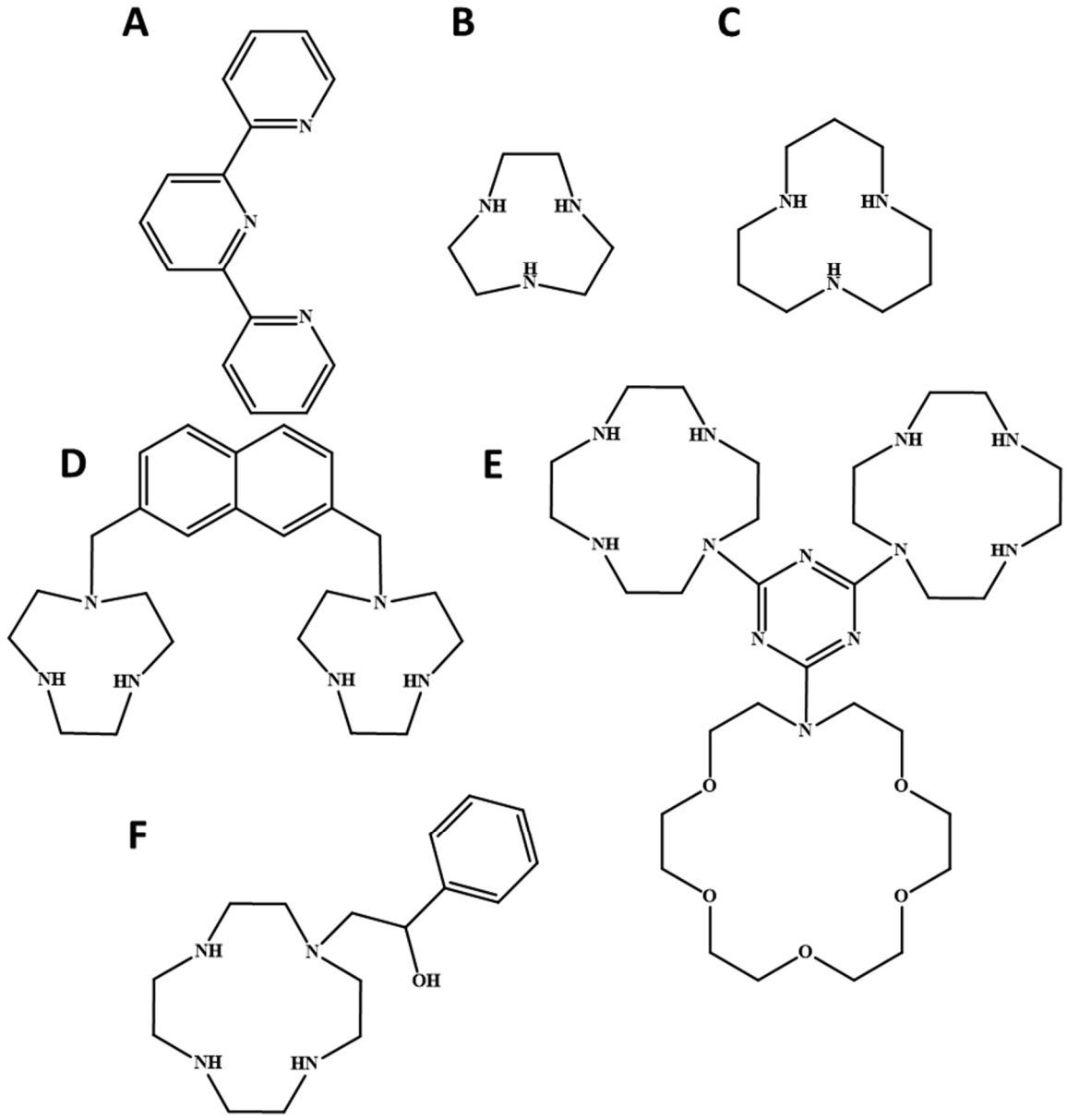


Figure 1.22: Examples of mononuclear and multinuclear metal ligands used as artificial metallonucleases. **A**) terpyridine; **B**) 1,4,7-triazacyclononane; **C**) 1,5,9-triazacyclododecane; **D**) dinuclear triazacyclononane derivative; **E**) trinuclear 1,4,7,10-tetraazacyclododecane derivative; **F**) 1,4,7,10-tetraazacyclododecane bearing a pedant alcohol.

Scrimin and coworkers reported an interesting example of polypeptide bearing two Zn^{2+} triazacyclononane hydrolytic units.[176] Folding of peptide is necessary to bring two catalytic rings sufficiently close together in order to enhance phosphodiester bond (Figure 1.23). Another artificial nuclease was developed by the same group using gold nanoparticles functionalized with hydrolytic moieties.[177] On particle surface, these units are sufficiently close to act cooperatively increasing hydrolysis rate.

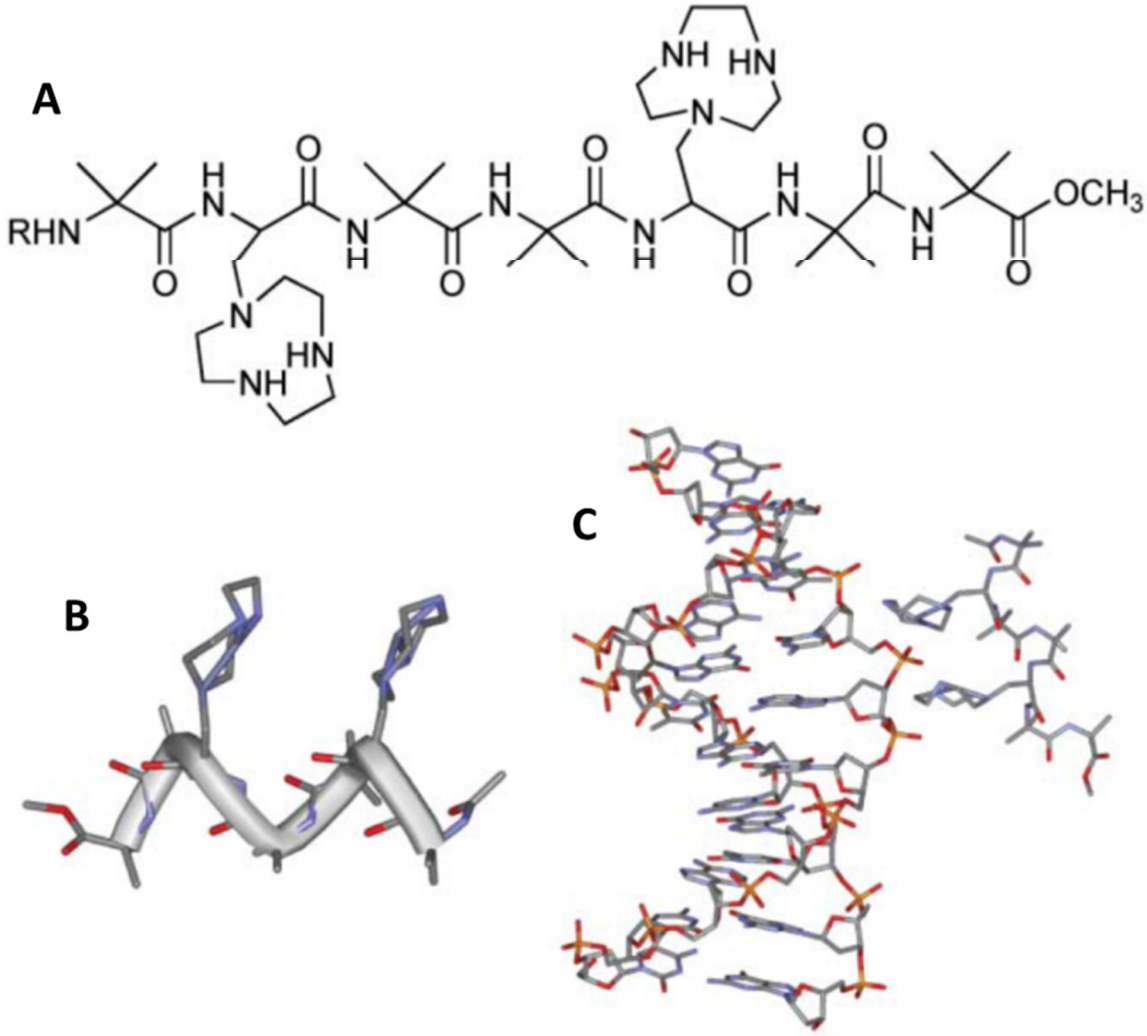


Figure 1.23: Modified Peptide bearing two hydrolytic units. **A**) Chemical structure; **B**) secondary structure; **C**) proposed mode of binding to DNA. Reprinted with the permission of Royal Society of Chemistry.[178]

During PhD thesis synthesis of a mononuclear ligand was performed in order to insert a hydrolytic unit in PNA strand. Discussion of PNA-based artificial nuclease will be reported in more details in Chapter 6.

2. Methodology background: Molecular Dynamics and Metadynamics

2.1. Simulations of large molecular systems

Theoretical basis for simulation of molecular properties have been developed since the very beginning of computer era.[179] However, only in recent years exponential growth of computational power permitted to widely expand the use of simulations to solve complex problems. *Ab initio* methods describe the motion of electrons in atoms and molecules using wave functions and avoiding any experimental data. In this way, it is possible to compute quantum mechanical properties of molecular systems at a high level of accuracy, but lacking in efficiency due to the usually high computational costs. For this reason, *ab initio* simulations are generally restricted to small molecules/systems with less than 100-1000 atoms. However, medicine, biology and biotechnology challenges require tools suitable for much bigger and complex systems such as proteins, enzymes, receptors, membranes, etc. with hundreds of thousands of atoms. Classical Molecular Dynamics (MD) is an alternative approach which can be used to address this type of problems; in MD atoms are described as spheres, represented by their Van der Waals radii, connected with rigid bonds and subjected to classical physics laws. This approach simplifies several calculation steps avoiding the resolution of large Hamiltonian matrixes as it happens in *ab initio* MD, thus neglecting quantum mechanical effects. MD techniques resulted in a successful approach for simulating large molecular systems, allowing to achieve precious insights in the molecular structures and motions. Many years have passed from the first simulation of 9 picoseconds, made by Karplus on bovine pancreatic trypsin inhibitor (BPTI),[180] and nowadays it is possible to reach simulation times ranging from tens to thousands of nanoseconds, thus approaching reliable simulated measurable properties. In this chapter we will describe the methodologies of the work carried out during a training period on computational tools spent at the Department of Chemistry and Applied Biosciences ETH Zurich, USI Campus, Lugano, in the research group of Prof. Michele Parrinello, one of the pioneers and most prominent scientist in the field of Molecular Dynamics of large systems, including Biosystems, under the supervision of Dr. Vincenzo Verdolino.

2.2. Introduction to Molecular Dynamics

Use of classic physic principles, neglecting quantum effects, allows to simplify energy calculation that is represented by the sum of kinetic (K) and potential (V) energy (**2.1**):

$$\boldsymbol{E}\,(\boldsymbol{r},\boldsymbol{p}) = \boldsymbol{K}(\boldsymbol{p}) + \boldsymbol{V}(\boldsymbol{r}) \qquad (2.1)$$

where $\mathbf{r}$ and $\mathbf{p}$ are the coordinate and momentum vectors, respectively. Kinetic energy has usually a very simple form (**2.2**) easy to calculate, while potential is a more complex function.

$$\mathrm{K} = \sum_{i=1}^{N} \frac{1}{2\boldsymbol{m}_i}(\boldsymbol{p}_{ix}^2 + \boldsymbol{p}_{iy}^2 + \boldsymbol{p}_{iz}^2) \qquad (2.2)$$

where $\mathbf{m_i}$ is the mass of $\mathbf{i^{th}}$ particle and $\mathbf{p_{ix}}$, $\mathbf{p_{iy}}$, $\mathbf{p_{iz}}$ are the vector components of $\mathbf{p_i}$.

In MD there are different way to express the potential $\mathbf{V(r)}$, but in any case it comes from the sum of different terms, defined as bonded ($\mathbf{V_{bond}}$, $\mathbf{V_{angle}}$, $\mathbf{V_{dihedral}}$) and non-bonded ($\mathbf{V_{improper}}$, $\mathbf{V_{VdW}}$, $\mathbf{V_{elec}}$).[181]

$$\boldsymbol{V} = \boldsymbol{V}_{bond} + \boldsymbol{V}_{angle} + \boldsymbol{V}_{dihedral} + \boldsymbol{V}_{improper} + \boldsymbol{V}_{Vdw} + \boldsymbol{V}_{elec} \qquad (2.3)$$

$$\boldsymbol{V}_{bond} = \boldsymbol{K}_b(\boldsymbol{b} - \boldsymbol{b}_0)^2 \qquad (2.4)$$

$$\boldsymbol{V}_{angle} = \boldsymbol{K}_\theta(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^2 \qquad (2.5)$$

$$\boldsymbol{V}_{dihedral} = \boldsymbol{K}_\chi[1 + \boldsymbol{cos}(\boldsymbol{n\chi} - \boldsymbol{\delta})] \qquad (2.6)$$

$$\boldsymbol{V}_{improper} = \boldsymbol{K}_\psi(\psi - \psi_0)^2 \qquad (2.7)$$

$$\boldsymbol{V}_{VdW} = \boldsymbol{\varepsilon}\left[\left(\frac{R_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{ij}}{r_{ij}}\right)^{6}\right] \qquad (2.8)$$

$$\boldsymbol{V}_{elec} = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \qquad (2.9)$$

where K_b is the bond force constant and $b - b_0$ is the distance from equilibrium position. K_θ is the angle force constant and $\theta - \theta_0$ is the angle from equilibrium between 3 bonded atoms. K_χ is the dihedral force constant, n is the multiplicity of the function, χ is the dihedral angle and δ is the phase shift. K_ψ is the force constant and $\psi - \psi_0$ is the out of plane angle. ε is the depth of the potential well, R_{ij} is the distance in which potential reach the minimum, and r_{ij} the distance between atoms. q_i and q_j are atoms charges, ε_0 is the dielectric constant and r_{ij} is the same distance of V_{WdW}. Bonds, angles and impropers terms are expressed in the form of harmonic potential (**2.4**, **2.5**, **2.7**). Usually spring constants assume quite high value, thus inducing a great energy variation for little displacements from equilibrium position. Dihedral potentials are conversely softer, allowing broader ranges and assuring necessary flexibility to the molecule (**2.6**). Electrostatic contribute is expressed in the familiar Coulomb equation (**2.8**) and dispersive interactions (Van der Waals) are regulated by Lennard-Jones potential (**2.9**). These two terms constitute long-range interactions and are the most expensive in terms of computational effort. In order to accelerate MD calculations a cutoff value for VdW interactions is usually set to 10 Å since this contribution decays as r^{-6} and at long distances is almost zero. For Coulombic potential is not possible to adopt the same cutoff since it decays as r^{-1} and the error generated would be fairly high. A common technique to this problem is the use of Particle Mesh Ewald (PME) method that consists in introducing a neutralizing charge distribution for every point charge of the molecule and subtracting this contribution to the total electrostatic potential.[182] This term decays faster and is then possible to apply a cutoff similar to that one used for dispersive forces. Use of PME protocol requires also application of periodic boundary conditions (PBC)[183] that means molecular systems replicated in the three spatial directions, forming and infinite *lattice*. This approach ensures that all particles leaving the simulation box from any side are re-entered from its periodic image (Figure 2.1). PBC box allows also to avoid surfaces effects[181] due to finite size systems and is therefore commonly used in Molecular Dynamics. All potential contribution described above must be defined for every molecule to simulate. Different simulation programs have been developed together with their own Force Field (FF) syntax and parameters attribution. Most used programs are GROMACS,[185] AMBER,[186] CHARMM[187] and NAMD.[188] Moreover, for common types of molecules and scaffolds, many different Force Fields are available for users avoiding manual parameter definition. In early versions, each different program used its parameters, leading to development of large variety of Force Fields for the same type of molecules. Nowadays, most of the programs share the code giving to users the possibility to choice; for example GROMACS allows to select it's proper FFs (GROMOS)[189] or among FFs developed for

other programs (AMBER,[190] CHARMM,[191] OPLS[192]). Since proteins are the most studied systems, in the past years Force Fields in this area were greatly developed.[193–195]

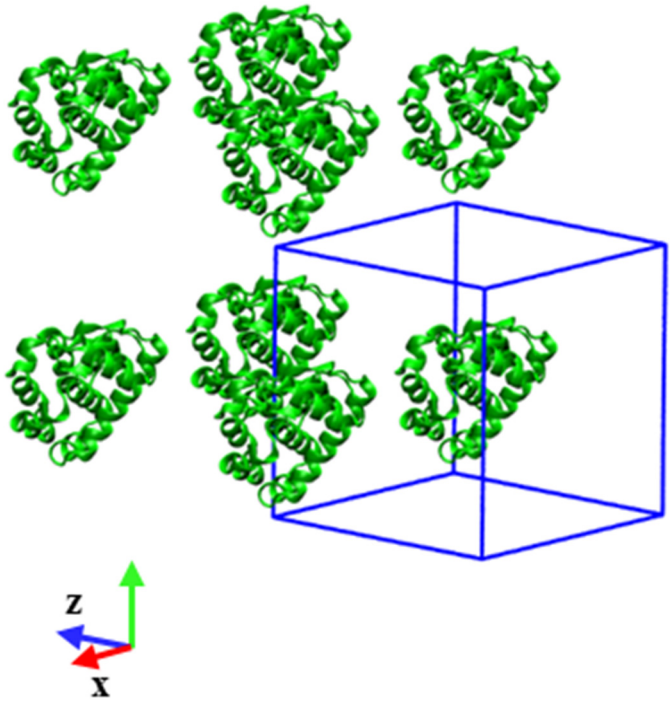


Figure 2.1: Representation of Periodic Boundary Conditions: simulation box is represented in blue (60.0 X 60 X 60 Å^3), original protein structure (3HTB)[184] is inside the box and outside +x, +y, +z, replicates are shown.

Nowadays, FFs for organic small molecules start to be expanded for including lipids,[196] sugars,[197] nucleic acids,[198] etc. The choice of the best performing FF is not trivial and usually is based on the experience of the user and on the literature. Compared to proteins, force fields for nucleic acids are barely explored and continuous improvements are in the process. Most recently used FFs for DNA and RNA are Amber *ff99sb*[199] improved with *parmbsc0*,[200] CHARMM27,[191,201] GROMOS53a6[189] and OPLS-AA,[202] but the first two gave the best results in reproducing the experimental data.[198,203,204] Generalized Amber Force Field (*GAFF*)[205] and CHARMM General Force Field (*CGenFF*)[206] are also popular FFs since they provide parameters for small organic molecules that can be applied to ligands or modifications.

Once the best choice of the FF is set along with all the simulation parameters, it is possible to start the time evolution of the molecular systems. This will follow trajectories that satisfy the solution of Newtonian motion equations under the field of forces selected. This achievement bare mechanistic, structural, thermodynamics and kinetic features that can be post processed. In particular, it is of great interest the potential or the Free Energy Surface (FES) describing the system under study. One effective approach is to define variables or combination of variables called Collective Variables (CV) that based on the experience or the physics of the system can optimally describe one or more

motions of interest. Moreover, the selected CVs have to univocally describe the motions under examination in order to better discriminate between different system states.

Another critical aspect of MD simulations is the conformational space exploration. In order to extract Free Energy Surface information out of the simulated trajectories, it is fundamental that the system under study spent a significant amount of time in the conformation of interest. This criterion that could appear as an arbitrary estimation can be quantitatively evaluated through the free energy surface (FES) convergence analysis. Indeed, if the simulation time is sufficiently long, the molecular system will have reached the equilibrium state. With this, the probability for the system to be in a given configuration is calculated by means of the time spent in that conformation and, most remarkably will be converged onto the free energy surface. Employing the collective variables formalism the FES can be expressed in terms of CV-dependent probabilities through the following equation (**2.10**):

$$\boldsymbol{F(s) = -k_B T ln(P(s))} \qquad \textit{(2.10)}$$

where $\mathbf{F}$ is the free energy, $\mathbf{P}$ is probability function of $\mathbf{s}$, $\mathbf{k_B}$ is Boltzmann constant and $\mathbf{T}$ temperature.

However, the equilibrium condition is difficult to practically satisfy and it would require infinite simulation time. With finite simulations, such approximation is particularly dangerous when considering systems with high degrees of freedom or for those where state transitions are relatively slow. The latter situation is common, for example, in proteins[207–209] in which passage from two different states may imply huge conformational rearrangements. In both cases, the restricted conformational sampling allowed by classical MD can bring to erroneous conclusions. To overcome to these problems, enhanced sampling methods[210–213] must be used. In particular, hereafter we describe the Metadynamics methodology.[214,215]

2.3. Principles of Metadynamics

Metadynamics has been defined as “filling the free energy wells with computational sand”.[216]

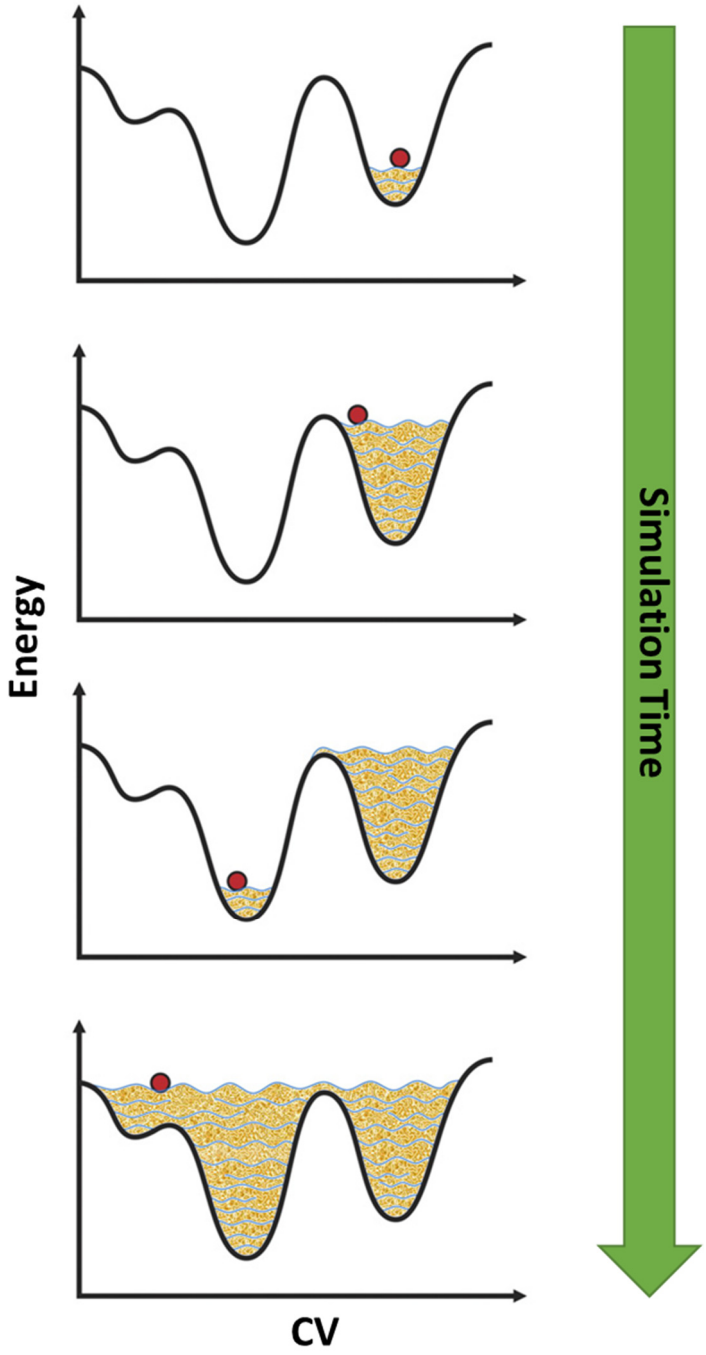


Figure 2.2: Graphical representation of Metadynamics.[217] Conformational space is expressed as function of a generic Collective Variable, which describes the system, and energy. Position of the system in this space is represented with a red circle. Applied potential is here represented as sand, which accumulates over time.

This statement simplistically represents the operative procedure employed by the Metadynamics in which simulated system is discouraged to visit previously populated states by means of a history-

dependent bias potential.[214] Figure 2.2 describes an example of energy profile as function of a generic CV in which there are two local minima and a global minimum. With a red circle we identify the *position* of the system under study at a given simulation time. We define this *position* in terms of Collective Variables that optimally describe the system conformational space. In this example, the starting position of the system is set in a local minimum. In order to escape from this basin a significantly tall energy barrier has to be crossed. The probability of populating other minima on the Free Energy Surface depends on the thermal contribution defined by the energy k_BT. Usually, standard MD at room temperature rarely allows for overcoming such tall energy barriers and the system will permanently oscillate in the starting basin only for temperature-induced fluctuation. Metadynamics adds, at regular user-defined intervals, a bias potential in terms of so called *hills* that help the system to populate higher energy configurations and eventually to fall in other local minima. Through this procedure, hills accumulate during the simulation "flooding" the free energy surface and allowing all possible minima in the conformational space.
The time-dependent potential assumes the form expressed in (**2.11**).

$$\boldsymbol{V(s(t)) = \sum_{t' \leq t} we^{\left(-\sum \frac{(s_i(t) - s_{i,t'})^2}{2\sigma_i^2}\right)}} \qquad \textit{(2.11)}$$

where $\mathbf{w}$ is the height and $\boldsymbol{\sigma}$ the width of a single hill, $\mathbf{s(t)}$ is a vector of CVs at time $\mathbf{t}$ and $\mathbf{s_{i,t'}}$ is the bias potential deposited by the hills at $\mathbf{t' < t}$, keeping the "bias-history" of the simulation. Recovering the visited states, it is thus possible to reconstruct the FES as the sum of the bias potential (**2.12**):

$$\boldsymbol{F(s) = -\sum_{t' \leq t} we^{\left(-\sum \frac{(s_i(t) - s_{i,t'})^2}{2\sigma_i^2}\right)}} \qquad \textit{(2.12)}$$

Reliable FES reconstruction requires an accurate choice of the collective variables. CVs should satisfy the following four characteristics to be used as system descriptors.[218]

- First of all, CVs must be simple or complex functions of atomic coordinates derived from trajectories collected during MD simulations ;

- Second, the function s(r) has to be continuous in the entire configurational space sampled. This condition is partially satisfied by the existence of ds(r)/dr which are the forces acting on each atoms at a given time and responsible for atomic motions. Failure in satisfaction of this condition causes crash of simulation;

These two conditions are of general validity and requested for all kind of MD simulations.

- Third feature: any efficient CV has to univocally discriminate the different states characterizing the system under study. Ideally, at a given value or combinations of CVs it one single system state or molecular structures within a certain confidential interval should be defined. This is a quite arbitrary evaluation that often gives rise to erroneous conclusions. In particular, for complex CV definitions, this condition can hardly be achieved and dedicated discussion should be derived. Nonetheless, if such condition is not satisfied, the statistics collected during the MD run can still be post processed by means of reweighting techniques that allows recovering the unbiased probability distribution of any variable from Metadynamics simulation.[219]

- Fourth, CVs should represent the "slow motion coordinates" of the system. In this way, "fast coordinates" follow a standard unbiased molecular dynamics whereas the time demanding motions will be facilitated by the Metadynamics. For example, if we consider the ligand binding to a protein it might be the case where the hot-site under study becomes available upon slow conformational rearrangement involving opening and closing of peptides functionalities. If that is the case, in order to accurately evaluate the binding energy, it is necessary to collect a sufficiently large number of statistics involving both processes (i.e. pocket opening/closing). Being this process the rate determining of the simulation it turns out that the optimal CVs have to be evaluated accordingly.

2.4. Well-Tempered Metadynamics

Classical Metadynamics employs constant height of the hills $\mathbf{w}$ along the whole simulation. This condition usually brings to overestimate the free energies and leaves to the user experience an arbitrary criterion for the statistic accumulation interruption. In particular, high hills allow to rapidly flood the energy surface, accelerating the MD, but resulting in rough FES estimation. On the contrary, low hills lead to accurate FESs requiring long simulation time. Another problem is that once a minimum is filled, the applied bias remains fixed, causing unrealistic sampling and in some cases preventing the observation of slow processes. An improvement of Metadynamics is the Well-Tempered Metadynamics (WT-MTMD).[220] This technique consists in applying variable height of the hills regarding the conformational space already visited. This algorithm, ensures that the hills deposition at a given value of the CVs is history dependent allowing for initial rapid filling of the local minimum and decreasing after several exploration of the same state. In this scheme, the molecular system will eventually explore all the free energy landscape avoiding a local overestimation of the energy barriers involved. Indeed, it is possible to demonstrate that Well-Tempered Metadynamics converges to an exact FES.[221]

The height of the hills is expressed in its time dependent formulation by the equation (**2.13**).

$$W = \omega e^{-\left[\frac{V_b(s,t)}{\Delta T}\right]} \tau_G \qquad (2.13)$$

where ω is the bias deposition rate, $V_b(s,t)$ is the potential at the current value of the CV for a given time step, and τ_G is the deposition time step.[222]

Metadynamics has been successfully used to simulate complex systems such as protein,[223] folding,[224] flexible docking,[225,226] nucleation[227] and mechanism of reaction pathways.[228] For this reason, we have selected this advanced technique for simulations of PNA-based systems.

2.5. Procedures for PNA simulations

All the MD simulations reported in this thesis were carried out with GROMACS 4.5.5,[229] patched with PLUMED 1.3[230] and PLUMED 2.0,[231]open source software, employed for the Metadynamics calculations.

Structure 176D,[143] taken from Protein Data Bank, was chosen as starting structure for the computational studies on unmodified PNAs (sections 3.2, 3.3, 3.5). For simulations where RNA moieties were present, we employed *ff99SB* force field improved with *parmbsc0*, since these force fields have been extensively used[232,233] leading to reliable and reproducible results.[203] PNAs are not parameterized in the available FFs and we retrieved missing parameters from R.E.DD.B. Server.[234,235] Details on the parameters used are reported in Appendix 8.1.

γ-modified PNA structures (sections 3.2, 3.4) were generated by modification of structure 176D. A force field for gamma serine modified PNAs was not available in the literature and its derivation will be discussed in section 2.6.

For the study of PNA containing modified bases, it was necessary to generate duplexes with sequences of interest besides the modification. Manual modification of duplexes is time consuming and can lead to unreal structural deformations of the system. The Nucleic Acid Builder (NAB)[236] module of AmberTools[186] is a module that allows to automatically generate DNA:DNA or RNA:RNA duplexes allowing to study of nucleic acids with desired sequence. To overcome problems described above we modified the code in order to automatically generate PNA:DNA and PNA:RNA duplex structures (see section 2.7). Charges and parameters derivation procedure are reported in section 2.8.

All systems were solvated in cubic boxes allowing at least 15 Å from the edges (i.e. 30 Å from the periodic boundary image) and filled by means of TIP3P water molecules. Topology and coordinate files were generated using t-leap module of AmberTools[237] and converted to GROMACS input files using Acpype script.[238] All the simulated systems, after a brief energy minimization procedure, were equilibrated at 300K (or differently where indicated) and 1 atm for 2 ns in the NPT (i.e. isothermal–isobaric) ensemble, employing the Berendsen barostat.[239] The production runs were generated in the NVT (i.e. canonical: isochoric-isothermal) ensemble, by means of the v-rescale thermostat.[240] The non-bonded cutoff scheme was set equal to 10 Å and the PME algorithm was employed for the electrostatics contribution.[182] The simulations were performed with a 2.0 femtosecond time step,

constraining all the bonds with the SHAKE algorithm.[241] Metadynamics simulations were conducted under the same conditions, applying the opportune bias to the CVs. Details are reported in the next chapter.

2.6. Parameters for γ-Modified PNAs

Simulating gamma modified PNA requires determination of charges and parameters for the force field. Restrained electrostatic potential (RESP)[242] is recognized as one of the standard in order to assign atom charges for MD simulations. It consists in a penalty-based fitting approach of electrostatic potential (ESP) calculated with *ab initio* methods. Use of different basis sets for calculation may results in different charges values, but using 6-31G* or larger basis set it has been demonstrated that results converge.[243] Therefore, this is the commonly used basis set because is a good compromise between accuracy and computational costs. Charges are generated fitting ESP by means of shell of points (with a density of 1 point/Å^2) at 1.4, 1.6, 1.8, 2.0 times the Van der Waals radius of atoms. This charge derivation allows to obtain reliable charge values, sampling "buried atoms" accurately and with reduced dependency to conformation. Better results are nowadays obtained increasing number of shells and of points; in the present work we have used 10 shells for each atom and a density of 17 points/Å^2. Charge derivation was carried out using the RESP ESP charge Derive (R.E.D.) Server[244] and following the protocol developed by Cornell *et al*.[245] In particular, to build a complete residue library, for each base there are three types of fragments: N-terminus, central and C-termins (Figure2.3). The N-terminus is the first monomer of the sequence (i.e. with free amino group) and C-terminus is the last monomer of the sequence (i.e. with free carboxylic group); into PNA chain, a generic monomer **n** is attached to amino group of residue **n-1** and to carboxylic group of residue **n+1** (central fragments). In order to mimic the peptide linkage between residues and to obtain reliable charges, their N-end tail is acetylated (ACE) whereas the C-end is capped with methyl amide (NMA). Structures were optimized with Gaussian 09 D.01[246] using DFT B3LYP level of theory and 6-31G* basis set. RESP charges were then calculated using R.E.D. program[244] implemented in R.E.D. server[247]. Schematic representation of charge derivation procedure is represented in Figure2.3. Concerning central fragments, RESP charges were obtained imposing neutrality condition (charge equal to 0) after capping groups removal (NMA and ACE). RESP charges on N-terminus were derived from capped monomers and methyl ammonium (NMA^+) imposing total charge equal to +1 after capping groups removal and joining of $\mathbf{NH_3^+}$ to the monomer.

Similarly, C-terminus is derived from capped monomer and acetate (ACE^-) imposing total charge of -1 after capping groups removal and joining of **COO^-** to the monomer. Moreover, charge fitting was performed imposing rotation of the bases along the linker. This last, homogenizes the fitted charge values as a function of base rotation, leading to more consistent and reliable RESP values.[234] Finally, atoms of the backbone and of the linker were forced to have the same charges except for C8' and relative hydrogens, allowing charge redistribution with different bases bonded. This procedure, the same used for unmodified PNA, permitted to construct a complete and consistent library of gamma modified PNA monomers. Details about R.E.D. usage are reported in the program site.[248] Structures and charges are reported in Appendix 8.2.

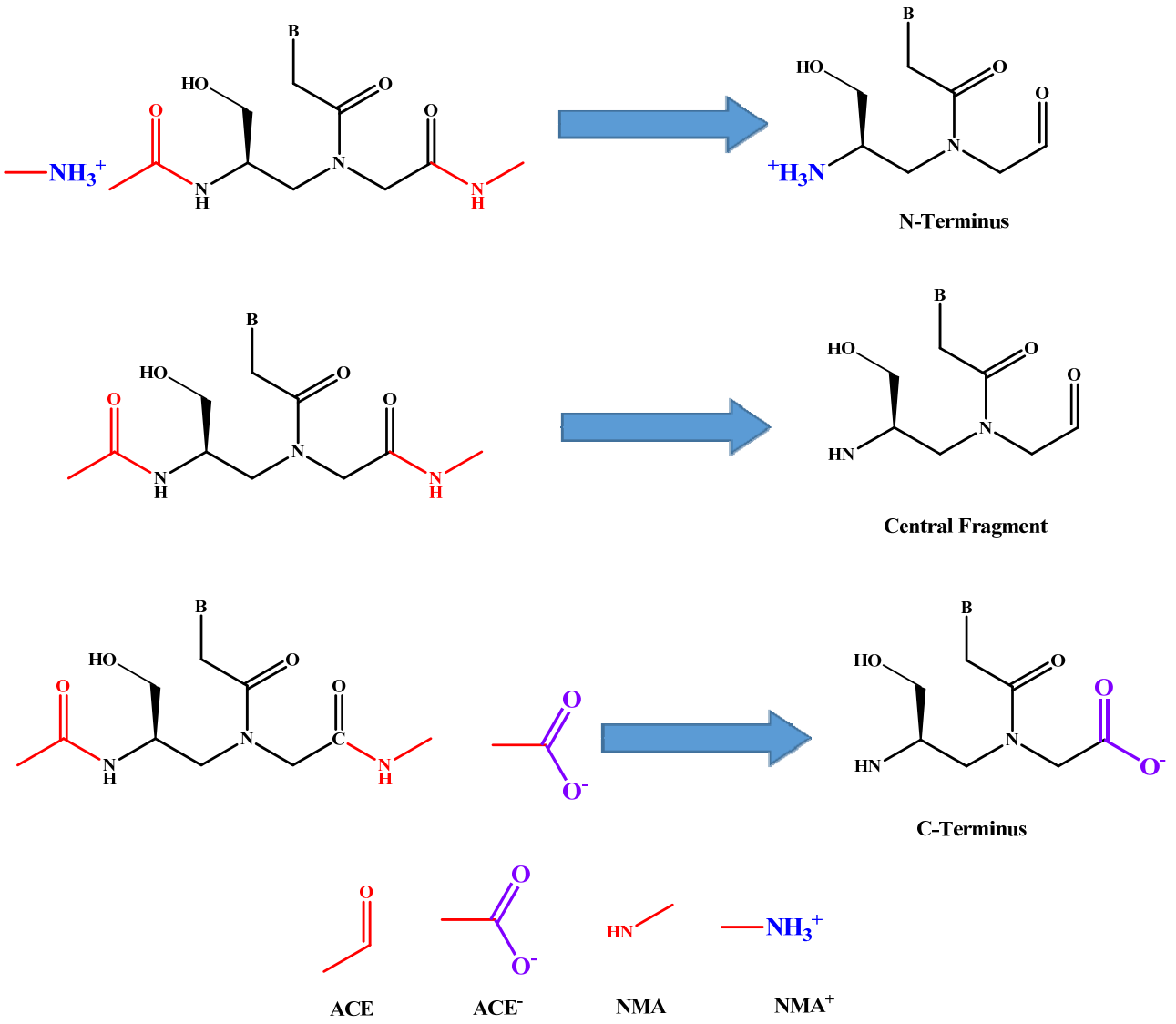


Figure2.3: Schematic representation of RESP derivation procedure. In red are reported groups that are removed to obtain final fragments. In blue NH_3^+ that is joined to backbone to generate N fragment and in violet COO^- that is joined to backbone to obtain C terminus.

Since the structure of gamma PNA is analogue to the unmodified one and the serine side chain is already implemented in *ff99sb,* it was not necessary to introduce new Force Field parameters for the related simulations.

2.7. Modified NAB

Despite of growing interest toward PNA, one limiting issue in PNA chemistry is the scarce availability of solid-state structures due to the difficulties in obtaining crystals of the duplexes, and the even scarcer number of solution structures obtained by NMR. In particular PNA:RNA duplexes are rarer than those with DNA. To design effective systems is important not only to test the effect of modifications, but also to evaluate the effect of introducing new PNA monomers in duplex with the desired DNA or RNA target sequence. Manual modification of duplex structures requires to substitute bases with the desired ones, both on PNA and DNA or RNA. Moreover, these must be placed in the correct orientation allowing for H-bonding and stacking interactions. Little displacements of atoms and functional groups can induce destabilization of the system leading to erroneous simulations. For all these reasons, a manual approach is strongly discouraged and time consuming. In order to automatize this process we propose a modified version of Nucleic Acid Builder (NAB) implemented in AmberTools14. This program works placing backbone atoms in the appropriate positions forming an undistorted helix. Lastly, the remaining atoms of the molecule are introduced trough the coordinates declared in the pdb files for each single monomeric unit. This procedure requires setting of two parameters, reads the helix rise and twisting. The current version of NAB already contains parameters optimized for A-DNA and B-DNA as taken from the literature,[249] allowing for studying different DNA forms. PNAs, as found in the already known duplex structures, show conserved conformation with helical arrangement similar to the DNA parents. Such a similarity, allows using the same approach to generate duplexes with PNA, just properly modifying the original code. We included the PNA monomers in the list of the recognized residues and we provided the corresponding structures to be recalled by the program. The parameters needed for the helix construction were taken from data available in literature[250,143] and the atoms employed for its generation are the corresponding atoms used for DNA and listed below.

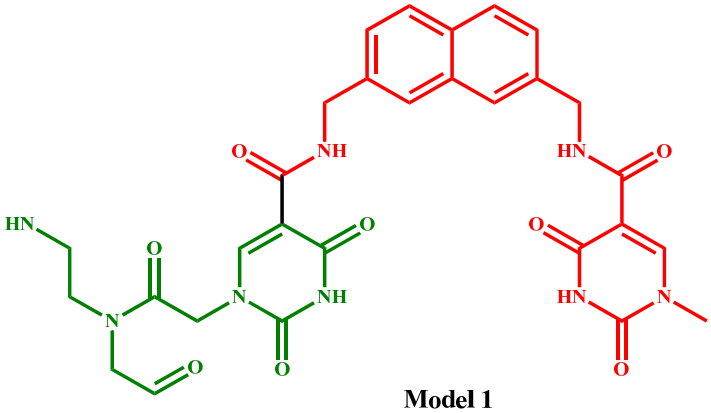
Table 2.1: List of atoms used to construct helix in original Nab program and corresponding PNA atoms used in our program.

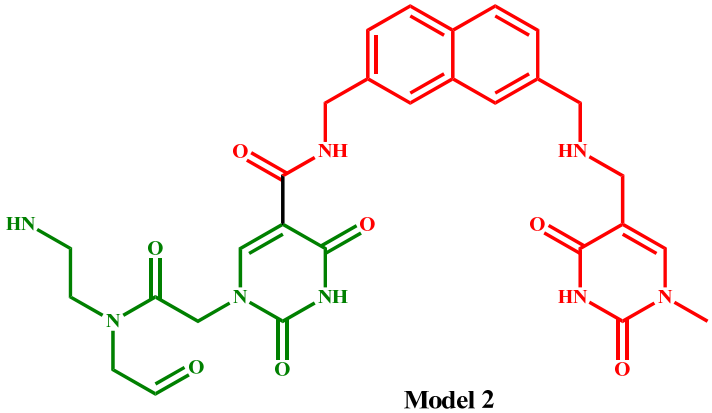
DNA atoms	Corresponding PNA atoms
P	N1’
O5	C2’
C5’	C3’
O3’	C’
C3’	C5’
C4’	N4’
C1’	C8’

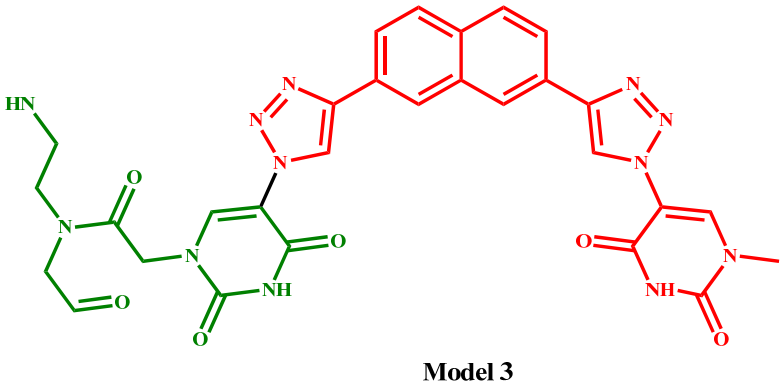
The advantage of our program is to automatize the whole scaffold construction saving a significant amount of time and gaining in accuracy and reproducibility. A full set of structures can be readily prepared in few minutes compared to several weeks of work as it used to be previously. Complete program list is reported in Appendix 8.3.

2.8. Modified bases parameterization

One attractive development is represented by PNAs made by modified monomers. The resulting building blocks can be designed with different complexity and functionalities in order to target the desired biochemical action. Computationally, this translates in more complex structures with non-standard moieties affecting on FF compatibility. Since modified residues are generally inserted in the central part of PNA chain, for each kind of modification (Figure 2.4) we decided to build a partial library focusing only on Central Fragments (Figure2.3). This allowed to save computational time. However, a full library (containing also N-terminus and C-terminus fragments) would be of great interest and it is currently in the optimization process in our group. Each monomer was capped with ACE and NMA moieties and the geometry optimized with Gaussian 09 using DFT B3LYP level of theory and 6-31G* basis set. RESP calculations were performed with Gaussian 09 at the HF level of theory and using 6-31G* basis set. In this case, we used 10 shells and a density of 10 point/$Å^2$. Missing force field parameters for MD simulations were taken for analogy from *ff99sb* and from *GAFF*.







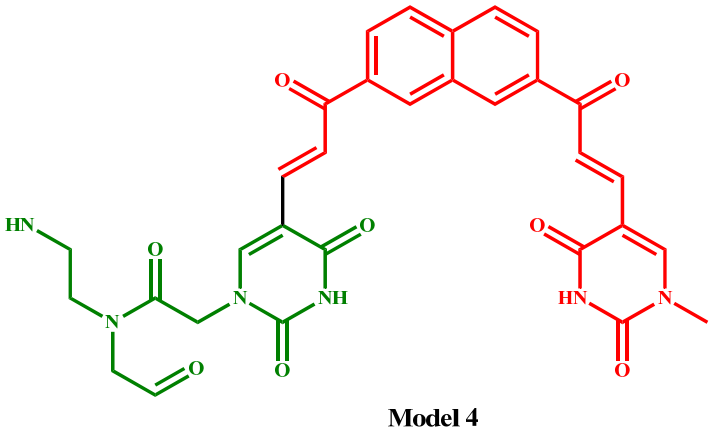


Figure 2.4: Structures of modified PNA monomers. Standard PNA fragments are colored in green and modifications are represented in red.

Starting duplex structures for MD simulations have been generated using NAB module. Modified bases can be inserted manually simply modifying the thymine structure or automatically inserted with NAB. In particular, **Model 1** has been already implemented in NAB code as a fundamental part of this thesis whereas all the other monomers will be included soon. A complete list of charges and parameters are reported in Appendix 8.6.

The Design and the relevance of these dimeric systems will be further described in Chapter 4 and 5.

3. Study of single-stranded PNAs and their interaction with RNA

3.1. Introduction

Understanding the role and the mechanism of action of large biomolecules normally relies on known existing solid-state structures as a starting point. However, as mentioned in chapter 1, in the literature only a limited number of structures containing PNAs can be found. For this reason, the possibility of performing reliable Molecular Dynamics simulation and virtual screening on single stranded PNA, as well as duplex structures containing PNA is highly desirable. In particular, this would enable prediction on the dependence of PNA activity upon structural modifications, by observing molecular details of interaction mechanism which are otherwise difficult to be probed by microscopy or spectroscopic techniques. Such approach could allow to selectively choose the most promising PNAs for the synthesis and experimental tests, based on binding properties prediction, thus saving much time and many resources spent on synthetic efforts.

Despite the growing interest of the scientific community on PNAs activity, the published computational studies are still very few. In particular, examples focusing on PNA:RNA models are very rare. Orozco and coworkers[251,252] were among the firsts to explore the possibility to use Molecular Dynamics as a technique capable to describe PNA properties. Their first study concerned the comparison between PNA:DNA:PNA triplex already reported in literature with their simulation scheme.[252] A successive research has focused on PNA:DNA and PNA:RNA duplexes.[251] Gantchev et al. studied γ-radiation induced cross-linking in PNA:DNA heteroduplexes by means of a 300 ps long MD.[253] Hatcher et al. performed studies of electron transfer in PNA and DNA single strands with 2 ns long simulations,[254] here highlighting the greater structural parameter fluctuations of the PNA compared to the DNA and the resulting influence on charge transfer. Mansawat et al. collected 10 ns long simulations concerning the use of pyrene moiety as an universal base in a acpcPNA:DNA duplexes,[255] revealing important structural information concerning the spatial arrangement of the inserted modification. Sanders et al. performed a 25 ns long simulation on modified PNA bearing a hypoxantine base for the targeting of KRAS miR and they validated simulated data with experimental evidences.[256] Autiero et al. studied PNA:DNA and PNA:RNA duplexes focusing on the structural parameters.[250,257] However, the understanding of the formation mechanism and of the factors governing stability of duplexes between PNA and RNA (or DNA) is still not complete, mainly

due to the multitude of degrees of freedom involved in this process. It is reasonable to believe that the duplex formation is somehow affected by the PNA single strand conformational freedom and to its ability to adapt to the helical conformation of the duplex. The validation of a MD–based procedure for the description of this process would lead to the dual benefit of providing insights into the duplex formation mechanism, and of developing a virtual screening tool for modified PNA moieties. Here we show our results concerning a systematic approach encompassing PNA single strand, and duplex PNA:RNA re-annealing and melting. Simulations described in this chapter were carried out in collaboration with the group of Prof. Michele Parrinello, at the Department of Chemistry and Applied Biosciences ETH Zurich, USI Campus, Lugano, using the computing facilities there available.

3.2. Validation of the force field by MD simulation of PNA:RNA duplexes

As discussed in Chapter1, our final goal is to prepare more effective systems able to target micro RNA. For this reason, as a starting model for assessment of the force field for PNA duplexes, we have chosen the 176D[143] NMR structure (reported in the Protein Data Bank database), which is one of the few PNA:RNA duplex structures reported in literature at the time of this study (crystal structure of PNA:RNA duplex was resolved for the first time at the end of December 2015).[258] Moreover, this 6 base long sequence (6 mer) has the appropriate length for our target modeling, since it is computationally non-expensive and its size is comparable to the putative duplex between a PNA and the seed region of miRs. The sequences of PNA and RNA are respectively H-GAACTC-O^- and 5'-GAGTTC-3' (Figure 3.1). This duplex was modified by removal of phosphate group bound to 5' residue of RNA, because the chosen force field (*ff99sb*) recognizes the RNA strand only without that group. As mentioned before, in the literature there are no consolidated force fields for PNA, and thus one of our aims was to improve their availability similarly to other biological cases as proteins or nucleic acids that are nowadays widely accepted. In order to test the parameters chosen (Appendix 8.1)[235] we performed a 200 ns long Molecular Dynamics simulation on the PNA:RNA duplex described above, using the protocol described in detail in section 2.5. We solvated the system with explicit solvent, in a cubic box of 41 X 50 X 45 $Å^3$, for a total of 9350 atoms.

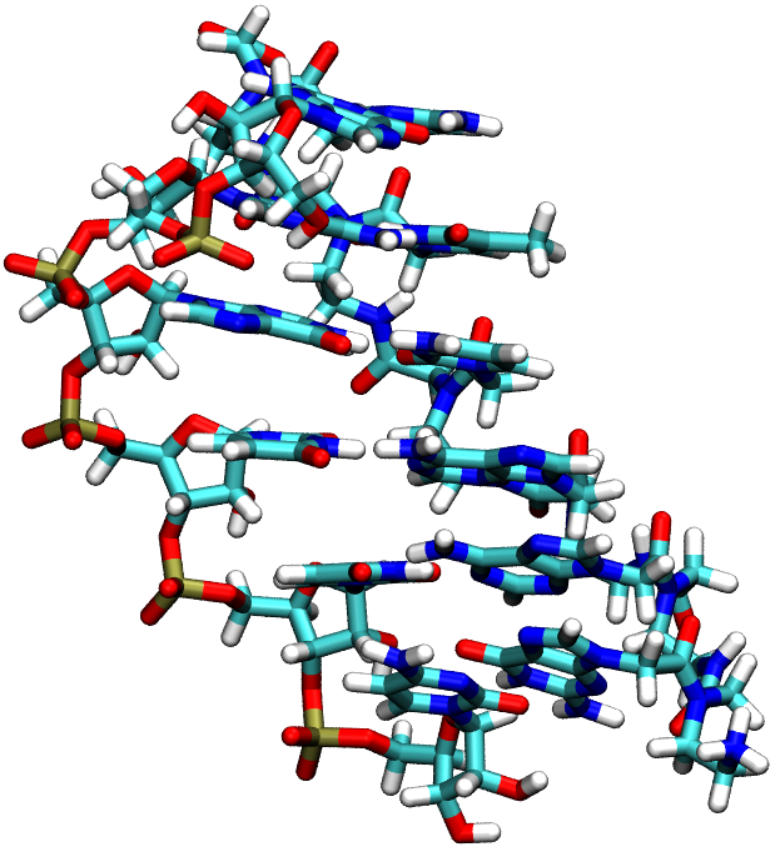
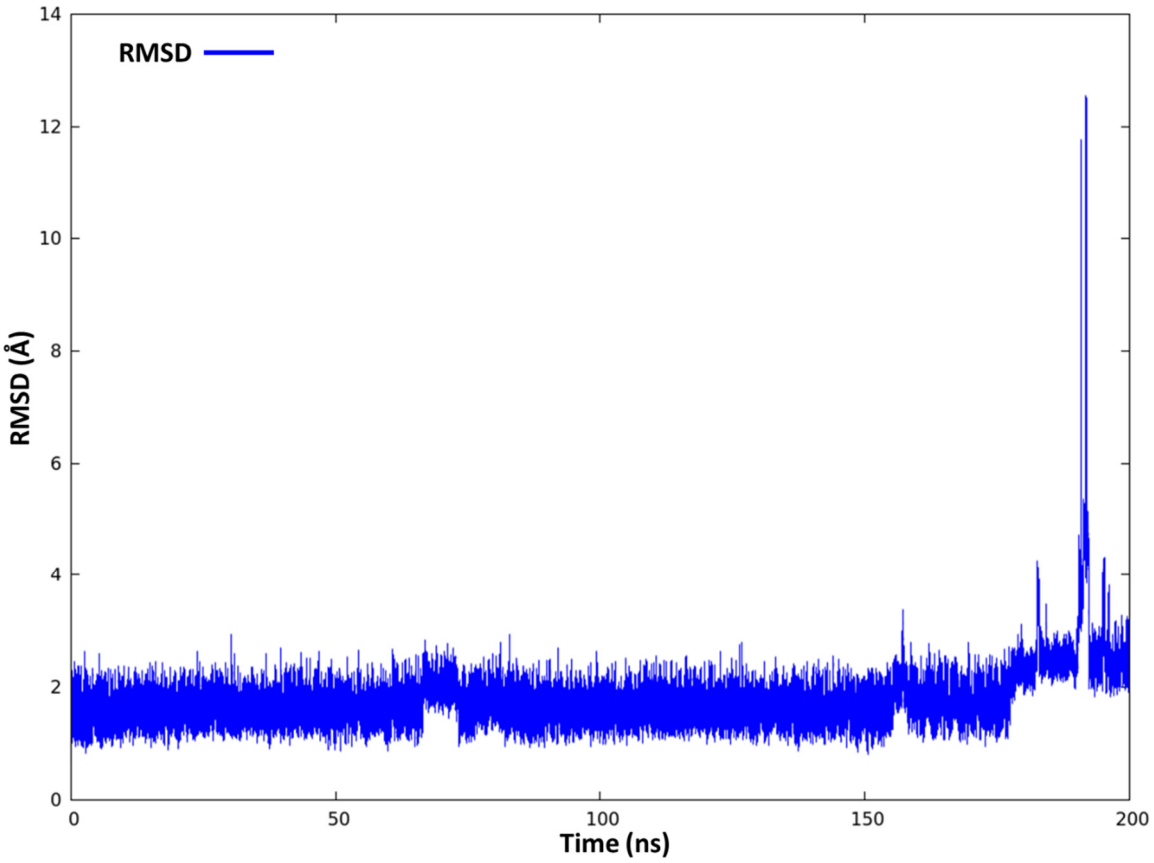


Figure 3.1: Structure of PNA:RNA duplex (176D), taken from Protein Data Bank, used for simulations.

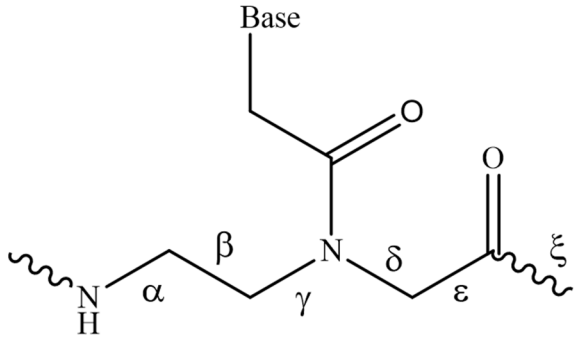


Graph 3.1: RMSD plot of simulated PNA:RNA duplex as function of time.

The duplex conformation was stable for all the simulation length with no significant structural modification. This can be inferred by examining the root mean square deviation (RMSD) that is less than 2 Å except for the very last few nanosecond (Graph 3.1).

The only significant degrees of flexibility of this duplex are represented by the H-Bonding disruption of a terminal couple of bases turning into stacking. This is in agreement with the general view that the two terminal sides of a nucleic acid helical duplex are more exposed than the inner base pairs to the solvent molecules that compete with the H-bond network. If the terminal hydrogen bonds are cleaved the system is destabilized and an efficient way to preserve the overall duplex structure is the stacking between the nucleobases of PNA:RNA terminal fragments. To better test the force field we checked the characteristic torsion angles of PNA obtained in the MD simulation, compared with those reported in literature.[259] The results are in good agreement with the experimental ones (Table 3.1).

Table 3.1: left, characteristic PNA angles; right, comparison between simulated angles and those reported in the literature.[259]



Angle	MD (°)	Experimental (°)
α	86.8±10.9	[40 - 100]
	250.8±15.1	[220 - 280]
β	70.7±11.2	[40 - 100]
γ	80.8±8.1	[60 - 100]
δ	95.6±9.8	[60 - 120]
ε	148.3±13.2	[130-180]
	339.2±16.9	[0 - 40;320 - 360]
ξ	170.3±12.1	[160 - 195]

Next, we considered the MD simulation of γ-modified PNA with the force field developed. In section 2.6, we discussed the procedure used for charge and parameters derivation for γ -modified PNA monomers. Also in this case we needed the force field was validated before using it for further investigations on PNA properties. Therefore, we performed a 50 ns long simulation on a duplex obtained by manual insertion of serine side chain in γ position of each monomer of the PNA:RNA duplex 176D used in the previous simulation. In this case, the duplex was solvated in a cubic box of 59 X 55 X 57 $Å^3$, for a total of 18100 atoms. This calculation was done with a shorter simulation time because, as for the previous one, duplex durableness was expected.

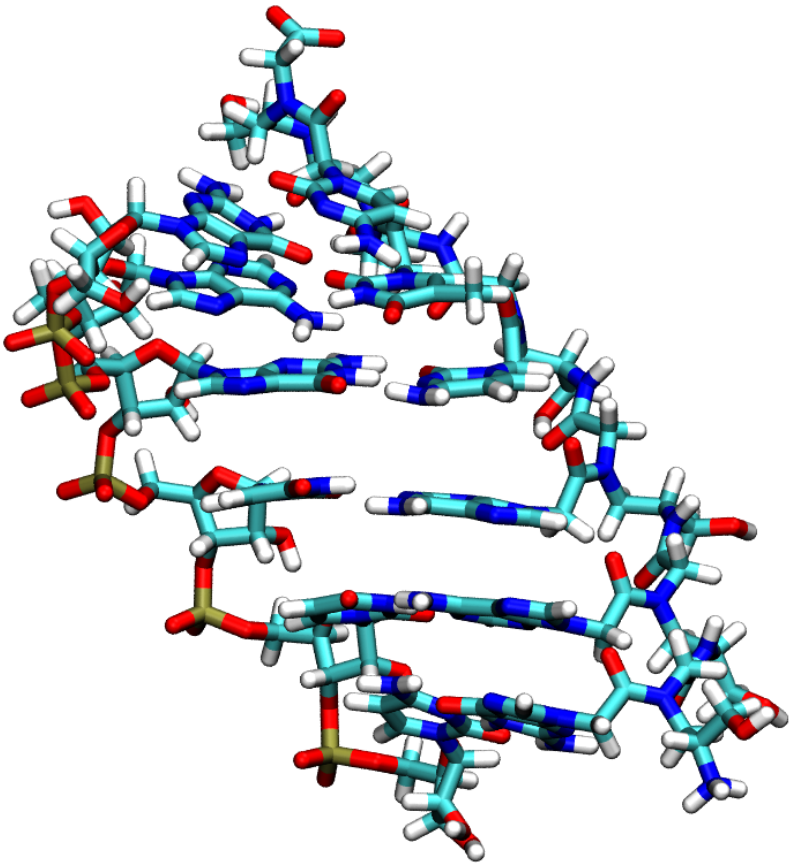
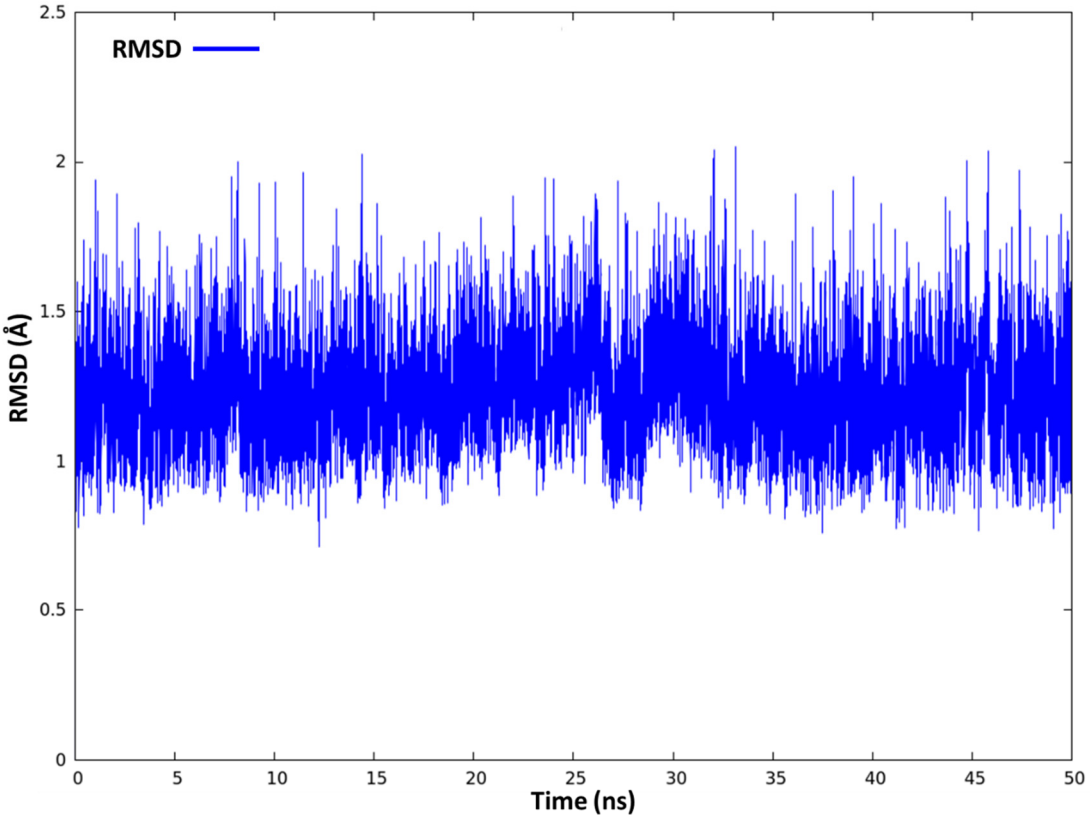


Figure 3.2: Structure of γ-modified PNA. The duplex was generated from system 176D retrieved from Protein Data Bank, by manual insertion of serine side chain in gamma (C5) position.

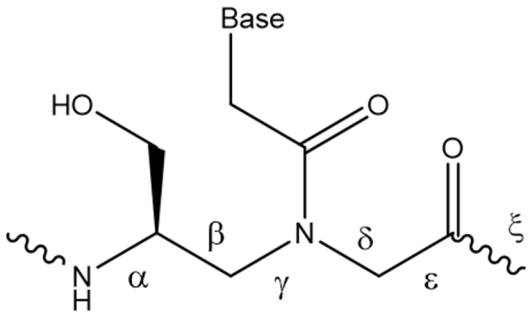


Graph 3.2: RMSD plot of simulated γ PNA:RNA duplex as function of time.

This system was expected to be even more stable than the unmodified one, according to the general properties of γ-modified PNA. Indeed, also in this simulation the duplex resulted perfectly paired for the entire simulation, as proven by the low RMSD (Graph 3.2).

Moreover, characteristic PNA angles were found also in this case to be compatible with those reported in literature (Table 3.2).

Table 3.2: left, PNA bearing serine side chain in gamma position and its characteristic PNA angles; right, comparison between simulated angles and those reported in literature.[259,138]



Angle	MD (°)	Experimental (°)
α	67.3±7.3	[40-100]
β	69.5±9.3	[40-100]
γ	81.1±7.9	[60-100]
δ	94.7±9.5	[60-120]
ε	165.8±10.3	[130-180]
ξ	170.1±8.1	[160-195]

3.3. Simulation of single stranded unmodified PNA (ssPNA)

Understanding the degrees of freedom that characterize the single stranded PNA in solution gives important information on its capability of interacting and eventually binding to RNA. Indeed, conformational freedom of single stranded PNA (**ssPNA**) is directly and strongly related to how the duplex is easily and quickly formed. Moreover, flexibility will also affect the energetic terms in the formation of the final duplex. We investigated the flexibility of a PNA single strand (Figure 3.3) by means of Molecular Dynamics and Metadynamics (MTMD) in order to guarantee a complete sampling of the whole configurational space. Due to the size of this system, we solvated PNA with a larger cubic box of 53 X 67 X 59 $Å^3$, including totally 20775 atoms. Initially, we performed a 200 ns long Molecular Dynamics simulation after thermal equilibration.

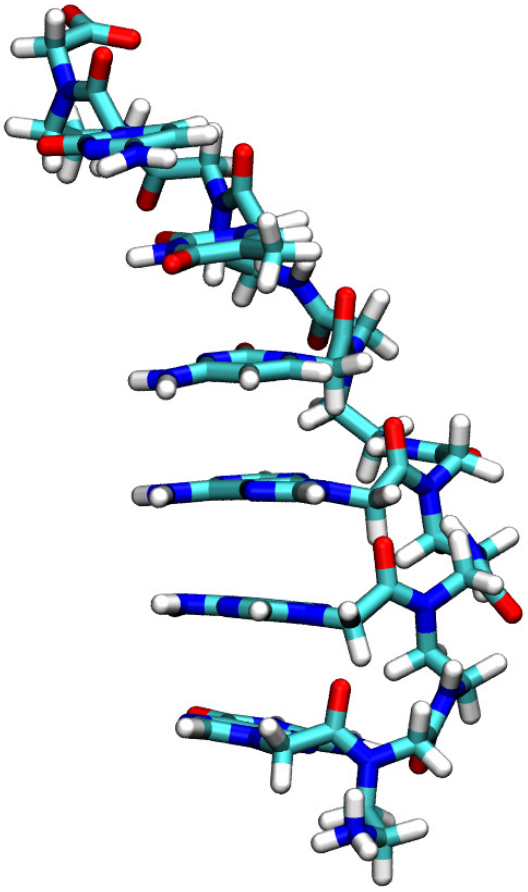
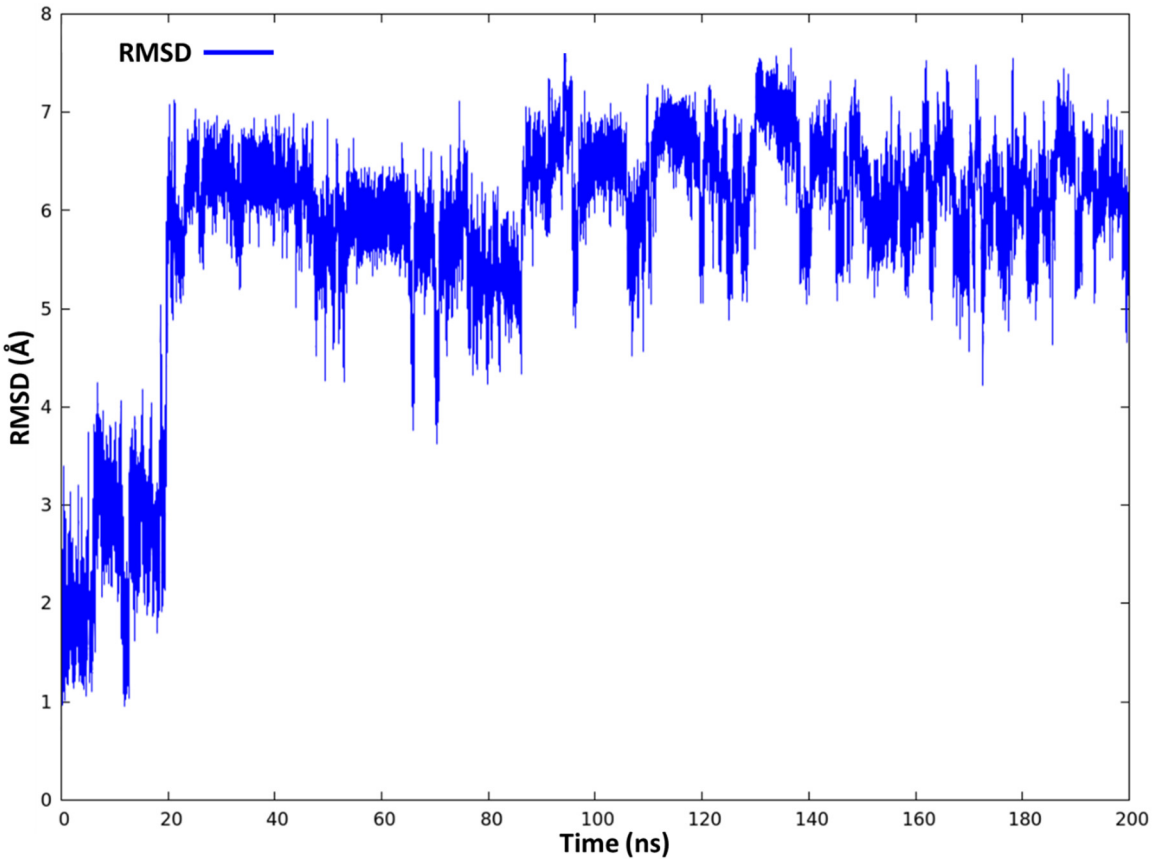


Figure 3.3: Single strand PNA, extracted from duplex 176 D, used as starting structure for simulation.



Graph 3.3: RMSD plot of simulated **ssPNA** as function of time.

In Graph 3.3, we plot the RMSD of the PNA, with respect the thermal equilibrated structure, in order to identify relevant structural transitions. Within the first 20 ns the RMSD is relatively low and centered around 2-3 Å denoting a certain degree of stability of the initial helical conformation. Afterwards, a steep variation of the RMSD denotes a critical conformational change that last until the end of the simulation. This rapid variation is connected to a kind of "folding" of the PNA. We analyzed this structural transition and we discovered that structures showing RMSD values oscillating between 5 and 8 Å correspond to several conformers among which the system is oscillating. This analysis suggests that the initial helical conformation does not last very long and that in water solution the single PNA strand can be found in a multitude of different conformations. In order to have a complete picture of these conformers, we also investigated the populated structures in greater detail using collective variables (CV) (general description in section 2.3). First, we used the head-to-tail (H-T) distance between the centers of mass of the first and of the last monomer (**R**). This variable describes the opening and closing motion of the system and it is related to "folding" and "unfolding" of the PNA. This CV is not sufficient to describe effectively the system because once this is folded there are many possible different stacking interactions. In order to better discriminate these stabilized stacked conformations, we decided to use a local order parameter previously developed for describing crystal nucleation.[260] For each base, we defined a vector lying in the plane of the rings. This CV takes into account the distance, the angle and the coordination number between these vectors. Coordination number represents the number of vectors close within a defined cutoff. This parameter is important when studying nucleation since it determines the presence of an aggregate. The PNA considered in our study is 6 mer long and the maximum theoretical coordination number for each vector is then 5. However, our purpose is not to define an aggregate, but to discriminate stacking interactions. Therefore, for each couple of bases, we defined variables in order to have the coordination function $\boldsymbol{\rho_i}$ (equation **3.2**) always set to one, thus ruling out association. Distances and angles between these vectors are extremely important in stacking description: when two bases are stacked their distance should be defined in a given cutoff range and the angles should assume discrete values. An appropriate definition of this combination allows determining whether two or more bases are stacked (Figure 3.4 A) or fully (Figure 3.4 B – C) and partially unstacked (Figure 3.4 D).

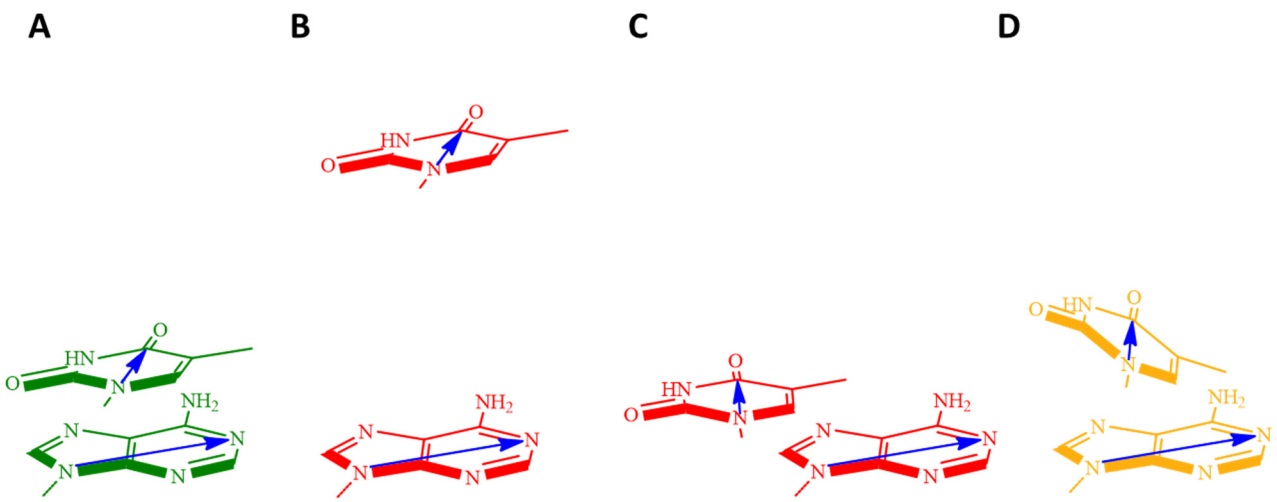


Figure 3.4: Schematic examples of stacking, not stacking and partial stacking arrangements. In blue are represented vectors used to describe local order parameter. **A**) Two bases are stacked; **B**) two bases are too distant for stacking interaction; **C**) two bases are close, but not stacked. In this case distance is favorable, but not the angle; **D**) Distance between bases is optimal, but angle not completely thus leading to partial stacking.

Going into the details, local order parameter is defined as the product of two sigmoidal and one single gaussian function:

$$f_{ij} = \frac{1}{\left(1+e^{a\left(r_{ij}-r_{cut}\right)}\right)} \qquad (3.1)$$

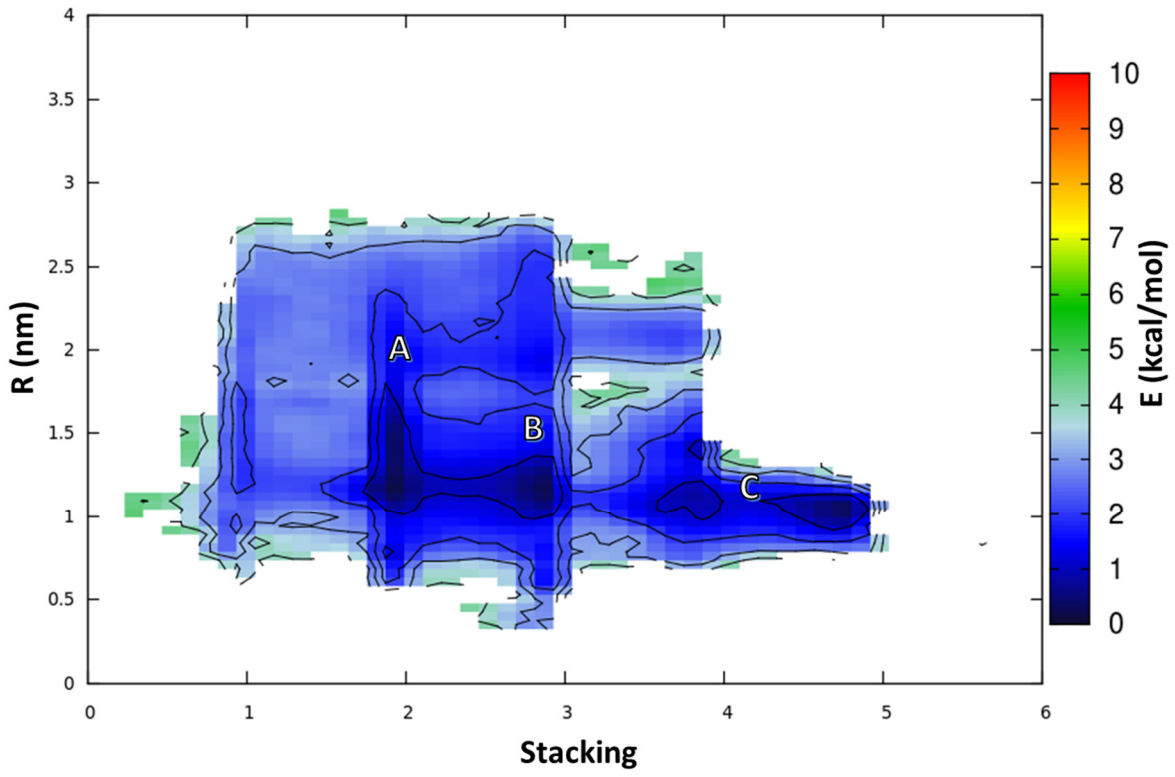
$$\rho_i = \frac{1}{\left(1+e^{-b\left(n_i-n_{cut}\right)}\right)} \qquad (3.2)$$

$$\theta_{ij} = \sum_{k=1}^{k_{max}} e^{-\left((\vartheta_{ij}-\vartheta_k)^2/2\sigma_k^2\right)} \qquad (3.3)$$

where r_{ij} are the distances between the above discussed vectors, r_{cut} is cutoff distance for the stacking interaction, n_i is coordination number, n_{cut} is cutoff for coordination number, ϑ_{ij} is the angle between the vectors, ϑ_k is a favorable angle for stacking and σ_k is the width of the gaussian applied on that angle. Lastly, **a** and **b** are exponential factors, determining how steep are the sigmoids.

These functions essentially monitor respectively the distance between bases (equation **3.1**), the coordination number (equation **3.2**) and angle between bases (equation **3.3**). As discussed above, the coordination parameters were set to obtain a value of $\rho_i = 1$ ($n_{cut} = 1$). The distance function f_{ij} was set to have a value of 1 for distance within a range from 5.0 to 6.5 Å, depending on the couple of bases *i* and *j* considered. Above the cutoff distance, the f_{ij} value rapidly decreases to 0. The angle

function θ_{ij} is a sum of functions, one for each characteristic angle chosen. For each angle ϑ_k a Gaussian function that exhibits maximum value of 1 for ϑ_k, is defined. In order to have stacking, bases should present opportune values of distances and angles. Based on MD simulation data, the coefficients (r_{cut}, ϑ_k, σ_k, a) were tuned in order to maximize f_{ij} and θ_{ij} when bases are stacked. The functions here described are referred to a single couple of vectors, but in our system we have several possible couples and also multiple bases coupled at the same time. Description of every single couple is not meaningful alone, so to describe entirely the system we used a linear combination of these local order parameters, defined for couples of bases, and we considered this combination a measure of total stacking (this function is here after indicated with **Stacking**). Then, we constructed the free energy surface as a probability distribution[219] (equation **2.10**) as a function of **R** and **Stacking**. In Graph 3.4, we report the Free Energy Surface (FES) obtained for the **ssPNA** (extracted from duplex 176D) within an energy range between 0 and 10 kcal/mol (referred to the lowest energy conformation), which can be considered wide enough at 300K.



Graph 3.4: Free energy surface (FES) as function of head to tail distance (**R**) and **Stacking** function obtained from the probability distribution of unbiased MD simulation.

The free energy surface appears quite flat: the local minima are less than 2 kcal/mol more stable than all the other conformations and a very tiny energy barrier separates them. This finding suggests that the single-stranded PNA is free to explore a multitude of conformations avoiding any pre-organized structure capable for potential RNA binding.

The structural analysis conducted on the MD trajectory highlights three principal clusters (Figure 3.5: A, B, C). All of them share the same tail-head distance (**R**~1.1 nm), but with different **Staking** contribution. The first cluster of structures (**A**) is characterized by two couples of bases stacked (**Stacking** = 2, violet and yellow, white and orange), in the second (**B**) three couples of bases (**Stacking** = 3, violet and yellow, white and orange, green and red) and in the third (**C**) more complicated structures with piles of bases leading to higher **Stacking** values (**Stacking** between 4 and 5, white and orange, piling of yellow, violet, red and green / red and green, piling of orange, white, yellow and violet). According to the calculated FES these structures differ in energy roughly by 0.5 kcal/mol (**A** and **B**) and 0.3 kcal/mol (**B** and **C**) and the barriers to be crossed are lower than 2 kcal/mol. Nonetheless, the evolution of the system is not completely random, as could be supposed by RMSD plot, but it is characterized by the passage between different stacked conformations.

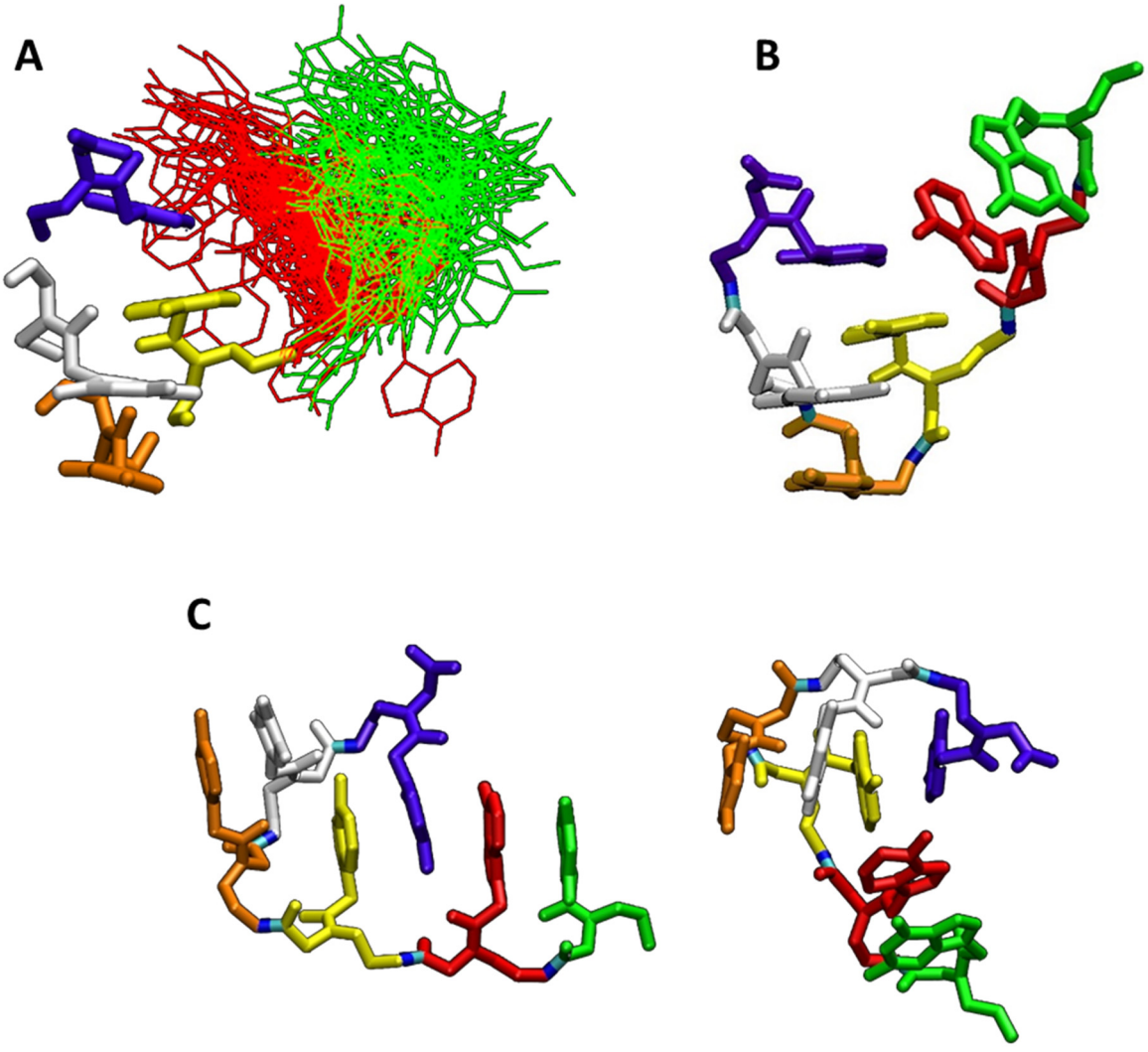
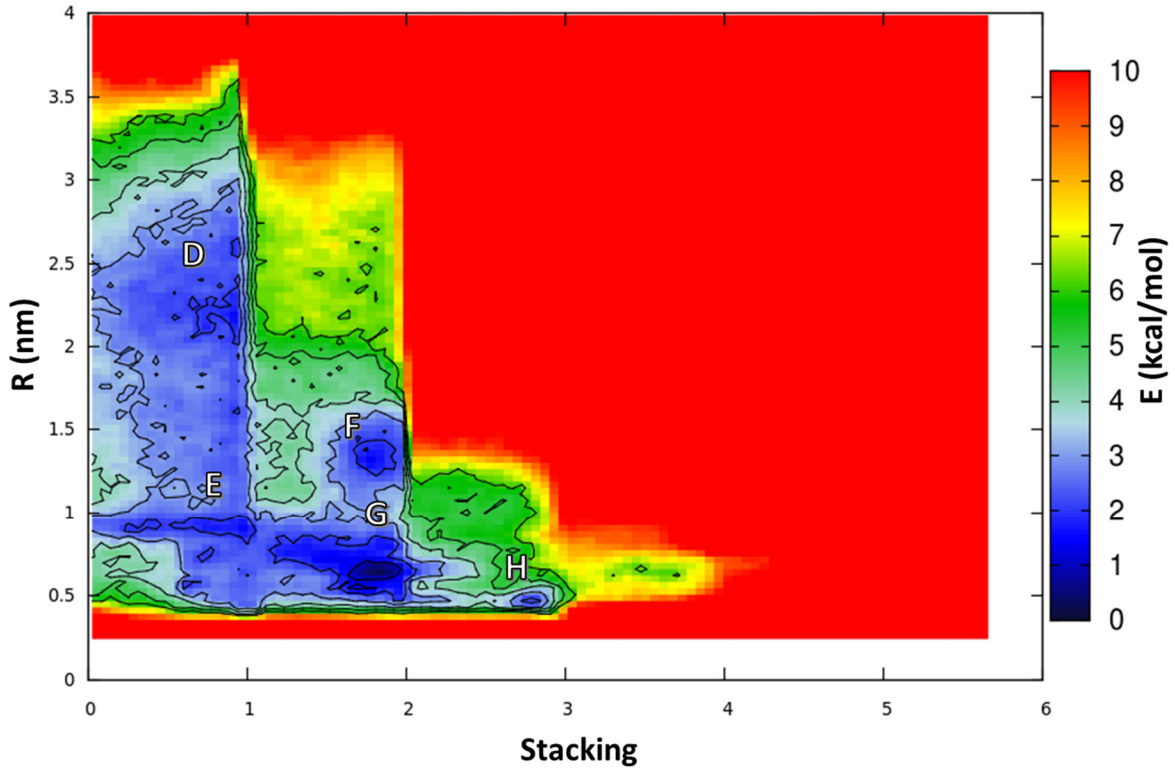


Figure 3.5: Local minima highlighted by Molecular Dynamic simulation. Color code: H-GAACTC-O$^-$.

As discussed previously, free energy surface obtained from distribution of probability is ideally correct only for infinite times, when the system is at equilibrium. However, such kind of systems, with high degree of freedom, is intrinsically complicate to be modeled by standard MD simulations and many important global minima can be missed. For this reason, we aimed to better explore the conformational space of the **ssPNA** by means of Metadynamics. To build the Metadynamics bias during the simulation we deposited every picoseconds a Gaussian hill with width of 0.01 (nm for **R** and dimensionless for **Stacking**) and height of 2.5 kJ (~ 1 k_BT) on the same CVs described above. To ensure the convergence of the Free-energy calculation, we adopted the Well-Tempered Metadynamics scheme (WT-MDMT), as discussed in section 2.4, with gamma equal to 20. External potential on the CVs allowed the systems to populate all the possible conformations of interest. We carried out a 950 ns long WT-MDMT, and the simulation was stopped after verifying that the local minima converged.



Graph 3.5: Free energy surface (FES) reconstructed from Gaussian hills deposition of MTMD. Quantities on x- and y-axes are the same plotted in Graph 3.4.

In Graph 3.5, we report the FES obtained biasing the collective variables described before. The flexibility of the PNA single strand is confirmed by the MDMT, which allows to explore wider configurational space if compared with the classical MD. In particular, the completely unfolded configuration (**R** >2 nm) has weaker stacking between 0 and 1. This suggests that the PNA single strand can also populate non-staked structures. On the contrary, higher levels of stacking imply an enthalpic gain and the corresponding conformations are stabilized when the PNA is significantly folded for **R** < 1 nm. **G** is the global minimum and it differs from others by less than 2.5 kcal/mol (**D** 2.2 kcal/mol, **E** 1.2 kcal/mol, **F** 1.5 kcal/mol, **H** 2.2 kcal/mol). This last FES computed through the Metadynamics expand significantly the configurational space including the one reported in Graph 3.4 and it provides an important added value to the classical MD simulation. A structural analysis of the conformations found inside the principal minima are reported in Figure 3.6. As observed, none of them shows a helical structure confirming the experimental data[261] for unmodified PNA and resulting in hardly obtainable crystal structures justified by the intrinsically high degrees of freedom. In conclusion, both classical MD and MDMT bring to similar results showing how this PNA single strand can freely assume a wide range of conformations avoiding any significant structural preorganization otherwise important for RNA binding. We believe that this finding can be reasonably extended to unmodified PNAs with other sequences.

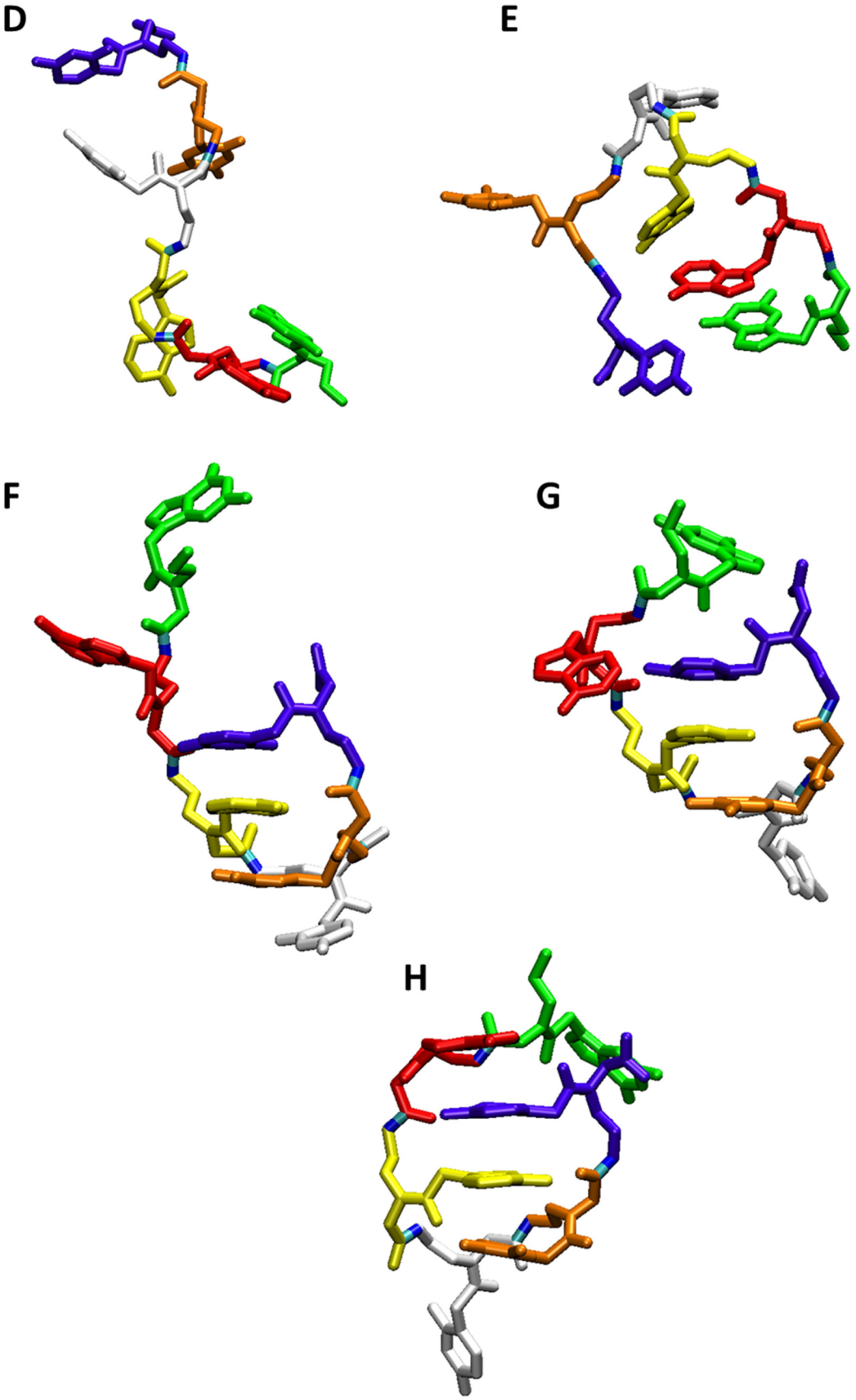


Figure 3.6: Structures of local minima from MTMD. Color code: H-GAACTC-O$^-$.

3.4. Simulation of single stranded gamma-modified PNAs (ss-γ-PNA)

As discussed briefly in Chapter 1, modification of the backbone usually affects preorganization of the system. γ-modified PNAs present a side chain on the γ-position (C5) (Figure 3.7).

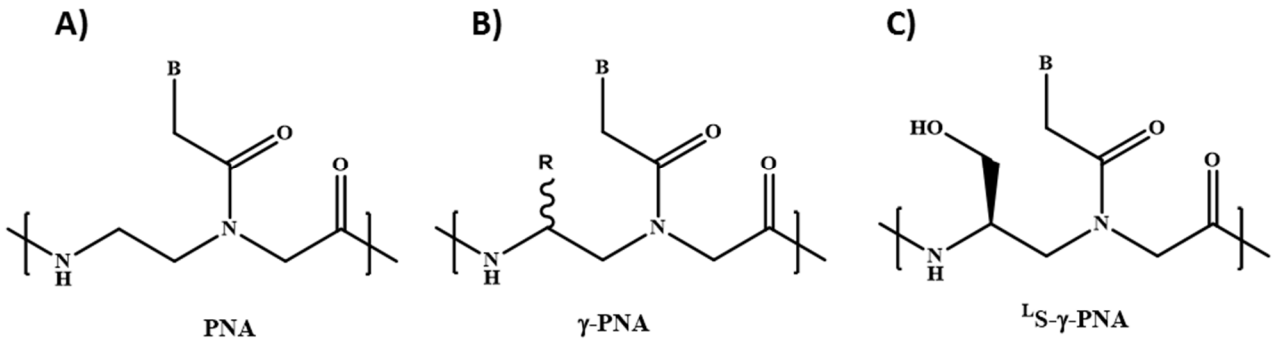


Figure 3.7: Comparison between **A**) standard PNAs, **B**) gamma modified PNAs and **C**) gamma serine modified PNA considered in this study.

Many examples of this type of PNA are reported in the literature.[107,262,263] Such modified PNAs form duplexes with DNA or RNA that are more stable compared to unmodified ones. The properties of these compounds have been related to the preorganization of the backbone.[114] Circular Dichroism spectra of γ-modified PNAs present signals in 200-230 nm range that are not present in unmodified PNAs. The CD signals in this region are due to the carbonyl chromophores of the backbone and were attributed to a helical disposition of these units. We then wanted to better study the effect of the introduction of gamma-modified PNA monomers in conformational properties of the single strand (**ss-γ-PNA**) as a test for the ability of molecular dynamics to predict this different behavior and to give molecular details accounting for the observed differences. Many groups have been attached in this position of the backbone in previous studies, but to reduce computational cost we selected a very simple serine side chain. Moreover, since experimental CD signals are proportional to the number of modified residues, we decided to modify every single monomer of the sequence.

In order to have a comparison with the unmodified PNA, we manually inserted serine groups in the helical structure used in section 3.3 (Figure 3.8). The duplex was solvated in a cubic box of 56 X 74 X 63 $Å^3$ for a total of 21100 atoms and then we performed a 200 ns long MD simulation in the conditions reported in section 2.5. The original helix undergoes some significant structural modification within the first 60 ns of our simulation as demonstrated by the RMSD reported in Graph 3.6.

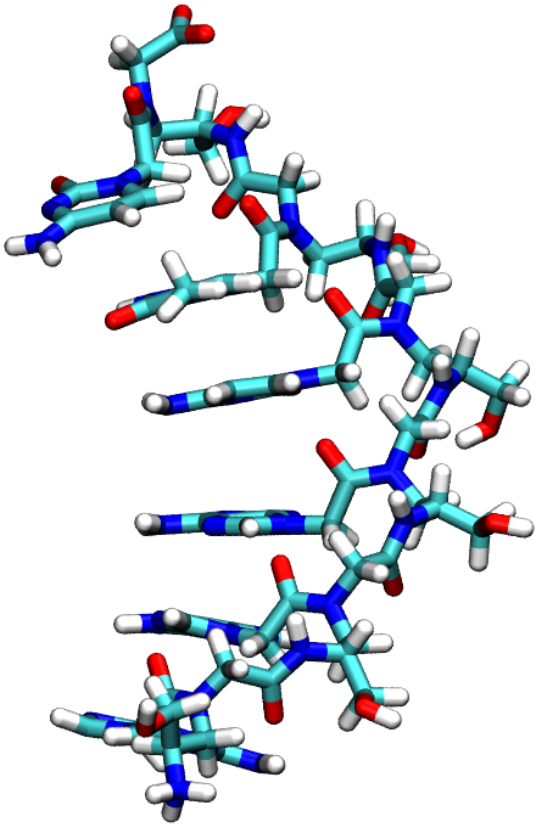
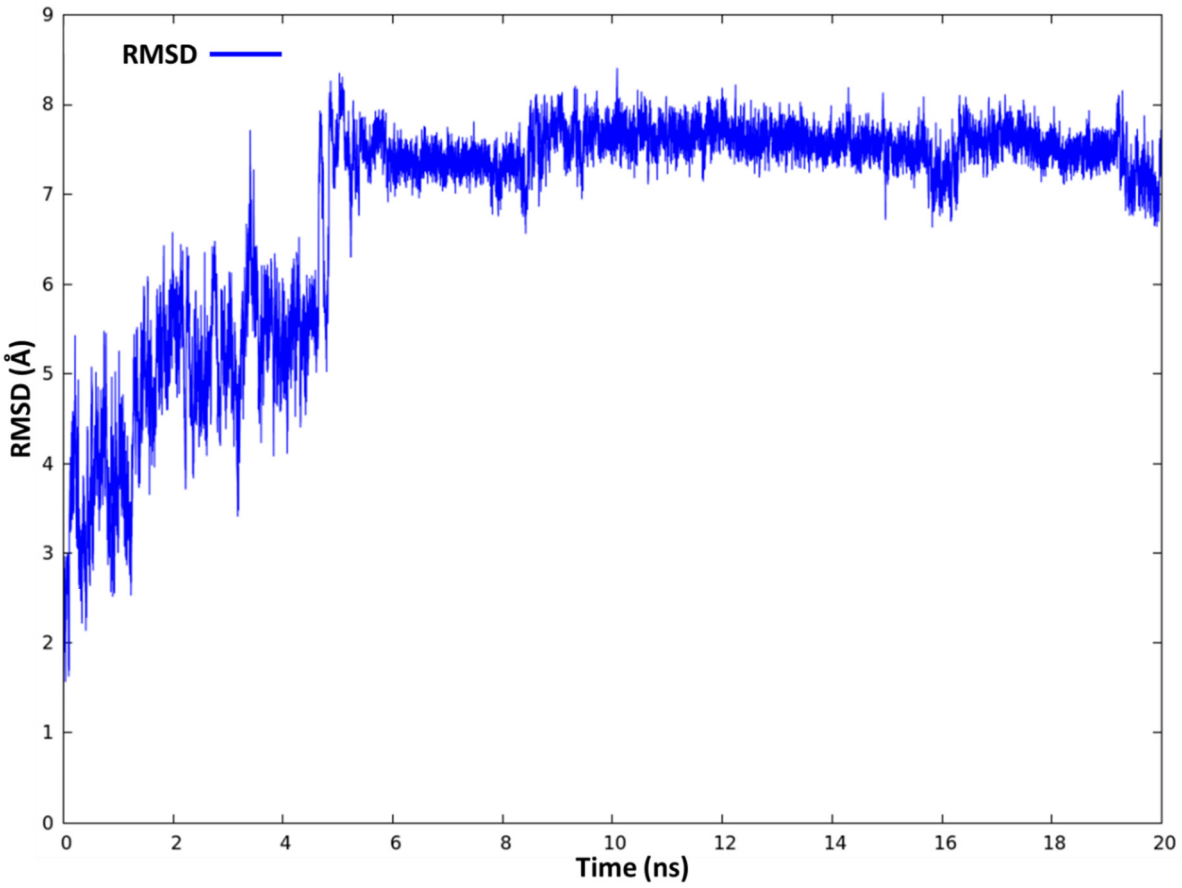
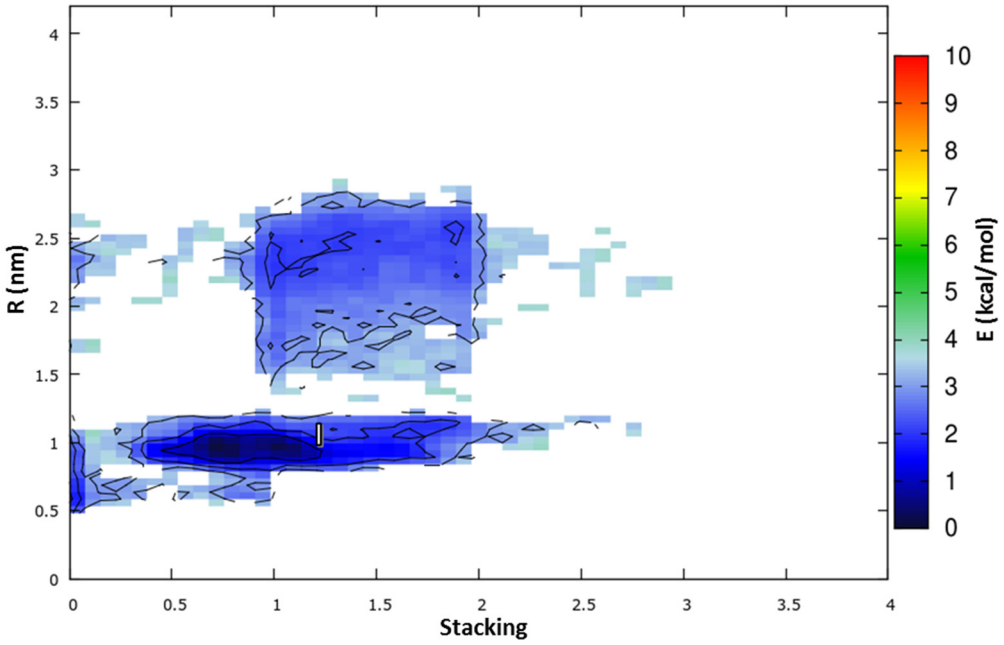


Figure 3.8: Analogue structure of **ssPNA** with γ - modified PNA bearing a L-serine side chain.



Graph 3.6: RMSD plot of simulated **ss-γ-PNA** as function of time.

However, the chemical functionalities positioned in γ induce such a lower flexibility of the backbone if compared with the **ssPNA** discussed in the previous section. The FES obtained with unbiased MD, reweighted on the two CVs discussed before for the unmodified PNA, shows one local minimum characterized by an average **R** (H-T distance) equal to 1 nm and a **Stacking** coefficient between 0.5 and 1 (Graph 3.7). From this basin we extracted several structures very close each other (Figure 3.9). Indeed, the γ-modified PNA quickly converge to a preferred conformation that last until the end of our simulation (Figure 3.9) and that is characterized by a relatively close-packed structure stabilized by the H-Bonding network established between the terminal guanine and the cytosine.



Graph 3.7: Free energy surface (FES) as function of head to tail distance (**R**) and **Stacking** function obtained from the probability distribution of unbiased MD simulation.

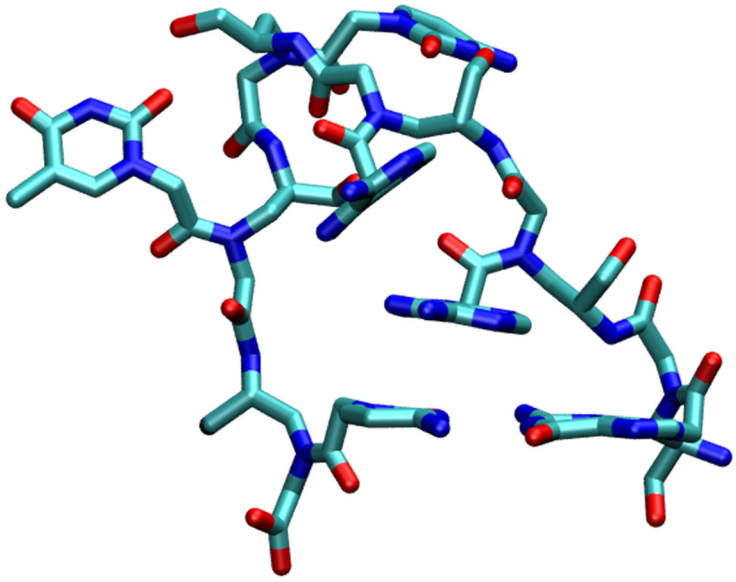
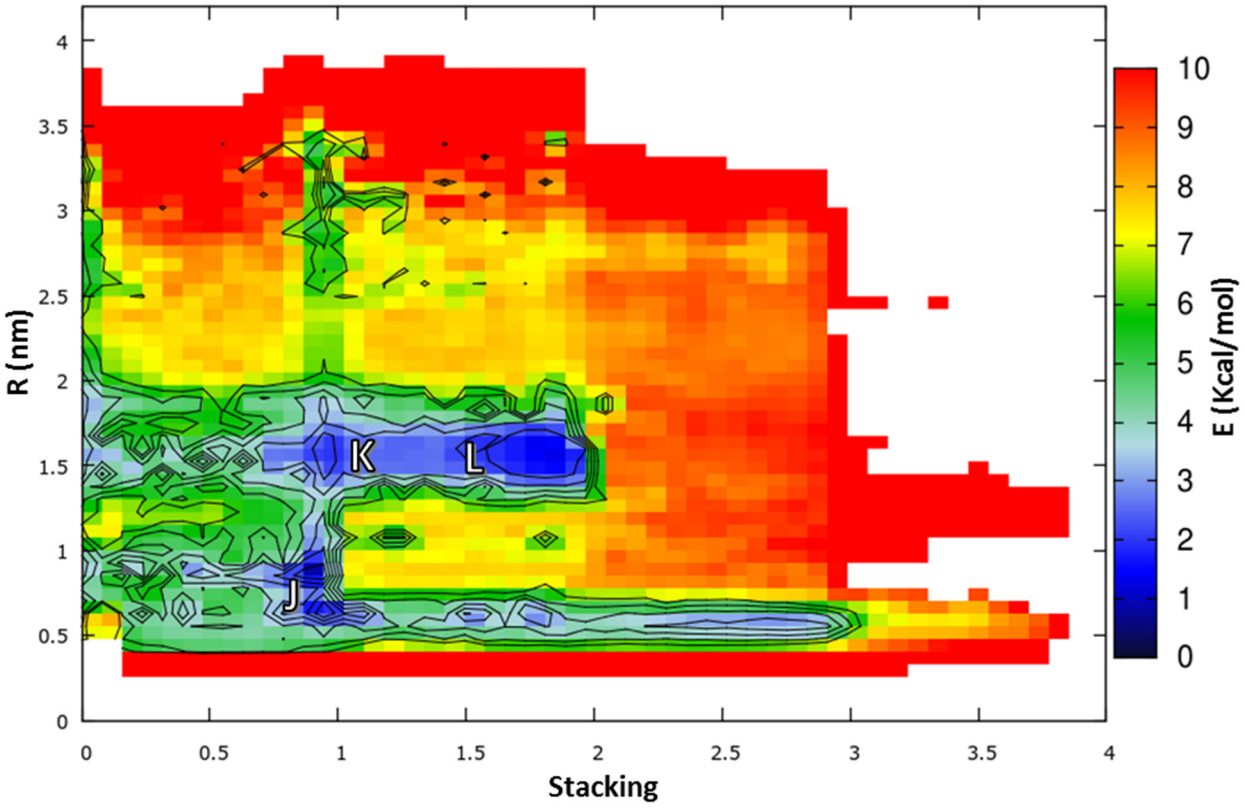


Figure 3.9: Representative structure of minimum **I**.

Even if conformational freedom is reduced compared to **ssPNA**, also in this case it is necessary to use Metadynamics to better explore conformational space. The PNA strand was solvated in a cubic box of 71 X 71 X 71 $Å^3$ for a total of 34700 atoms. To build the Metadynamics bias we deposit every picoseconds a Gaussian hill with width of 0.01 (nm for **R** and dimensionless for **Stacking**) and 2.5 kJ (~ 1 k_BT) of height on the same CVs described in section 3.3. To ensure the convergence of the FES we adopt the Well-Tempered Metadynamics scheme, with gamma equal to 30. External potential on the CVs allows the systems to populate all the possible conformations of interest. We carried out a 400 ns long WT-MDMT verifying, also in this case, the local minima to be converged. In Graph 3.8, we report the FES obtained biasing the usual collective variables.



Graph 3.8: FES reconstructed from Gaussian hills deposition of WT-MTMD.

The rigidity of **ss-γ-PNA** is confirmed by the WT-MDMT and the system populates a reduced, almost isoenergetic (**L** global minimum, **J** +0.1 kcal/mol, **K** + 0.4 kcal/mol) number of states compared to **ssPNA**. Nevertheless, both from unbiased and biased simulations appear that the initial helical conformation is not preserved. However, we observed structures in the three main minima of the

FES and those in basin **L** are particularly interesting (Figure 3.10).

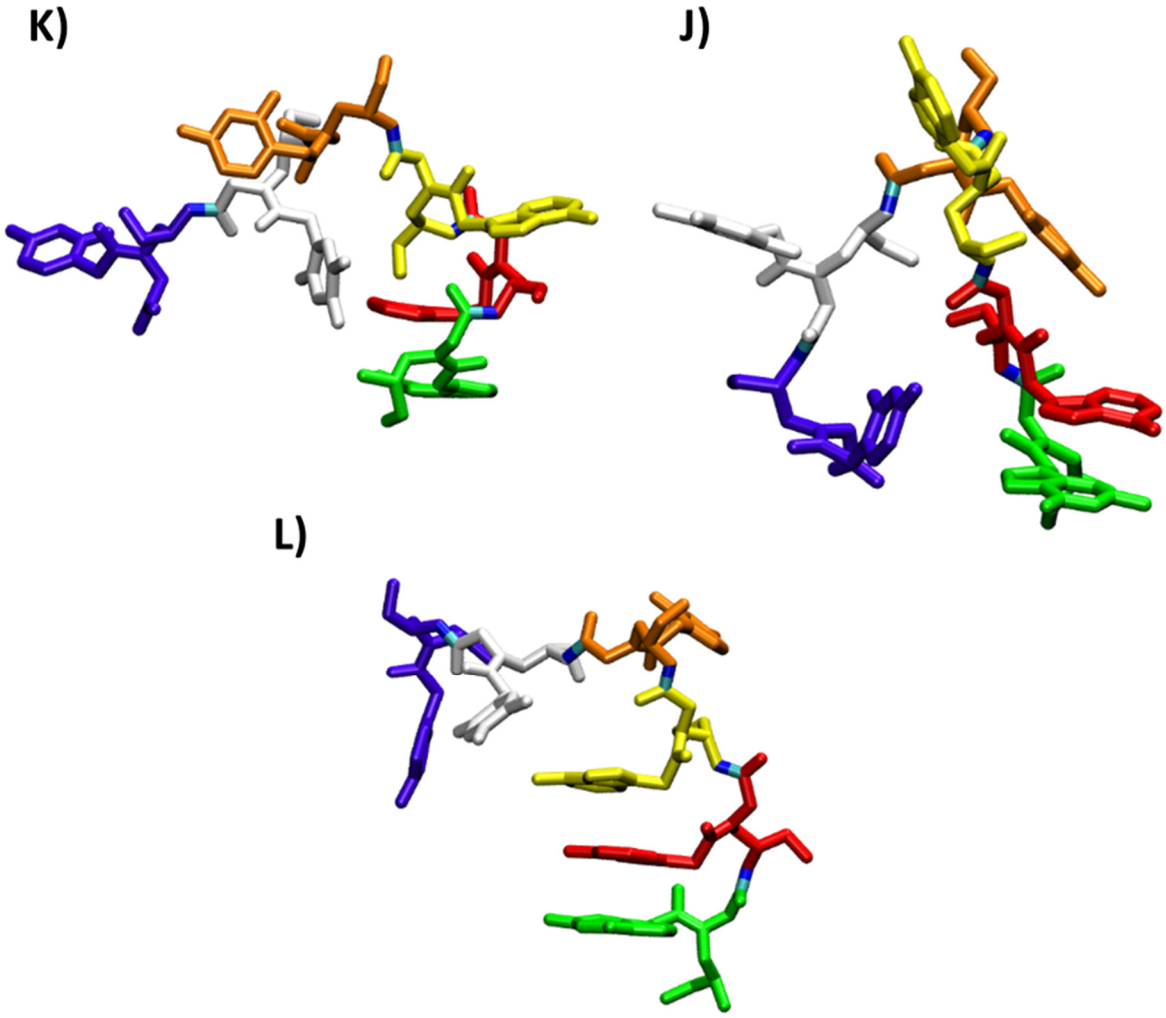


Figure 3.10: Representative structures of main minima of MTMD. Color legend is the same of unmodified PNA.

Portions of the original helical conformation are preserved in this type of structures, in particular three bases starting from the N terminus. The origin of this asymmetrical behavior is connected to the specific sequence, with a three purine on one side and a three pyrimidine on the other side. We therefore suggest that Circular Dichroism spectra reported in literature for this type of PNA[114] does not necessarily represent a stretched helical conformation of the PNA as that present in the duplex (starting conformation in the simulation), but rather a cumulative effect of local helically shaped conformation favored by backbone rigidity.

Another experimental evidence that can be explained by an induced pre-organization of the PNA single strand is represented by the melting temperature (Tm) registered in order to quantify the relative stability of PNA:RNA duplexes. It is known that for each chiral unit an increase of ~ 3°C in

Tm for PNA-DNA duplex and ~ 2°C for PNA-RNA occurs.[114] This can be correlated with the balance between the enthalpy and the entropy of the PNA single strand in solution: higher rigidity translates in a lower internal entropy of the system increasing the possibility of establishing a most durable H-bonding network and consequently requiring higher temperature to melt. The balance between ΔH and ΔS is important also for duplex formation: single stranded PNA is a highly disordered system whereas the resulting duplex is a well-organized complex stabilized by Watson and Crick Bonds. The formation of the duplex is therefore accompanied by a large entropy loss, which is compensated by the favorable enthalpic gain due to hydrogen bonding and stacking interactions. However, if the starting state has a less disordered structure, as in the case of the γ-PNA, due to restricted conformational freedom, the entropy variation upon duplex formation will be less negative than in the case of the flexible unmodified **ssPNA**.

For example, using melting data reported in the work of Ly and coworkers[114] and fitting these with the model described by Marky and Breslauer for a two-state transition,[264] it is possible to calculate the entropy loss for an unmodified PNA and for γ-modified PNA during duplex formation. For the 10 mer PNAs, considered by Ly, results a ΔS of -883.98 J/molK for unmodified PNA whereas it becomes -683.88 J/molK for γ-modified PNA. Modification in gamma position determines then a reduction in entropy loss of about 200 J/molK, in agreement with the above reported findings of MD simulation. Studies are underway to repeat these data using the same sequence used for the simulation.

3.5. MD studies on the interaction between PNA and RNA

Study on single strands permitted to evaluate interesting properties of two PNA systems, highlighting the crucial role of entropy in the behavior of the molecule, but of course, the structure and stability of the duplex and the mechanism of its formation are also crucial points. For this reason we have studied the PNA:RNA duplex in two different processes: duplex re-annealing and duplex melting. These are complementary process and in our study they lead to concordant conclusions that will be here discussed.

3.5.1. Duplex re-annealing

In order to understand possible mechanisms that lead to the formation/dissociation of PNA:RNA duplex we applied our methodology to the study of double strand re-annealing after an induced external stress. As a general strategy, we chose to use two strands (PNA and RNA) with partially base pairing, since a simulation of annealing from totally dissociated strands is computationally too demanding (even for high performance computing tools), since the two strands should a) encounter one another; b) initially pair (nucleate) with all the possible base combinations; c) anneal from different nucleation sites.

We evaluated the molecular dynamics of the system starting from five different distorted geometries (Figure 3.11) in which the two strands were placed in close contact and some of them with partial base pairing. **DIST1** and **DIST2** are spread like a fan on 5' and 3' ends, in **DIST3** strands are shifted for 5-6 Å, in **DIST4** strands are crossed and in the **DIST5** the two strands were stretched using MTMD and put close together. We thermally equilibrated each structure in a cubic box filled with TIP3P water molecules (box size and atoms: (**1**) 45 X 53 X 52 $Å^3$, 12300 atoms; (**2**) 49 X 59 X 47 $Å^3$, 10000; (**3**) 53 X 54 X 55 $Å^3$, 11900; (**4**) 52 X 54 X 51 $Å^3$, 10550; (**5**) 57 X 53 X 75 $Å^3$, 17800) and performed a 200 ns long MD simulation, in the conditions reported in section 2.5. Within this time frame, we can investigate the evolution of these stressed configurations understanding key steps of re-annealing.

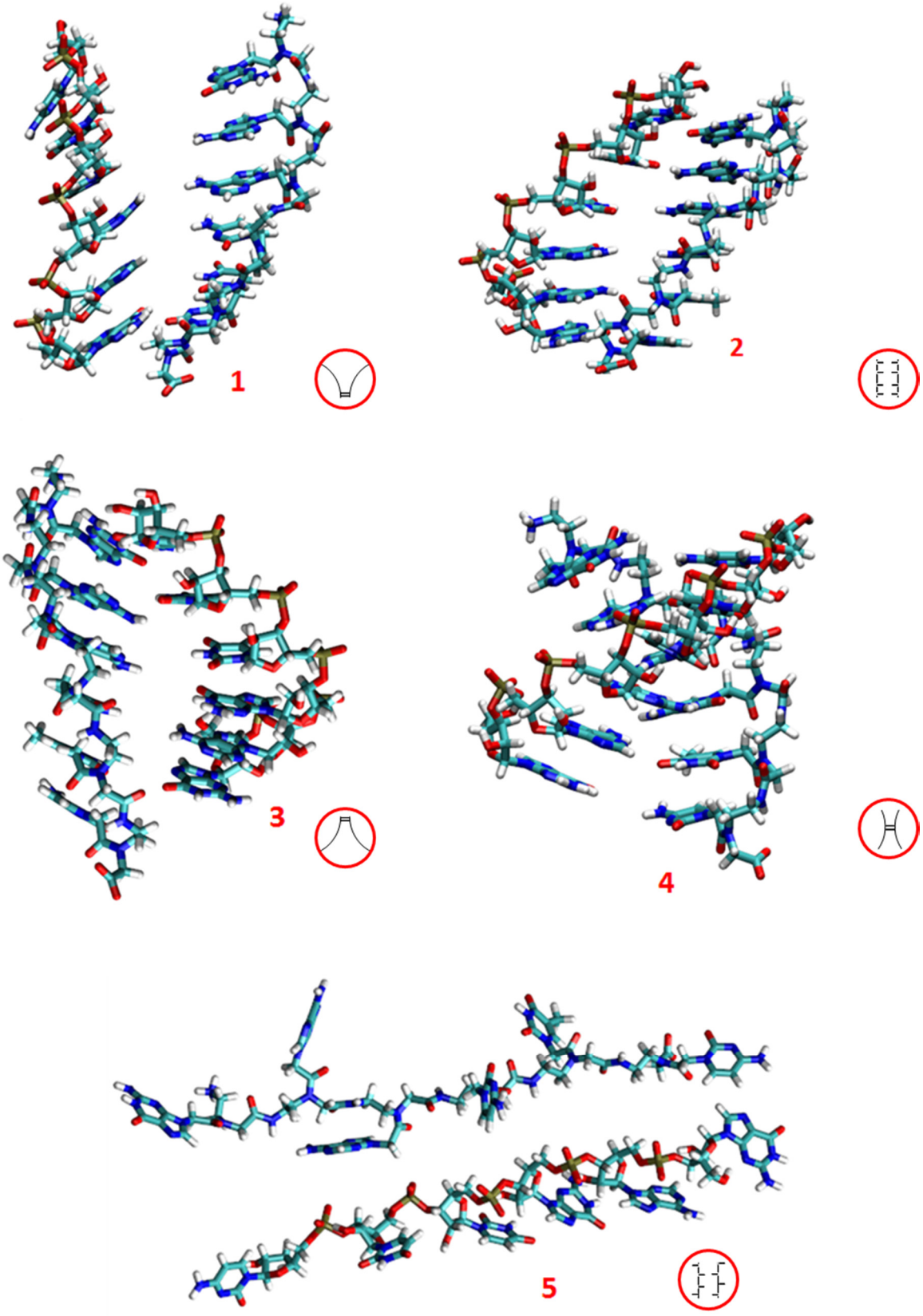
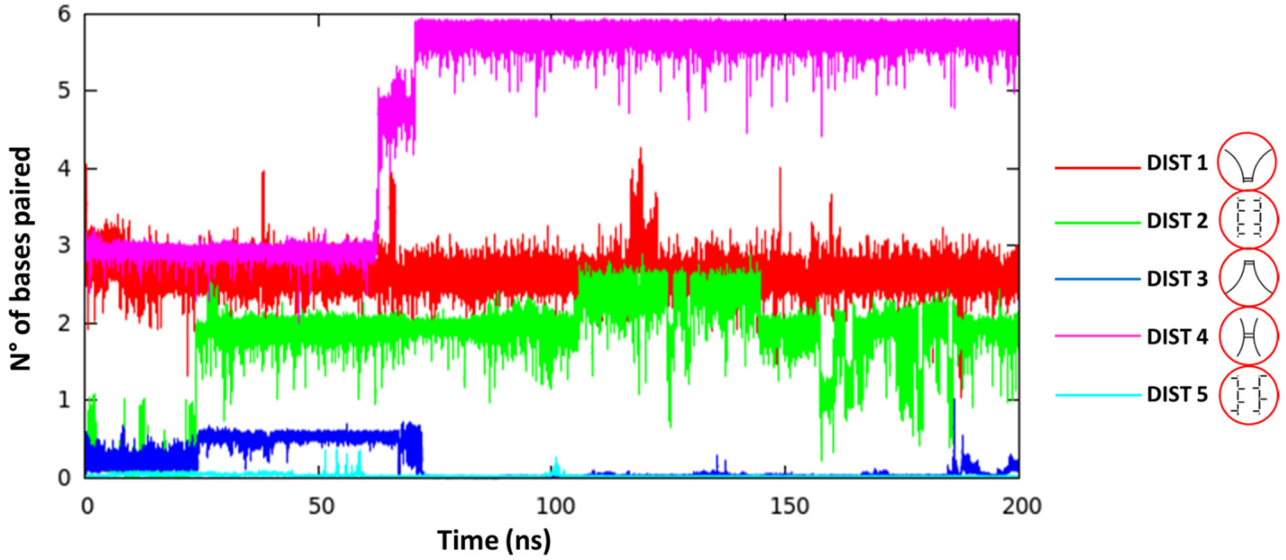


Figure 3.11: Distorted duplex structures: **1**) spread like a fan on CCP:RG5 end; **2**) strands shifted of 5-6 Å; **3**) spread like a fan on NGP:RC3 end; **4**) twisted strands **5**) stretched strands. In red circles the symbolic representation of stress induced in the system are represented.

The re-annealing kinetics is highly dependent on the kind of external stress induced on the duplex. In the ideal case of an infinitely long simulation, all these cases would converge to the same stable duplex structure as investigated in section3.2. However, for finite simulation times the accessible time scale allows to understand some crucial dynamics only that are responsible for triggering the duplex formation. In order to track the duplex time-evolution we computed the number of bases paired. This parameter was calculated dividing the H-Bonding between base pairs, calculated using the CV "Hydrogen bond" as implemented in the Plumed 1.3 library for each couple, by the maximum possible H-Bonding number (complete base pairing). In Graph 3.9 we plot this function with respect to the simulation time.



Graph 3.9: Number of base paired as function of time for **DIST** systems described above.

Among the five different scenarios, only **DIST4** reaches complete base pairing within 200 ns. The distortion initially induced on this structure brings PNA and RNA strands to be interacting by the central bases only. This has consequences on the pairing mechanism evolving simultaneously toward the two extremes of the duplex. The kinetic of the duplex formation is driven by the H-bonding formation between base paring and the base intra- and inter-molecular stacking. The constrain of the central bases paired in the structure of **DIST4** allows to obtain the right balance between H-Bonding formation and stacking for the next bases. **DIST1** and **DIST3** initially present the same base baring number and interaction are in both cases located at the PNA extremity (respectively C-terminus and N-terminus). However, **DIST1** retain a constant number of bases paired during all the simulation, while **DIST3** immediately loses initial favorable interactions and turned

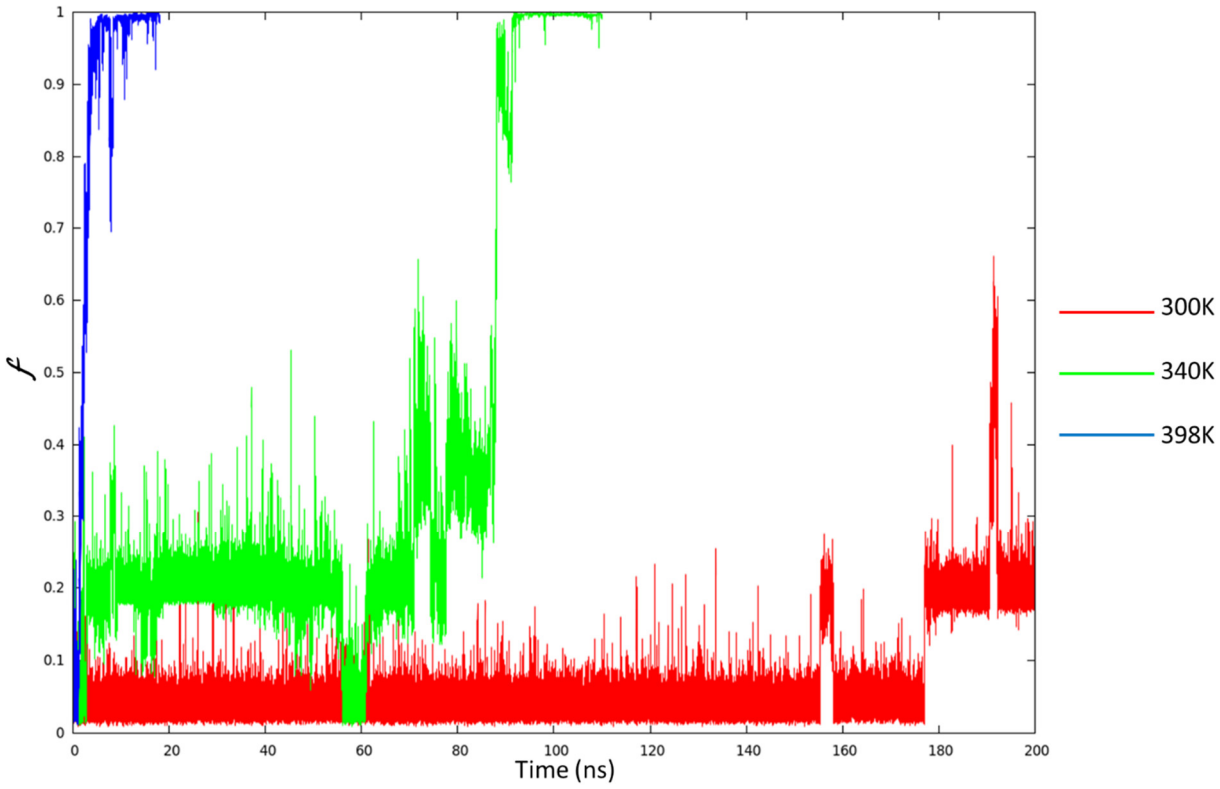
into a structure completely mispaired (Graph 3.9, red and blue traces). This asymmetric behavior is probably due to the specific sequence considered (GAACTC): the first half of the duplex presents a total of 7 hydrogen bonds while the second half 8, slightly favoring C5' edge for an initial interaction. In **DIST2** the two strands are initially shifted of about 5 – 6 Å which means that H-bonding interactions are almost zero. However, the systems readily forms inter-molecular stacking interactions that allow after about 25 ns to establish a partial H-bonding network. **DIST5** is a limit case in which the two strands start from a completely disordered situation. Without favorable interactions, the competition between H-bonding and stacking leads to a very slow formation of the duplex and within 200 ns pairing is almost zero.

The interaction models considered in these simulations are not certainly sufficient to describe all possible ways of contact between the two strands. However, this study demonstrates that the orientation of the interface between bases are crucial; central bases exhibited unique properties in favoring duplex formation, suggesting a possible mechanism of formation of the duplex starting from this position.

3.5.2. Duplex melting

The backward process of duplex re-annealing is the dissociation (i.e. *melting* process). Experimentally, this is achieved increasing the temperature up to a value that is conventionally called melting temperature $\mathbf{T_m}$, that is the temperature at which half of the duplex population is still associated and the other half is not.[265] The experimental data are normally interpreted as derived from a two-state model: duplex and dissociated single strands, without intermediate states, due to the highly cooperative effects of binding interactions. The cooperative effect is attested by the sigmoidal curves obtained by spectroscopic techniques (UV-CD) upon heating of the duplexes.[266] $\mathbf{T_m}$ is generally used as a measure of duplex stability, though in principle it only describes the thermal stability of the duplex. This experimental parameter only gives comparative information on the stability of different duplex, but no information about the mechanism behind the dissociative process induced by the temperature. In order to provide a molecular description we performed several MD simulations on the same duplex PNA:RNA, described in section 3.2, at different temperatures (i.e. between 333K-398K). We define f as the ratio of broken hydrogen bonds with respect the total presented by the system after thermal equilibration. This value ranges between 0, when the duplex is optimally paired, and 1, when the two PNA and RNA single strands are non-

interacting through hydrogen bonds. In Graph 3.10, we report the f value for three representative cases simulated at 300, 340 and 398 K.



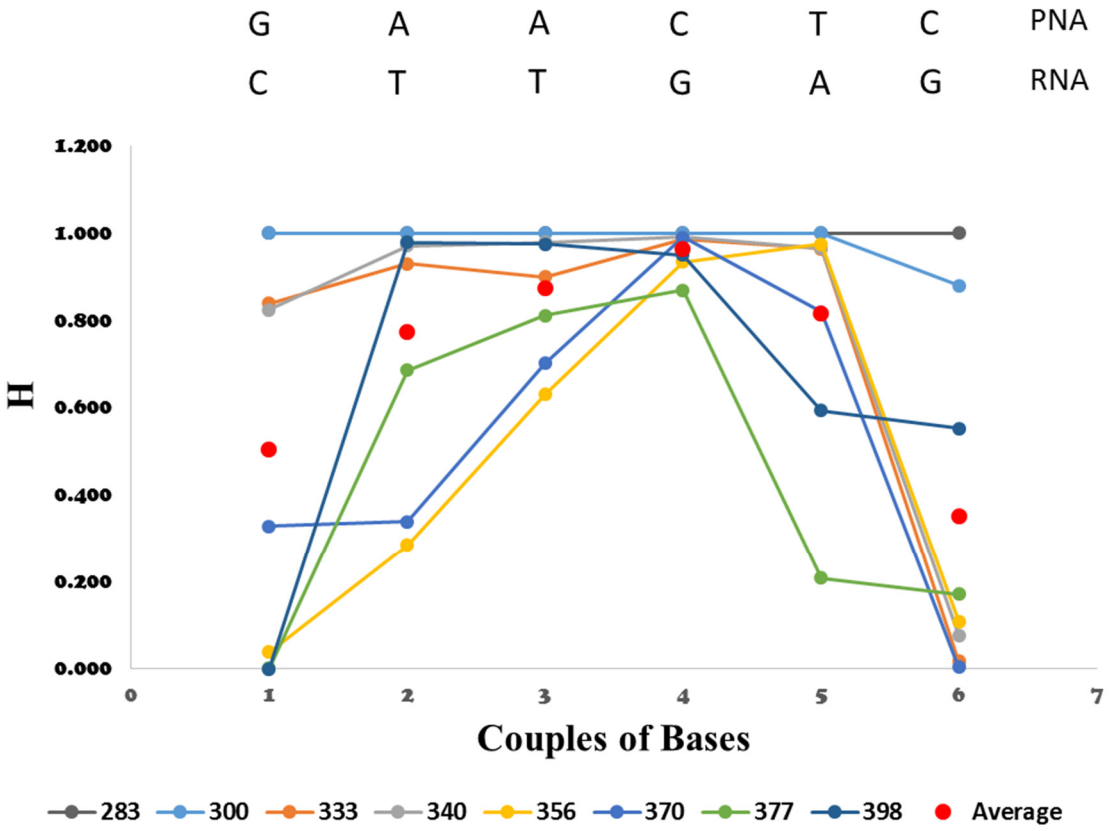
Graph 3.10: f factor as function of time at different temperatures.

This duplex at 300K is stable and only in the very last part of the simulation we observe pair of bases switching from hydrogen-bonded to stacked state, increasing mildly the f value. At 340K we observe an intermediate case in which thermal denaturation starts at the very beginning of the simulation and it remains quite ineffective until 80 ns, when the system shows a sudden increase of f due to the complete opening of the duplex. At 398 K denaturation is very quick and the dissociation occurs in the very first nanoseconds. This temperature dependence was expected and provides important insights to be analyzed in order to extract the contribution of each couple of bases into the denaturation process. Similarly to f, for each base pair we can define the hydrogen bonding parameter **H** as the ratio between the time being paired and the total simulation time before melting condition (i.e. $f \leq 0.5$). This value gives us the opportunity to analyze the most persistent base pairs in duplex dissociation. In Table 3.3, we report the **H** parameter calculated for each base pair at the different temperatures analyzed.

Table 3.3: **H** factor for each base pair as function of temperature.

	283	300	333	340	356	370	377	398
GC	1.000	1.000	0.839	0.825	0.039	0.328	0.002	0.000
AT	1.000	1.000	0.931	0.972	0.284	0.339	0.686	0.980
AT	1.000	1.000	0.900	0.979	0.630	0.702	0.811	0.976
CG	1.000	1.000	0.988	0.991	0.934	0.991	0.870	0.950
TA	1.000	1.000	0.965	0.966	0.975	0.819	0.208	0.593
CG	1.000	0.880	0.017	0.077	0.109	0.004	0.171	0.552

As a reference, we took duplex at 283 K, which is perfectly stable along the entire run as demonstrated by each base pair contributing with an **H** factor equal to one. By increasing the temperature, the kinetic energy of the system grows along with the probability for base un-pairing. This induces the duplex to dissociate due to the loss of hydrogen bonds and the increasing of entropy. In Graph 3.11 we plot data from Table 3.3 for a clear interpretation of the phenomenon. **H** factor is on the average lower for those bases placed at the two extreme sides of the duplex while it is higher in the middle. This suggests that the duplex dissociation occurs starting from the terminations and then favoring water invasion on central bases.



Graph 3.11: **H** factor at different temperatures for each couple of base. PNA sequence is written from N-terminus to C-terminus, RNA sequence is written from 3' to 5'.

These results are complementary and in agreement with the study on re-annealing presented in the

previous section. Moreover, they further suggest the importance of central bases in both formation and dissociation process. A direct consequence of this finding is that in order to improve PNA binding properties, functional modifications must be focused on stabilizing central part of the PNA. The further effect of introducing modifications in different position of the sequence has not been specifically studied, but can be deduced by many experimental data.[106,267] A complete experimental study on base mismatched as a function of modification positioning[268] is available in the literature. In this study, on a 17 mer PNA:DNA duplex the authors showed that a single mismatch in the central part of the PNA leads to a destabilization that causes a drop in the melting temperature up to 25.7 °C, while mismatches in terminal positions destabilize much less the system (range 0.4 – 10 °C).

3.6. Conclusions

Crystal structures represent an important point of reference for the study of PNA-related systems. However, they often offer a static point of view of the problem, while these types of compounds are extremely dynamic and can adopt different conformations. In this chapter, we presented a novel approach using Molecular Dynamics and Metadynamics that permitted to focus on molecular properties of the systems and to observe evolution and adaptability of PNAs to different conditions. Unmodified single-stranded PNA (**ss-PNA**) presented a great flexibility, populating several conformations: stretched conformations with low stacking, stabilized by entropic factors and closed conformations with high staking, stabilized by enthalpic factors. This can explain unique binding properties of these compounds that can adapt to the substrate to bind. As expected, **γ-ss-PNA** presented a reduced flexibility compared to unmodified one. However, this system is not rigid, having the possibility to populate different states and not, as previously proposed, fixed in a rigid helix disposition. Nonetheless, among these well-defined conformations we were able to identify chirally organized structures that could justify the experimental CD signals and a smaller number of populated conformations, which account for the higher binding capability of γ-PNA compared to unmodified PNA discovered experimentally.

Our simulations on PNA:RNA interactions suggested a possible mechanism of formation for the duplex which starts from central bases pairing. Simulated annealing and melting processes both highlighted the importance of these bases for duplex formation and stability. Thus functional modifications, aiming to reinforce sequence selectivity, should be designed by focusing in the central part of PNA strand and must either favoring kinetics of duplex formation or stabilizing base pairing through extra interactions.

4. Synthesis and study of "Adenine-clamp" base

4.1. Triplex structure and TAT mimicking base

As discussed in Chapter 1, usually PNA binds to DNA forming a helix similar to DNA:DNA duplex. When the target sequence is homo-purine and the PNA sequence is homopyrimidine, a triplex structure, in which two PNAs bind one DNA strand using both Watson-Crick and Hoogsteen bonds, can be formed.

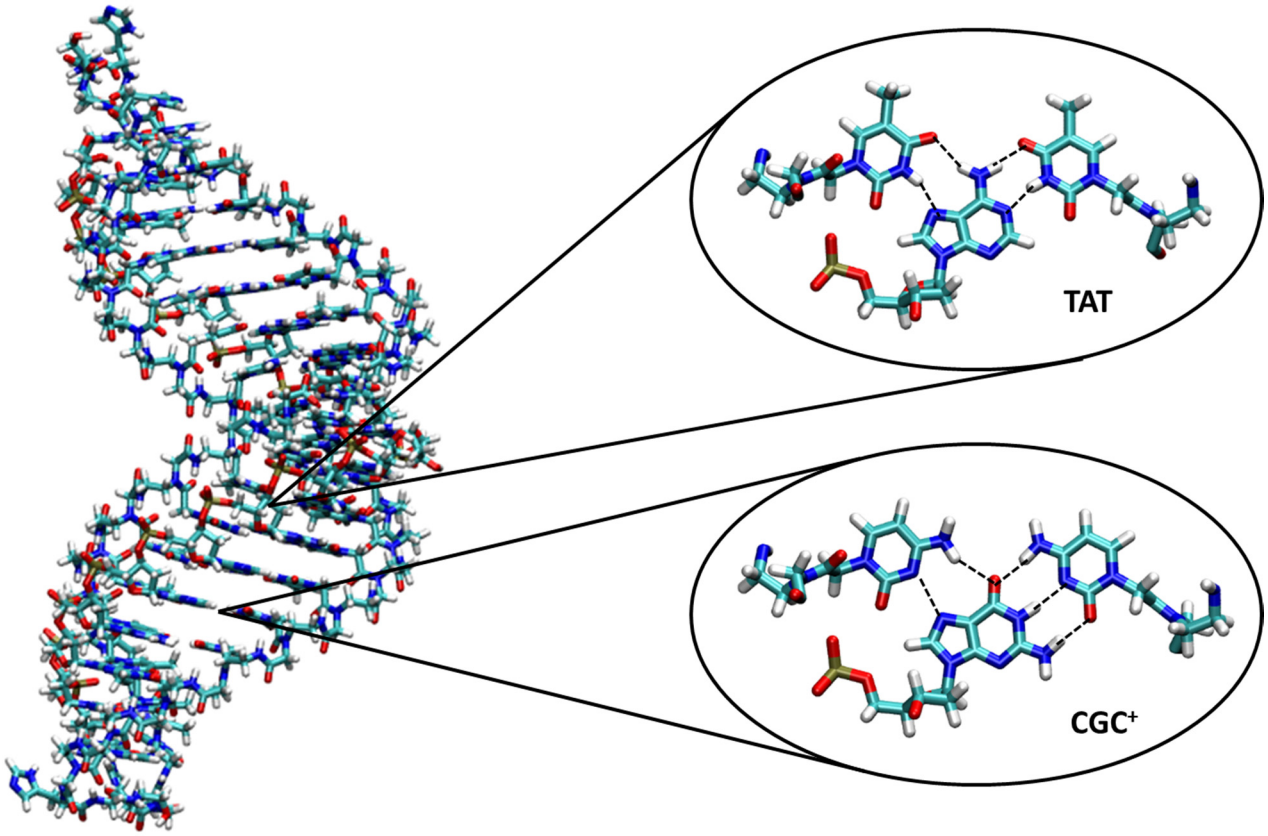


Figure 4.1: Crystal structure of a triplex taken from PDB (1PNN) and enlargement of a TAT and CGC^+ triplet.

In Figure 4.1 an example of PNA:DNA:PNA triplex taken from the Protein Data Bank (1PNN)[140] is shown, displaying the typical H-bonding patterns described above. Triplexes are very stable structures showing high melting temperatures. By controlling the pH it is possible to observe, upon melting, the transition from triplex to duplex and then to single strand or directly from triplex to single strand;[269] the pH dependence is primarily due to the protonation of cytosine needed for the CGC+ triplet; under optimal conditions, the binding of the two strands is cooperative, thus the

binding of one strand favors the binding of the second one by Hoogsteen base pairs. The distance between methyl groups of thymine in triplex structure is roughly the same in all TAT triplets and has a mean value of 9.5 Å. In a previous PhD thesis work,[270] this allowed to design a dimeric base (figure 4.1) by connecting two uracil moieties with an opportune linker based on 2,7-bis(aminomethyl) naphthalene; this was chosen as a bridge because it can in principle allow to obtain a distance between the methylene groups protruding from uracils of about 10 Å, similar to the distance of the thymine methyl groups in the triplex structures. A rigid spacer like naphthalene is necessary also to avoid collapse of the uracil moieties by formation of intramolecular hydrogen bonds.

4.2. Dimeric uracil base: testing previously reported synthetic route

In Figure 4.2 the structure of the PNA monomer bearing the previously designed dimeric base is reported. Synthesis of such modified compound can be achieved by four different synthons, depicted in Figure 4.2 and represented with different colors.

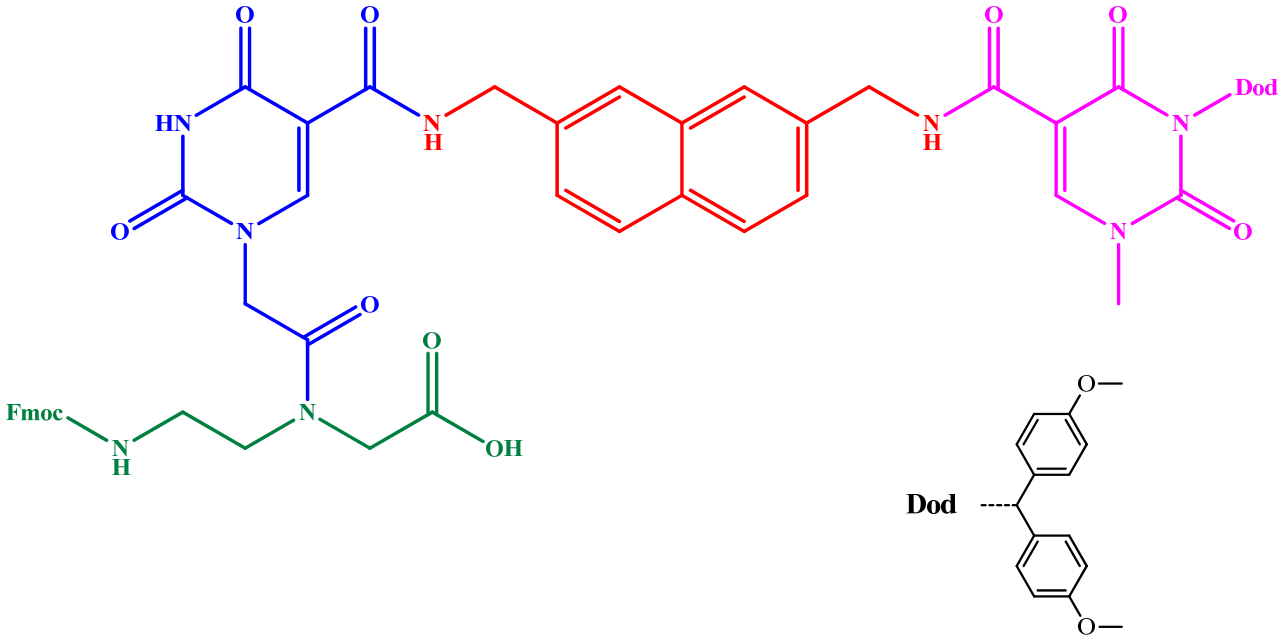
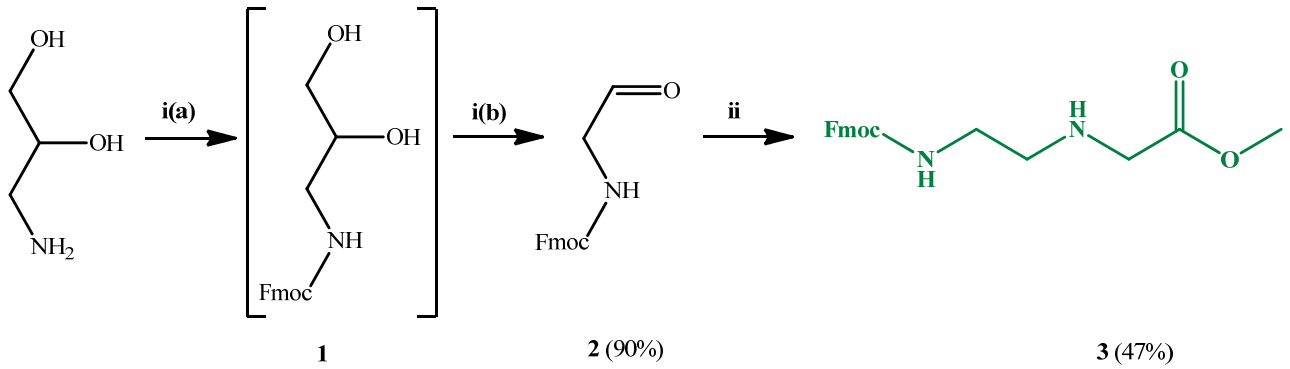


Figure 4.2: Structure of the dimeric base previously designed.

In this thesis, the steps previously used for the synthesis of constituent units[270] were checked and optimized, where possible.

The synthesis of the PNA backbone (Scheme 4.1) was achieved starting from 3-amino-1,2-propanediol which was Fmoc-protected to give the intermediate **1,** which was readily oxidized to

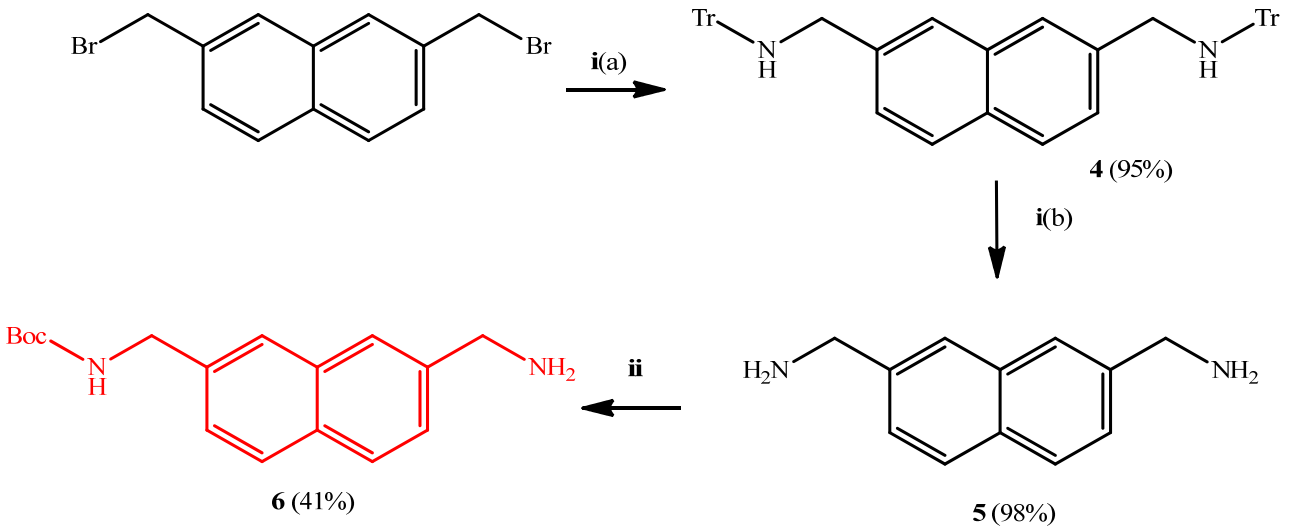
the aldehyde **2** using mild conditions (KIO_4). The protected backbone **3** was obtained with reductive amination of **2** with glycine methyl ester.[271]



1 **2** (90%) **3** (47%)

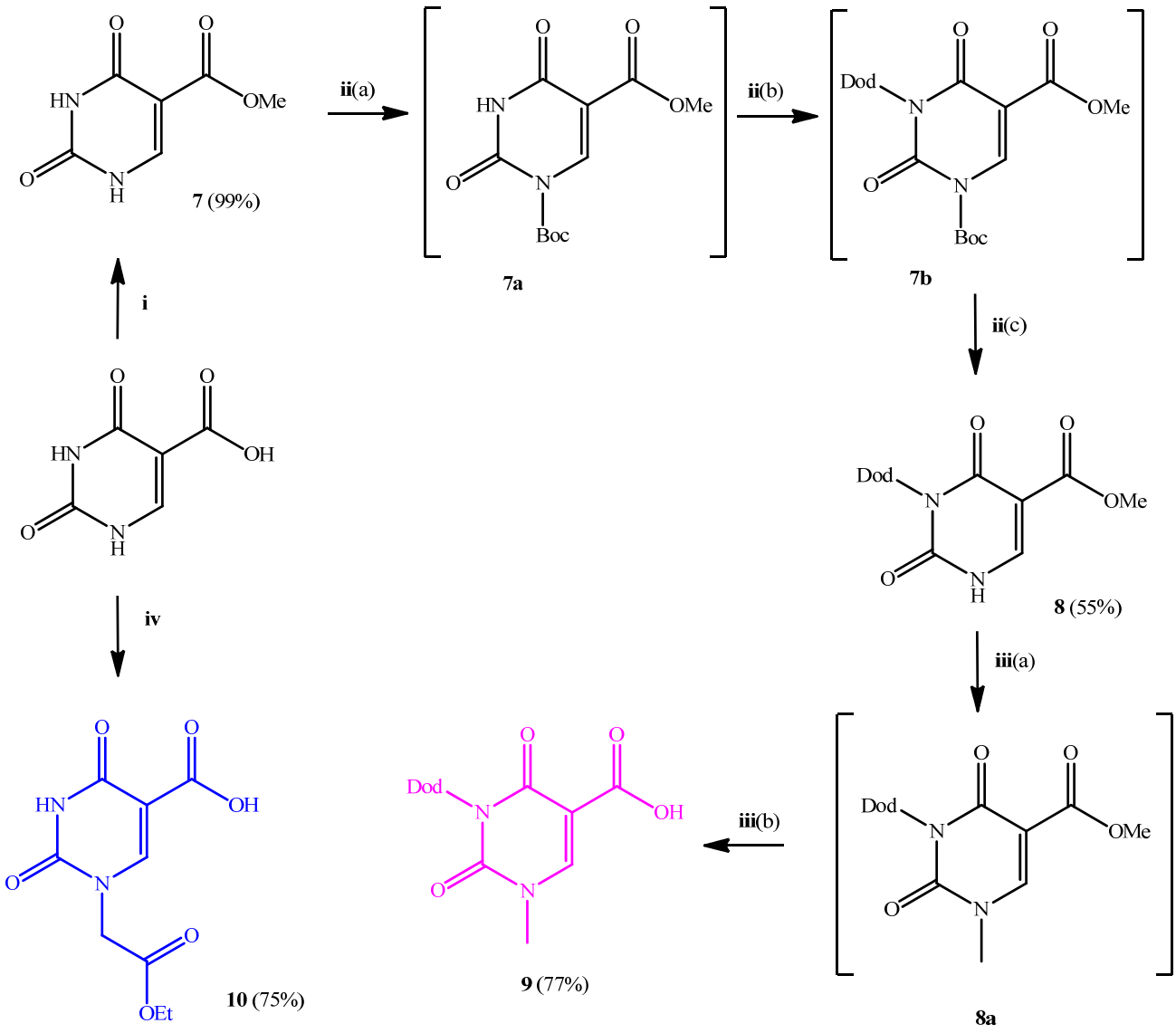
Scheme 4.1: Synthesis of the Fmoc-aeg-OMe backbone. (i) [(a) Fmoc-OSu, Na_2CO_3, dioxane, water, 75 min, rt; (b)KIO_4, acetone, water, 12h, rt]. (ii) [GlyOMe, $NaBH_3CN$, MeOH, 6h, 0°C to rt].

The synthesis of the central linker (Scheme 4.2) was carried out starting from 2,7-bis(bromomethyl)naphthalene that underwent a nucleophilic substitution with tritylamine.[272] Amine moieties were obtained by deprotection of trityl group under acidic conditions. In order to couple two different uracil units to this spacer, unsymmetrical protection of the two amino groups was needed; this was achieved by mono protection with Boc, using a mild Boc donor phenyl-t-nutylcarbonate (BocOPh), which guaranteed better yields compared to $(Boc)_2O$.



Scheme 4.2: Synthesis of the mono protected linker. (i) [(a)Tr-NH_2, K_2CO_3, MeCN, 96h, 55°C; (b) TFA, DCM, MeOH, 40 min, rt]. (ii) [BocOPh, EtOH, overnight, reflux].

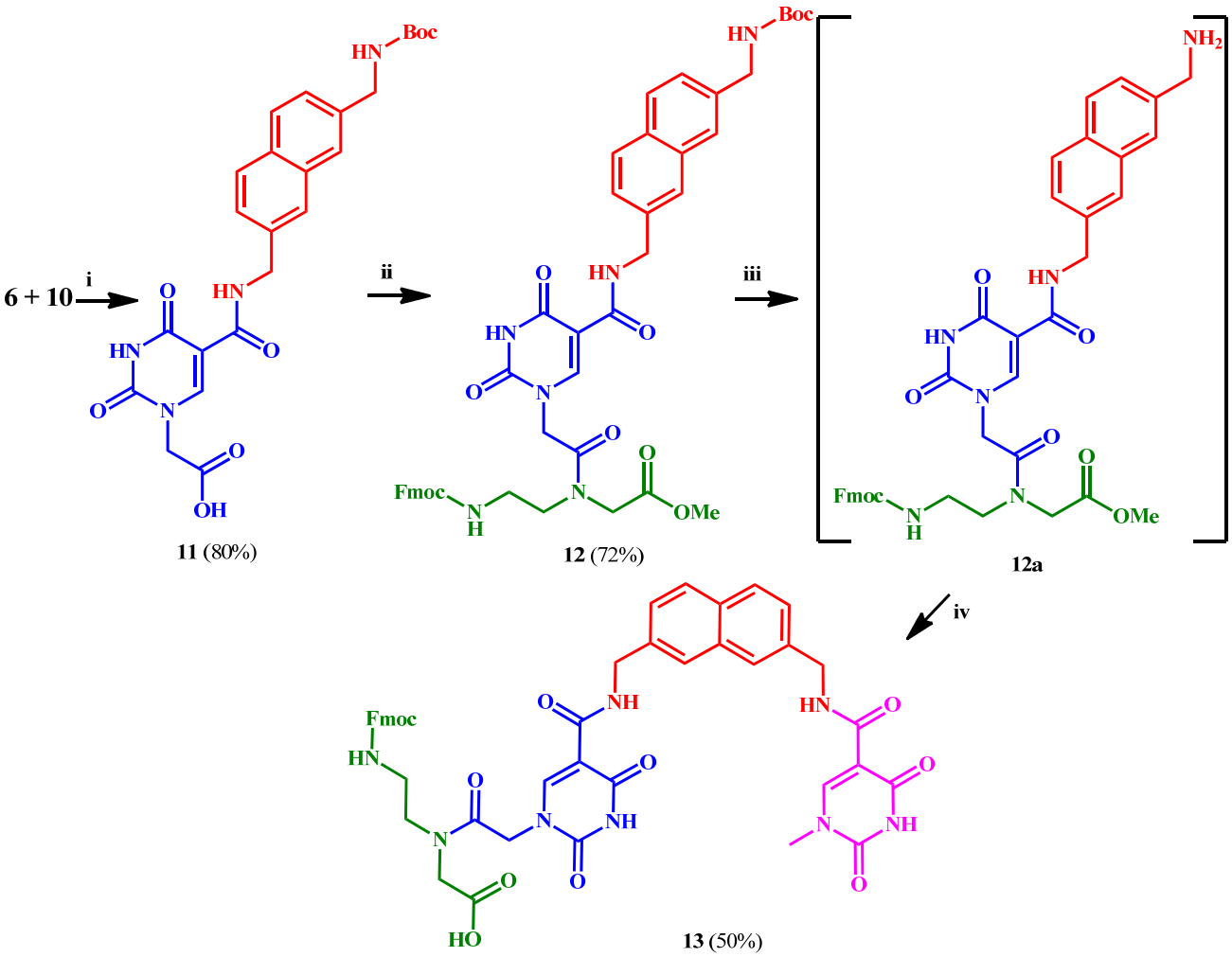
The two different uracil moieties were synthetized starting from isoorotic acid (Scheme 4.3). Selective methylation of N1 is difficult to achieve directly due to high reactivity of alkylating agents and to small steric hindrance of the methyl group. After esterification of isoorotic acid, alkylation proceeded then through Boc-protection of N1 nitrogen. The crude product **7a** was treated with Dod-Cl, thus protecting N3 position. N1 Boc showed an enhanced reactivity due to electron withdrawing effect of the uracil ring, allowing its removal under basic conditions.



Scheme 4.3: Synthesis of Uracil moieties. (i) [$SOCl_2$, MeOH, overnight]. (ii) [(a) Boc_2O, DMAP, MeCN, 5h, rt; (b) Dod-Cl, NaH, DMF, overnight, 0°C to rt; (c) K_2CO_3, MeOH, 2h, rt]. (iii) [(a)MeI, K_2CO_3, MeCN, 4h, rt; (b) 1M NaOH, MeOH, 2h, rt]. (iv) [(a)HMDS, TMS-Cl, 4h, reflux; (b) Br-CH_2COOEt, o.n., reflux; (c) water/AcOH 1:1, 20 min, rt].

The intermediate **8** was then methylated and final product **9** obtained for hydrolysis of methyl ester. **10** was easily obtained through direct alkylation of sylilated isoorotic acid with ethyl bromo acetate.[273]

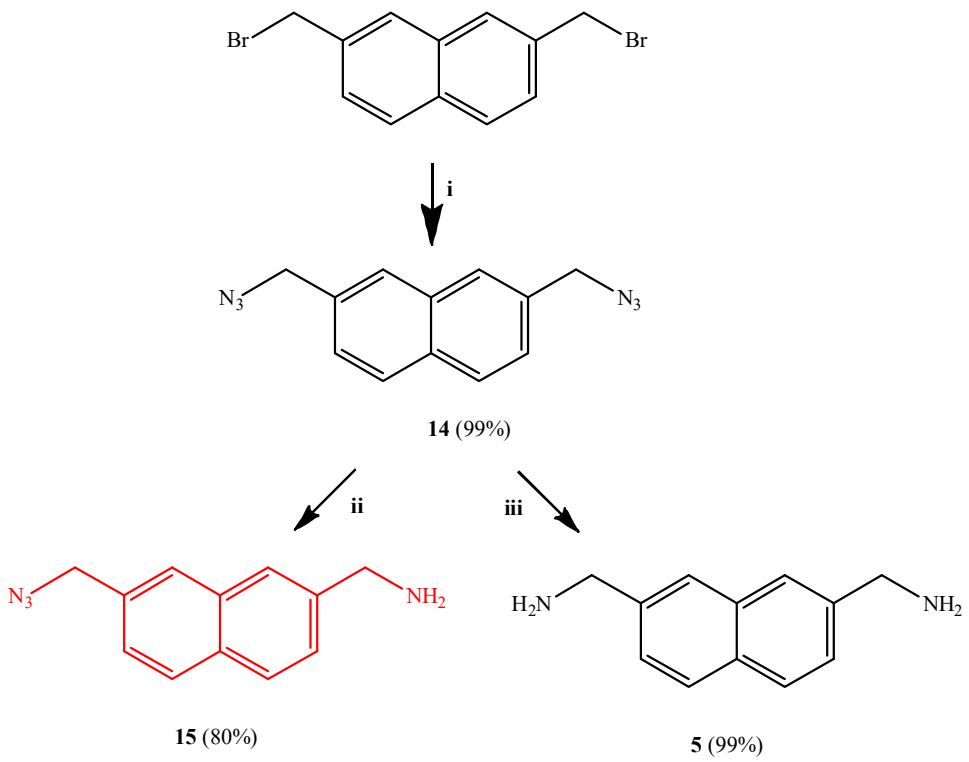
Linker **6** was coupled with uracil **10** using HBTU and then hydrolyzed to obtain **11**. This was coupled with backbone **3** using DCC/DhBtOH to afford **12**. Derivative **12** was deprotected using TFA and intermediate **12a** was coupled with **9** using HBTU. The final product **13** was obtained after hydrolysis using $Ba(OH)_2$ and controlling reaction time in order to avoid Fmoc deprotection (Scheme 4.4).



Scheme 4.4: Synthesis of dimeric base monomer. (i) [(a) HBTU, DIEA, DMF, 4h, rt; (b) 2M NaOH, THF, 90 min, rt]. (ii) [**3**, EDC, DhBtOH, DIEA, DMF, overnight, 0°C to rt]. (iii) [TFA/DCM 4:6, 1h, 0°C to rt]. (iv) [(a) HBTU, DIEA, DMF, 4h, rt;. (b) $Ba(OH)_2$, THF, 1.5h, rt].

4.3. Dimeric uracil base: new synthetic strategies

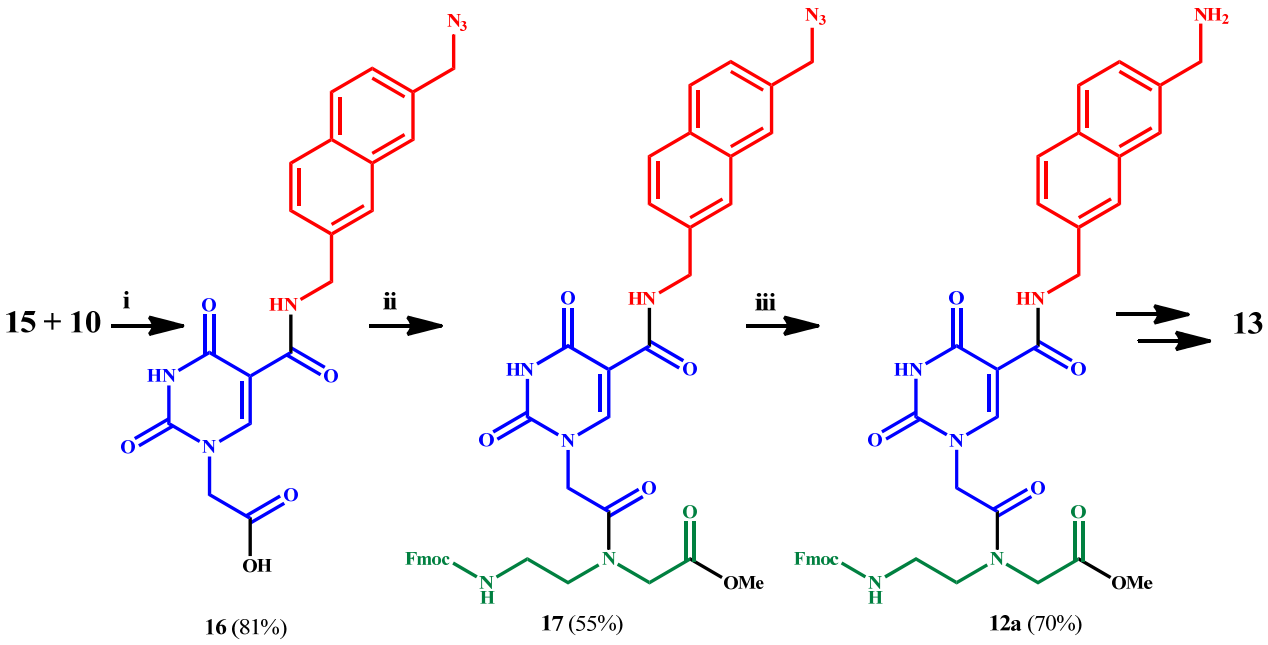
In the schemes of reaction described above, desymmetrization of central linker was made by mono protection of an amino group with Boc. Despite of the reduced activity of BocOPh, desired product was achieved in slightly more than statistical distribution; thus, we revised the synthetic strategy to allow an improved overall yield, using 2-methylazide-7-methylammino naphthalene as linker for obtaining final monomer **13** or for new solid phase strategy. These new synthetic routes were developed under the supervision of Dr. Alex Manicardi.



Scheme 4.5: synthesis of azido-amino linker. (i) [NaN_3, DMF, r.t., 30']; (ii) [PPh_3 in Et_2O, 1M HCl/AcOEt 1:1 (then THF/H_2O 1:1), r.t., 5h + 3h + o.n.]; (iii) [PPh_3, H_2O, THF, r.t., o.n.].

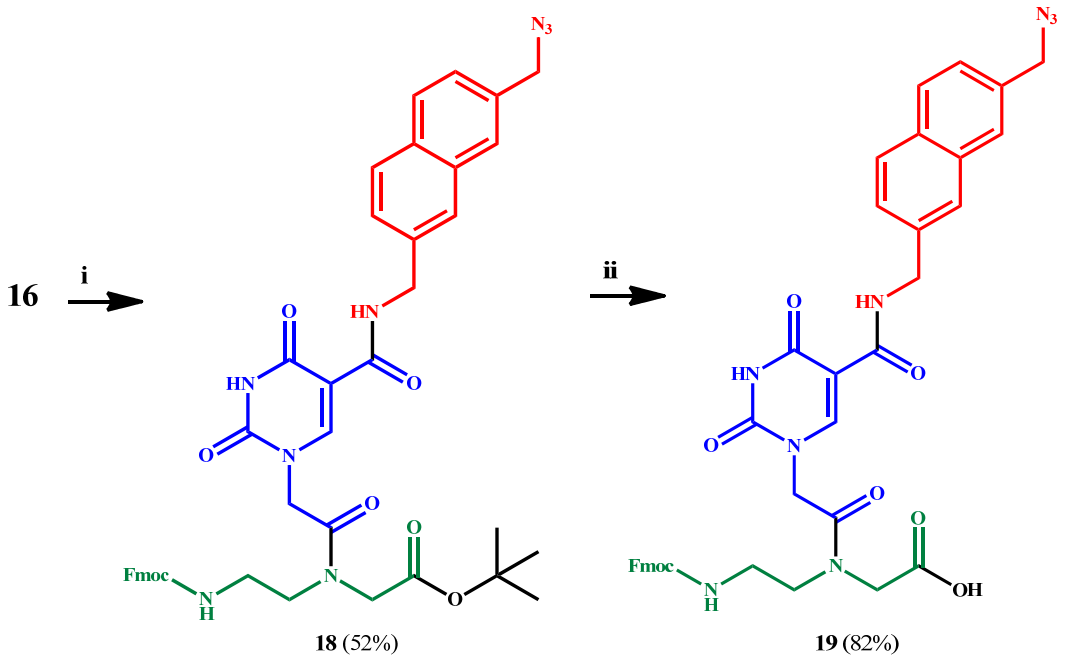
2,7-bis(bromomethyl)naphthalene readily reacted with sodium azide to give the product **14**. Sodium azide is much cheaper than tritylamine, leading in the same way to quantitative reaction. The key point is also in this case desymmetrization of intermediate (Scheme 4.5). Azide can be efficiently reduced to amine using Staudinger reaction.[274] The reaction proceeds through formation of phosphazide, which decomposes to an imminophsphorane (phosphazene), generating N_2. Reaction of the imminophsphorane with water leads then to desired amine product and to formation of phosphine oxide[275]. In order to obtain reduction of only one azide group, reaction was performed

in a biphasic system using a 1M water solution of HCl and ethyl acetate. Under vigorous stirring, intermediate **14** was dispersed in the biphasic solution and triphenylphosphine, dissolved in Et_2O, was slowly added. Reaction between azide and phosphine occurred in organic phase, and as soon as the imminophsphorane was reduced to amine the latter was protonated, due to acidic conditions, subtracting product from organic phase and avoiding the further reaction leading to the double reduction byproduct. In this way it was possible to obtain the product **15** with very high yield (80%). Following this reaction scheme was possible to achieve linker synthesis with a yield double of that obtained by the previous pathway (from 38% to 79%). Intermediate **14** can be used also in alternative synthesis of substrate **5**, through reduction of both azide groups. Compound **15** is useful since it presents a masked amino group that can be used for the synthesis of **13** using a modified version of the previous synthesis (Scheme 4.6).



Scheme 4.6: Synthesis of methyl/tert butyl intermediates. (i) [(a) HBTU, DIEA, DMF, 4h, rt; (b) 2M NaOH, THF, 90 min, rt]. (ii) [3, DCC, DhBtOH, DIEA, DMF, overnight, 0°C to rt]. (iii) [PPh_3, H_2O, THF, r.t, o.n.].

Indeed product **15** was coupled with **10** using HBTU to obtain **16**, after ester hydrolysis. Coupling with backbone led to **17**, a product analogue to **12** shown in Scheme 4.4. The azido group can then be reduced by Staudinger reaction obtaining the intermediate **12a**. Reactions to obtain final monomer are then the same discussed above.



Scheme 4.7: (i) [Fmoc-aeg-O^{t}Bu, EDC, DhBtOH, DIEA, DMF, overnight, 0°C to rt]. (ii) [TFA, DCM, 0 °C to r.t., 30 min].

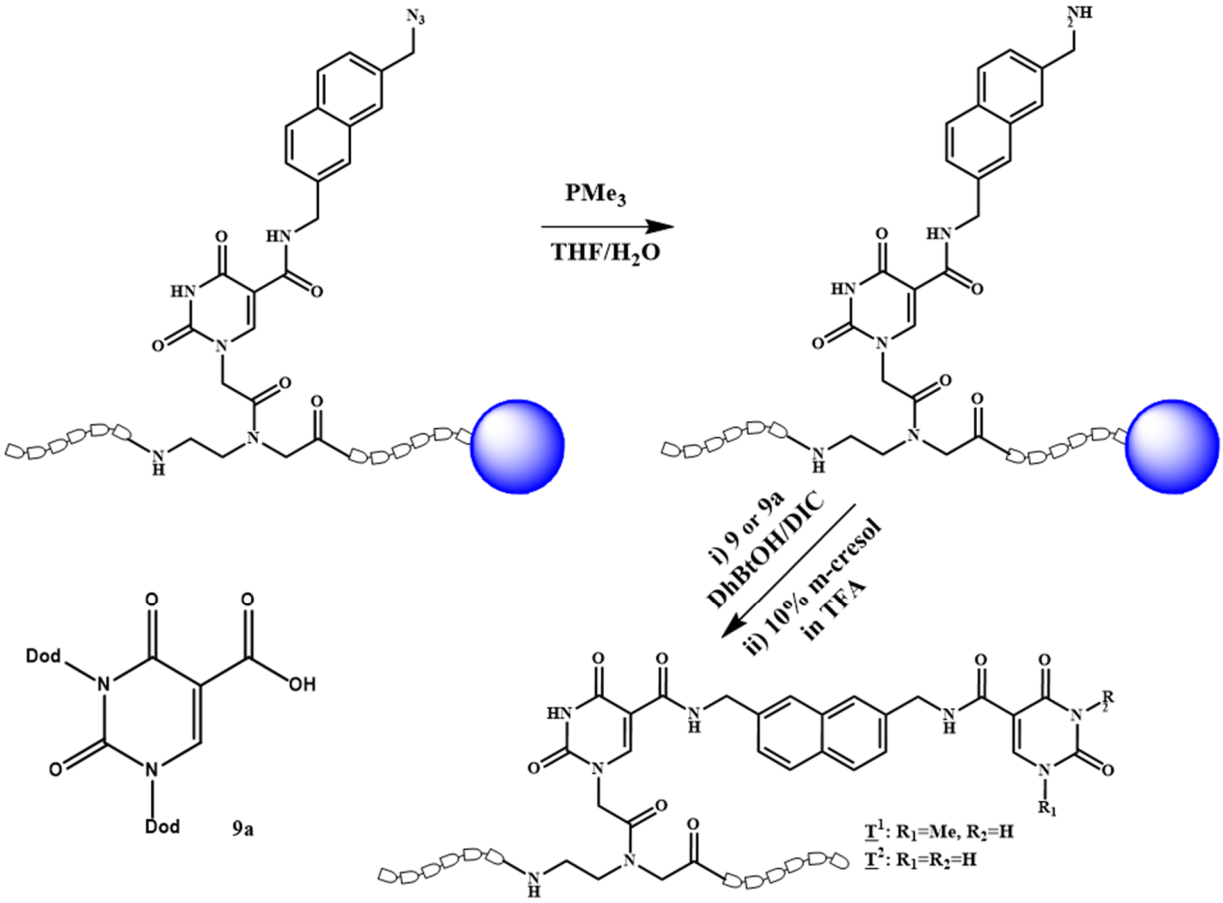
The intermediate **16** was coupled also with commercially available *N*-[2-(Fmoc-amino)-ethyl]glycine tert-butyl ester to give **18** (Scheme 4.7) that was then hydrolyzed to obtain compound **19**. This monomer can be used for a solid phase approach, in which second uracil is coupled directly on resin after PNA chain completion. Details about this synthetic strategy are reported in the following section (Scheme 4.8).

4.3.1. On resin assembly of the dimeric base.

The use of the monomer **13** in PNA solid-phase synthesis give rise to poor yields, due to high steric hindrance of the modified monomer; an alternative route can be to use the precursor **19** for the PNA synthesis, and then to obtain the dimeric base by insertion of the second uracil unit after the completion of the synthesis of the entire PNA sequence. Since we wanted initially to determine if self-association could be induced only by presence of AT doublet, we specifically designed a sequence containing this doublet without possibility of any other complementarity: H-TTATCC-Gly-NH_2. The synthesis of this PNA oligomer was performed using Fmoc strategy. Monomer **19** was inserted in the growing sequence of PNA using a longer coupling time (overnight instead of 30'). After insertion of modified monomer the sequence was completed using standard protocols. Fmoc was not removed at the end of the synthesis in order to allow the attack the second uracil moiety

on the solid phase (Scheme 4.8) without competing reactions. The azide group was reduced performing a Staudinger reaction on the solid phase, in which a crucial point was the use of trimethylphosphine in THF/water. Trimethylphosphine was chosen in place of commonly used triphenylphosphine due to its reduced steric hindrance and thus favoring access to reactive sites. Despite these measures, reduction of azide resulted difficult due to a formation of a biphasic system in the presence of water. Further studies suggested the possibility to reduce this problem by vigorously stirring the solution before adding it into reaction vessel. The reduced PNA was coupled with second uracil moiety using DhBtOH/DIC. We used two different uracils, compound **9** and compound **9a** (Scheme 4.8), which was obtained as byproduct of **8**.

Cleavage from the resin with 10% m-cresol in TFA allowed to recover PNA and, at the same time, to remove Dod protective group. The final products were obtained by two purification cycles with HPLC. First, the Fmoc-protected PNA was purified on RP-HPLC, then the Fmoc group was removed and a second purification was carried out, using RP-HPLC with different conditions.[276] Details are reported in experimental section.



Scheme 4.8: Solid phase synthesis of dimeric bases. Symbols in the chain represent PNA monomers. **T**[1] derives from coupling with **9** (**PNA1**, **PNA3**) and **T**[2] with **9a** (**PNA2**).

Due to problems in the reduction step described above, the two PNAs were obtained with low yields (1.1% in case of **T**^{1} and 0.5% in case of **T**^{2}). Modifications of this solid phase protocol, in particular refining Staudinger reaction conditions, must be introduced in future in order to achieve better yields. We synthetized also two unmodified PNA as control (**NM1**, Yield = 12%; **NM2**, Yield = 20%).

4.4. Solution properties of PNAs bearing modified dimeric base

PNAs containing a modified monomer with the dimeric uracil base (**13**) were tested in a previous work in our laboratory.[270] The sequences studied were H-GTAGA**T**^{1}CACT-NH_2 (**PR210**), which is a sequence widely used in the literature as a reference for comparing PNA properties as a function of structure,[106,113] and H-TCCT**T**^{1}CACT-NH_2 (**PR212**), which was used in the development of sensing approaches in order to recognize a single point mutation (W1282X) of the cystic fibrosis (CF) gene.[277,278] The central thymine was substituted with modified monomer **T**^{1} in both cases. Preliminary studies revealed tendency to self-association in the case of single strand **PR210**, but not in the case of **PR212** (Figure 4.3 top).

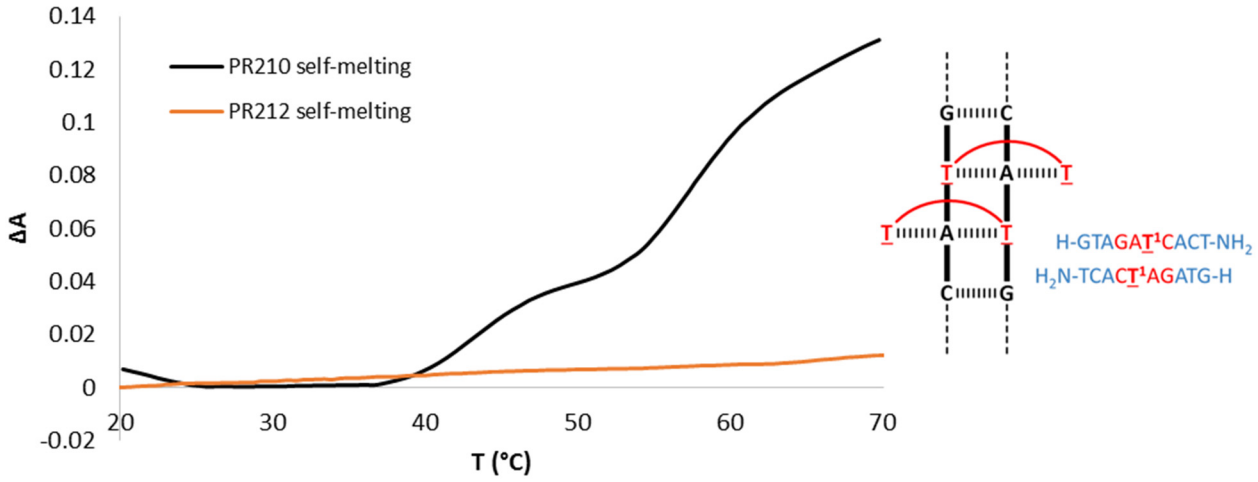


Figure 4.3: Top) UV melting profile expressed as variation of absorbance as function of temperature; Bottom) schematic representation of PR210 self-aggregation. **T** is the modified monomer **T**^{1}.

Sequence of **PR210** presented actually four complementary bases in the middle of the sequence, but so high melting temperatures are not typical of a 4 mer duplex and were probably due to the presence of dimeric base (Figure 4.3). **PR212** is not affected to this problem of partial complementarity, but contrary to what was expected, it didn't lead to great duplex stabilization (T_m +2 °C compared to unmodified PNA). This system exhibited anyway interesting properties of

selectivity toward single mismatched base.[270] We then wanted to better explore properties of PNA bearing this modified monomer using both experimental and computational approach (see next chapter).

To evaluate the presence of auto-association of PNA single strands we performed UV melting experiments on 5 µM **PNA1** and **PNA2** solutions in PBS buffer (Figure 4.4). Samples were annealed at 90 °C and then cooled at 10 °C.

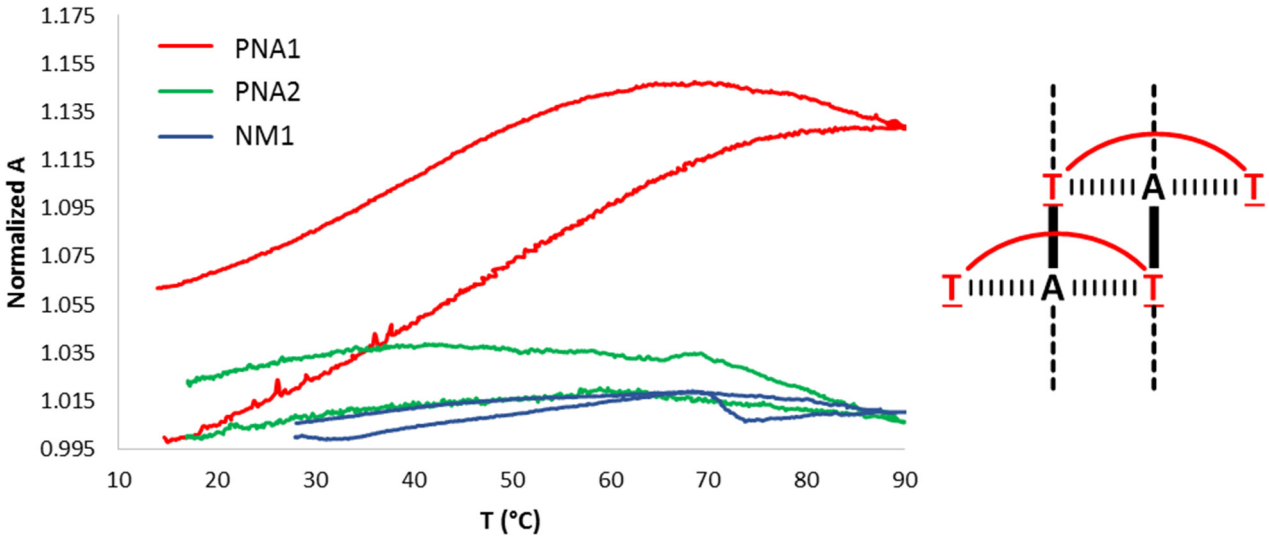


Figure 4.4: Left) UV melting curves of non-modified PNA (**NM1**) and PNAs containing modified base obtained from coupling with uracil **9** (**PNA1**) and **9a** (**PNA2**); right) schematic representation of **PNA1** self-aggregation.

PNA2 and **NM1** presented low absorbance variations as function of the temperature, not correlated to a melting. Irregularity in the shape of the curves was addressable to instrumental problems and not to the samples, as verified by several tests on different samples. **PNA1** melting temperatures were comparable in both heating and cooling cycles (T_m^{up} = 46.8 °C, T_m^{down} = 39.2 °C) and they resulted quite high considering that, in the case of self-association, were due to only two couples of bases (Figure 4.4 right). **PNA2** presented completely different properties compared to **PNA1**, even if the only difference was in the methyl group in N1 position. We supposed that interaction of N1 with solvent molecules could disturb Hoogsteen bonds of the second uracil and steric hindrance of methyl group could reduce water access. A 3D model of this difference is shown in Figure 4.5.

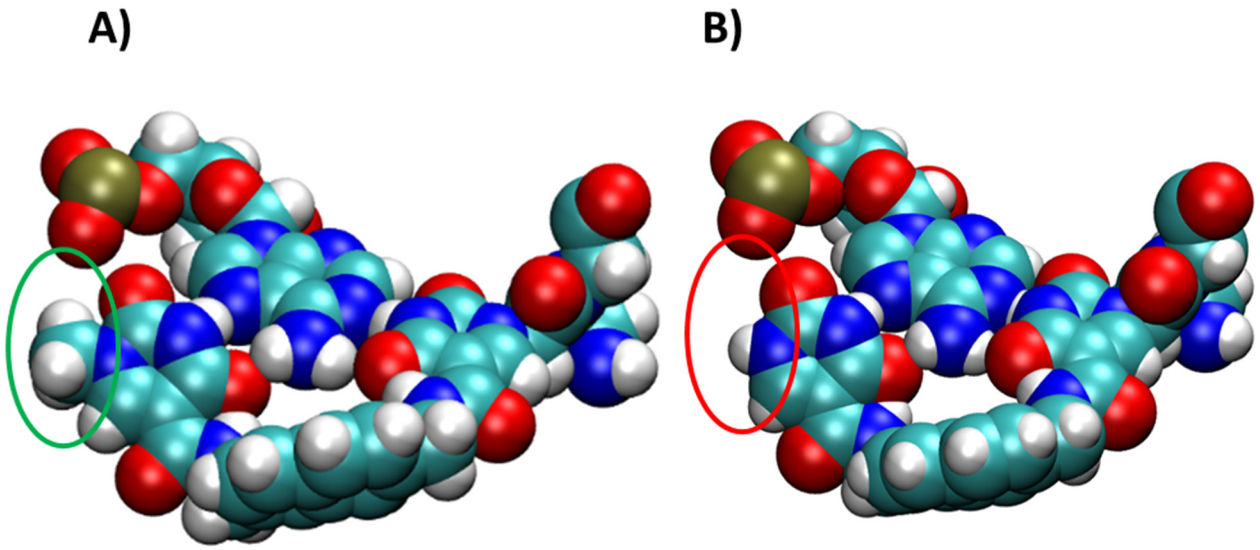
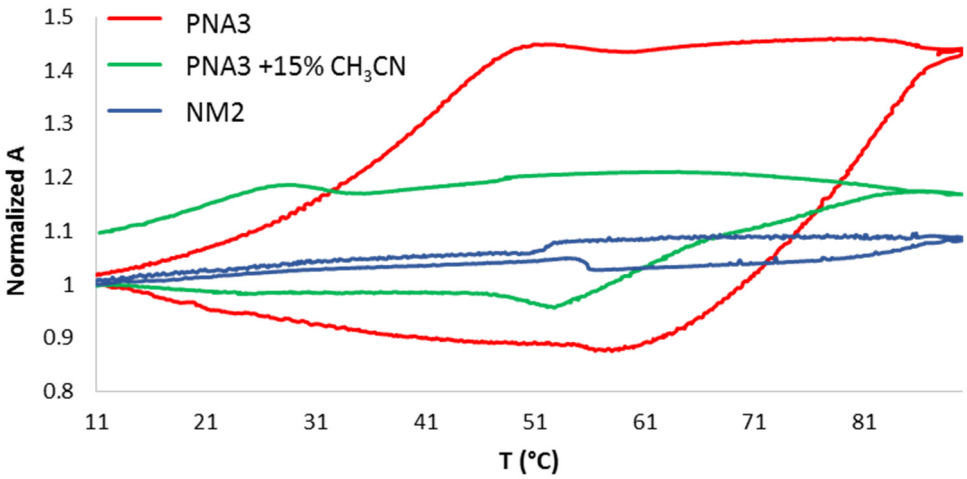


Figure 4.5: Dimeric base bound to adenine. **A**) monomer with uracil N1 protected; **B**) monomer with N1 not protected. Van der Waals radius representation of the atoms allows to recognize a protective function of methyl group attached to N1 compared to simple hydrogen.

We then wanted to observe if the same behavior was presented in a sequence with reversed doublet order (TA). In this case we have studied a sequence complementary to seed region of miR 221 and 222 (H-ATG**T**^{1}AGC-Gly-NH_2, **PNA3**). This target has a biological relevance since has been classified as onco-miR[279,280] in many different pathologies. This system allowed us to study both possible association effects and a real interesting target.

PNA3 was synthetized with the same strategy described above for **PNA1** and **PNA2**, using uracil **9**. The yield was in this case particularly low (0.2%) due to difficulties in purification with HPLC.

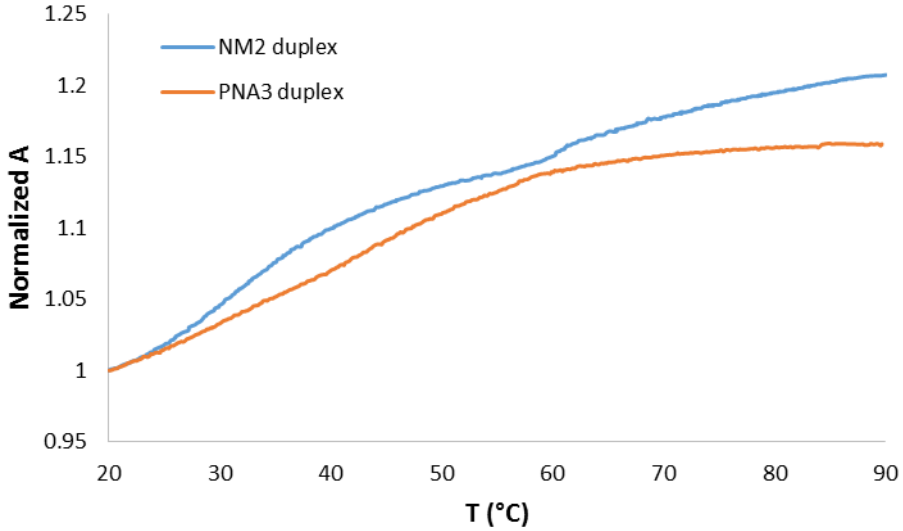
Unmodified PNA (**NM2**), chosen as reference, didn't show any transition and irregular shapes were due, also in this case to instrumental problems (Graph 4.1). **PNA3**, bearing modified base, presented a wide hysteresis compared to **PNA1** with melting temperatures far apart (T_m^{up} = 82.1 °C, T_m^{down} = 42.7 °C). After the heating-cooling cycle, there was the formation of a precipitate; heating again the sample to 90 °C afforded a complete transparent solution. We therefore believe that in this case there was a sum of effects, solubilization of the dimer and then melting.



Graph 4.1: UV melting curves of non-modified PNA (**NM2**) and PNA containing modified base (**PNA3**) in water and in a 15% solution of acetonitrile.

In order to better solubilize the complex we added 15% of acetonitrile to the sample. In this case solution appeared completely transparent. Melting curve presented a hysteresis similar to that one observed in only water, but of course shifted to lower values due to denaturizing effect of acetonitrile. The first tract of the curve is quite flat compared to the other. We can suppose that, without acetonitrile, dimer precipitates on the walls of the cuvette inducing a local increased absorbance that falls down with solubilization of the complex. After 60 °C, the acetonitrile easily evaporates, probably determining a reduction of this solvent content. Indeed, at the end of the experiment precipitate was found in the cuvette. Anyway, the presence of the hysteresis in the curve showed a low kinetics of aggregate formation.

For these PNAs (**NM2** and **PNA3**) we have studied also association with DNA. In this case, melting temperature of modified PNA is about 5 °C higher compared to unmodified one, confirming a limited, but reproducible, stabilizing effect of the dimeric base (Graph 4.2).



Graph 4.2: UV melting curves of **NM2** and **PNA3** duplexes with complementary DNA.

4.5. Conclusions

We proposed an alternative pathway for the synthesis of compound **13** allowing to achieve better yields for central linker unit. We proposed also a new synthetic strategy for the solid phase synthesis of PNAs containing dimeric base. This protocol must be improved in order to obtain better yields, in particular changing Staudinger reaction conditions, but is more flexible and can allow to introduce multiple variations in the structure starting from the same PNA precursor.

PNAs containing this type of modified monomer confirmed their tendency to self-associate in presence of AT doublet in the sequence. Melting temperatures of self-associated species were found to be high if one considers that pairing occurred only on this two couple of bases. Also the PNA bearing reversed doublet order (TA, **PNA 3**) showed evidences for self-association, though accompanied with low-solubility. This behavior must be considered in target selection, avoiding sequences that presents adenine close to modified thymine or choosing longer sequences in order to have competition between DNA:PNA duplex and PNA:PNA aggregate. However, it is possible to imagine a possible use of this dimeric base as non-covalent, self-assembly unit (Figure 4.6).

Study on PNA:DNA heteroduplex showed a modest stabilization induced by dimeric base compared to unmodified PNA, thus confirming the results previously obtained with a different sequence (**PR212**).

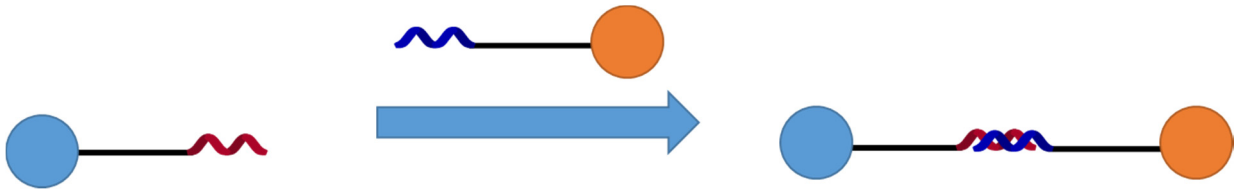


Figure 4.6: Schematic representation of self-assembly induced by AT doublet.

Though this stabilizing effect (and even more, the high sequence selectivity) of this dimeric base can be important in biological applications, it was not as large as expected. For this reason, in the next chapter, an attempt to address the design of this system by molecular modelling is reported.

4.6. Experimental section

General. Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, LINK and used without further purifications. TLC were run on Supelco 56524-25EA silica gel on aluminium foils with fluorescence indicator 254 nm. Column chromatography were performed as flash chromatography on Merck silica 60 (0.040 – 0.063) under 0.1-0.2 bar of air overpressure. NMR spectra were obtained on Bruker Avance 300 MHz or 400 MHz spectrometers. δ values are expressed in ppm. FT-IR were recorded on Thermo Nicolet 5700. ESI-MS spectra were recorded on a Acquity Ultra Performance LC. Meting points were measured with GALLENKAMP melting point apparatus.

Synthesis and characterization of products from **2** to **13** were carried out using the conditions reported in previous PhD thesis work.[270] Yields were reevaluated and were found to be those reported in the results and discussion part.

1,3-bis (bis (4- methoxyphenyl) methyl)-2,4- dioxo-1,2,3,4- tetrahydropyrimidine-5- carboxylic acid (9a). Ester of compound **9a** was obtained as a byproduct of compound **8** synthesis as a part of Boc deprotection of intermediate **7b** (Yield 35%, 417 mg). Hydrolysis was performed by action of 1 M NaOH (aq) and THF 1:1 during 2 h reaction. Final product was precipitated by adjusting pH of solution to 3 with 1 M HCl. White solid was filtrated and dried under vacuum (88%, 408.4 mg).

(9a). ^{1}H-NMR (300 MHz, 25°C, d_6-DMSO) δ(ppm): 7.87 (s, 1H, CH uracil), 7.16 (s, 1H, CH Dod), 7.14 (d, 4H, J = 8.8 Hz, CH aromatic Dod), 7.07 (d, 4H, J = 8.8 Hz, CH aromatic Dod), 6.97 (d, 4H, J = 8.9 Hz, CH aromatic Dod), 6.84 (d, 4H, J = 8.8 Hz, CH aromatic Dod), 6.80 (s, 1H, CH Dod), 3.77 (s, 6H, OMe), 3.73 (s, 6H, OMe). ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ(ppm): 164.2, 161.7, 159.4, 158.6, 150.7, 145.3, 130.8, 130.3, 130.0, 127.8, 114.8, 114.2, 114.1, 113.8, 109.7, 62.9, 58.0, 55.6, 55.5. FT-IR ν (cm^{-1}) 2930, 2837, 1749, 1718, 1627, 1610, 1585, 1511, 1462, 1441, 1419, 1391, 1304, 1248, 1175, 1113, 1031. m.p.: decomposes at 190 – 200 °C.

2,7-bis (azido methyl) naphthalene (14). 2,7-bis(bromomethyl)naphthalene (800 mg, 2.55 mmol) was dissolved in dry DMF (7 ml) and then NaN_3 (347.8 mg, 5.60 mmol) was added to the solution.

After 2.5 h the reaction was completed (TLC on silica gel, Hex-AcOEt 9:1, R_f = 0.43), reaction mixture was diluted with AcOEt (250 ml) and washed with brine (3 X 225 ml). Organic layer was dried over Na_2SO_4 and evaporated under vacuum to afford pure product (light brown solid, 99%, 607 mg).

(14) ^{1}H-NMR (400 MHz, 25°C, d_6-DMSO) δ(ppm): 7.98 (d, 2H, J = 8.5 Hz, CH napht.), 7.94 (s, 2H, CH napht.), 7.53 (dd, 2H, J_1 = 8.5 Hz, J_2 = 1.7 Hz, CH napht.), 4.64 (s, 4H, CH_2); ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ(ppm): 134.29, 133.07, 132.57, 128.81, 127.63, 127.19, 54.16. Compound decomposes in mass; MS (ESI+) m/z: found 406.11, 235.26, 211.6. FT-IR ν (cm^{-1}) 2908, 2852, 2101, 1635, 1510, 1441, 1429, 1370, 1325, 1267, 1175, 1145. m.p. 84 – 86 °C.

2-azidomethyl-7-aminomethyl naphthalene (15). Compound **14** (300 mg, 1.21 mmol) was dissolved in a mixture of 1 M HCl (150 ml) and AcOEt (75 ml). Solution was vigorously stirred and triphenyl phosphine (313.8 mg, 1.15 mmol), dissolved in diethyl ether (75 ml), was slowly added over a period of 5 h. Mixture was kept stirring for additional 3 h followed by organic solvent removal under vacuum. THF (150 ml) was added and reaction proceeded overnight. Organic solvent was removed and water phase was washed with ether (3 x 150 ml). Water phase was removed and residue dissolved in the minimum amount of 1 M HCl (~ 100 ml). pH was adjusted to about 12 with NaOH pellets and mixture kept in refrigerator for 2 h. Precipitate (white solid) was collected to afford pure product (138 mg). The remaining product (65.4 mg) was obtained by flash chromatography on silica gel (AcOEt + 1% NH_3, R_f = 0.2 in AcOEt-MeOH 9:1). (total yield: 80%, 203.4 mg).

(15) ^{1}H-NMR (300 MHz, 25°C, MeOD) δ(ppm): 7.94 (dd, 1H, J = 1.8, CH napht.), 7.92 (s, 1H, CH napht.), 7.87 (bs, 1H, CH napht.), 7.56 (dd, 1H, J_1 = 8.4, J_2 = 1.8, CH napht.), 7.51 (dd, 1H, J_1 = 8.4, J_2 = 1.8, CH napht.), 4.56 (s, 2H, CH_2-N_3), 4.15 (s, 2H, CH_2-NH_2). ^{13}C-NMR (75 MHz, 25°C, MeOD) δ(ppm): 128.2, 128.1, 126.8, 126.7, 126.2, 126.0, 54.3, 44.1. MS (ESI+) m/z: $[M + H]^+$ calcd. 213.11, found 213.18. FT-IR ν (cm^{-1}) 3348, 3270, 2908, 2853, 2107, 1632, 1560, 1511, 1473, 1440, 1430, 1388, 1352, 1326, 1278, 1171, 1144. m.p. 93 – 95 °C.

2,7-bis(aminomethyl) napthalene (5) (alternative synthesis)**.** Compound **14** (301.8 mg, 1.27 mmol) was dissolved in THF (20 ml) in presence of triphenyl phosphine (697.8 mg, 1.267 mmol) and water (114 μl, 6.34 mmol). To complete the reaction, after 4 h, additional amount of triphenyl phosphine (348.9mg, 0.63 mmol) and H_2O (228 μl, 12.67 mmol) were added and mixture stirred overnight. THF

was removed under vacuum, crude dissolved in 2 M HCl (150 ml) and washed with DCM (3 X 100 ml) to eliminate phosphine. Then, pH was adjusted to about 12 and product extracted with DCM (7 x 200 ml). Organic layers were combined, dried over Na_2SO_4 and evacuated under vacuum to afford pure product (99%, 236 mg).

(5). ^{1}H-NMR (300 MHz, D_2O, 25 °C) δ (ppm) 4.33 (s, 4H), 7.59 (dd, 2H, J_1 = 8.5 Hz, J_2 = 1.7 Hz), 7.98 (d, 2H, J = 1.7 Hz), 8.02 (d, 2H, J = 8.4 Hz). ^{13}C-NMR (75 MHz, D_2O, 25°C) δ (ppm) 46.5, 130.2, 131.6, 132.3, 134.5, 136.1, 136.2. Full characterization of this compound is reported in pervious thesis work.[270]

2-(5-(((7-(azidomethyl)naphthalen-2-yl)methyl)carbamoyl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl) acetic acid (16). Compound **10** (171.2 mg, 0.71 mmol) was dissolved in dry DMF at 0°C, under nitrogen atmosphere. HBTU (268.1 mg, 0.71 mmol) and DIPEA (246.1 μl, 1.41 mmol) were added and stirred at 0 °C for 15 min and at room temperature for further 15 min. Compound **15** (100 mg, 0.471 mmol) was then added and mixture stirred overnight at room temperature. At the end of the reaction, solution was partitioned between saturated $KHSO_4$ (200 ml) and AcOEt (200 ml), and then organic phase washed with $NaHCO_3$ and brine. Solvent was evaporated under vacuum and crude dissolved in a mixture of methanol (6 ml) and 1 M NaOH (6 ml). Solution was stirred for 4 h and then pH adjusted to 1 using 1 M HCl. Precipitate was separated on a Büchner funnel and dried under vacuum to afford pure product (81 %, 140 mg).

(16) ^{1}H-NMR (400 MHz, 25°C, d_6-DMSO) δ(ppm): 13.31 (bs, 1H, CO<u>OH</u>), 12.05 (s, 1H, NH uracil), 9.24 (t, 1H, J = 6.1 Hz, NH amide), 8.64 (s, 1H, CH uracil), 7.94 (d, 1H, J = 8.5 Hz, CH naphth.), 7.91 (d, 1H, J = 8.7 Hz, CH napht.), 7.89 (s, 1H, CH napht.), 7.80 (s, 1H, CH napht.), 7.49 (dd, J = 8.0, 1.6 Hz, 1H, CH napht.), 7.47 (dd, J = 8.2, 1H,1.6 Hz, CH napht.), 4.68 (d, 2H, J = 6 Hz, <u>CH_2</u>NH), 4.63 (s, 2H, <u>CH_2</u>N_3), 4.61 (s, 2H, <u>CH_2</u>COOH). ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ(ppm): 169.7, 164.1, 162.2, 152.3, 150.5, 138.0, 133.9, 133.2, 132.1, 128.7, 128.4, 127.5, 126.7, 126.6, 125.8, 105.2, 54.2, 49.8, 42.7. FT-IR ν (cm^{-1}) 3011, 2818, 2125, 1706, 1632, 1605, 1555, 1460, 1420, 1381, 1331, 1268, 1234. Decomposes at 270 – 280 °C.

Methyl 2 -(N-(2-((((9H-fluoren-9-yl) methoxy) carbonyl) amino) ethyl)-2-(5-(((7-(azidomethyl) naphthalen -2-yl)methyl)carbamoyl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetate (17). Compound **16** (70 mg, 0.17 mmol) was dissolved in dry DMF under nitrogen atmosphere and stirred at 0 °C with EDC (32.9 mg, 0.17 mmol) and DhBtOH (27.9 mg, 0.17 mmol) for 15 min and at r.t. for additional 15 min. Backbone **3** (72.9 mg, 1.2 eq) was then added together with DIPEA (34 µl, 1.2 eq) and mixture stirred overnight. Solution was diluted with AcOEt (150 ml) and washed with saturated $KHSO_4$ (2 x 150 ml), saturated K_2CO_3 (2 x 150 ml) and brine (1 x 150 ml). Organic layer was dried under vacuum and crude purified by flash chromatography on silica gel (AcOEt-Hexane 8:2) to afford pure product (55 %, 70 mg).

(17) ^{1}H-NMR (400 MHz, 25°C, d_6-DMSO) δ(ppm): 12.00 (s, 1H, NH uracil), 9.25 (t, 1H, J = 6 Hz, NH amide), 8.48 (M) and 8.39 (m) (s, 1H, CH uracil), 7.92 (m, 2H, CH napth), 7.89 (d, 2H, J = 7.1 Hz, CH aromatic Fmoc), 7.88 (s, 1H, CH napth.), 7.80 (s, 1H, CH napth.), 7.69 (bd, 2H, J = 7.3 Hz, CH aromatic napth.), 7.49 (dd, 1H, J_1 = 6.5 Hz, J_2 = 1.5 Hz, CH napth.), 7.47 (dd, 1H, J_1 = 6.5 Hz, J_2 = 1.6 Hz, CH napth.), 7.41 (t, 2H, J = 7.4 Hz, CH aromatic Fmoc), 7.32 (td, 2H, J_1 = 7.4 Hz, J_2 = 0.7 Hz, CH aromatic Fmoc), 7.27 (t, 1H, J = 5.9 Hz, NH Fmoc), 7.95 (M) and 4.76 (m) (s, 2H, CH_2 linker), 4.68 (d, 2H, J = 5.8 Hz, Napth.-$\underline{CH_2}$-NH), 4.61 (s, 2H, $\underline{CH_2}N_3$), 4.35 (M) and 4.29 (m) (d, 2H, J_M = 6.8 Hz, J_m = 3.6 Hz, CH_2 Fmoc), 4.27 – 4.18 (m, 1H, CH Fmoc), 4.32 (m) and 4.08 (M) (s, 2H, CH_2COOMe), 3.73 (m) and 3.63 (M) (s, 3H, Me), 3.45 (M) and 3.37 (m) (t, 2H, J_M = 6.4 Hz, J_m = 6.8 Hz, CH_2 backbone), 3.27 (M) and 3.09 (m) (q, 2H, J_M = 6.1 Hz, J_m = 6.1 Hz, CH_2 Fmoc). ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ(ppm): 170.2 (m) and 169.9 (M), 167.7 (m) and 167.5 (M), 164.1, 162.3, 156.8 (M) and 156.6 (m), 152.6 (M) and 152.4 (m), 150.5, 144.4, 141.2, 138.0, 133.9, 133.2, 132.1, 128.7, 128.4, 128.1, 127.5, 126.7, 126.6, 125.8, 125.6, 120.6, 105.0, 66.0, 54.2, 52.8 (m) and 52.3 (M), 49.2, 49.0 (M) and 48.7 (m), 48.2, 47.4 (m) and 47.2 (M), 42.7, 38.7 (M) and 38.1 (m). MS (ESI+) m/z: $[M + H]^+$ calcd. 745.27, found 745.31. FT-IR ν (cm^{-1}) 2818, 2097, 1718, 1687, 1683, 1611, 1538, 1464, 1450, 1375, 1341, 1239, 1213, 1145, 1108, 1003. m.p. 155 – 157 °C.

Tert-butyl 2-(N-(2-((((9H-fluoren-9-yl) methoxy) carbonyl) amino) ethyl)-2-(5-(((7-(azidomethyl) naphthalen-2-yl) methyl) carbamoyl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl) acetamido) acetate (18).[317,318] Compound **18** was prepared analogously to product **17** but using commercial backbone with tert-butyl ester. Purified with flash chromatography on silica gel (DCM-MeOH 98:2) to afford pure product (52%, 185.4 mg).

(18) ^{1}H-NMR (300 MHz, 25°C, $CDCl_3$) δ(ppm): 9.04 (t, 1H, J = 9 Hz, NH amide uracil), 8.46 (s, 1H, NH uracil), 8.38 (s, 1H, CH uracil), 7.90-7.70 (m, 6H, CH napth. and CH aromatic Fmoc), 7.62 (t, 2H, J = 7.3 Hz, CH aromatic Fmoc), 7.50-7.37 (m, 4H, CH napth. and CH aromatic Fmoc), 7.32 (d, 2H, J = 7.5 Hz, CH aromatic Fmoc), 5.90 (s, 1H, NH Fmoc), 4.77 (d, 2H, J = 5.6 Hz, CH_2 napth.), 4.65 (s, 1H, CH_2 linker), 4.50 (s, 2H, CH_2N_3), 4.39 (d, 1H, J = 7.2 Hz, CH Fmoc), 4.30-4.18 (m, 2H, CH_2 Fmoc), 3.96 (s, 1H, $\underline{CH_2}COO^tBu$), 3.64-3.33 (m, 4H, $NH\underline{CH_2}\underline{CH_2}NH$), 1.48 (s, 9H, ^{t}Bu); ^{13}C (100 MHz, 25°C, $CDCl_3$) δ(ppm): 157.0, 143.8, 141.3, 127.7, 127.0, 125.0, 120.0, 66.8, 53.46, 47.2, 42.1, 40.6, 40.3, 27.8, 27.7, 23.4. NMR signals are relative to major rotamer. MS (ESI+) m/z: $[M + Na]^+$ calcd 809.3, found 809.3.

Methyl 2-(N-(2-((((9H-fluoren-9-yl) methoxy) carbonyl) amino) ethyl)-2-(5-(((7-(aminomethyl) naphthalen-2-yl) methyl) carbamoyl)-2,4- dioxo-3,4- dihydropyrimidin-1(2H)-yl) acetamido) acetate (12a). Compound **17** (30 mg, 0.04 mmol) was dissolved in THF (1 ml) in a tube and stirred with PPh_3 (11.6 mg, 0.04 mmol) and water (36 µl, 2.00 mmol) at room temperature overnight. Then TFA (1 ml) was added and solvent removed under vacuum. Residue was dissolved in the minimum amount of TFA and then the tube was filled with Et_2O. Precipitate was centrifuged and washed two time with ether. White solid was dried under vacuum to afford pure product (70%, 20 mg).

(12a) ^{1}H-NMR (400 MHz, 25 °C, d_6-DMSO) δ(ppm): 12.00 (s, 1H, NH uracil), 9.22 (t, 1H, J = 6.1 Hz, NH amide uracil), 8.49 (s, 1H, CH uracil), 8.00 – 7.80 (m, 4H, CH naphth. and CH aromatic Fmoc), 7.72 – 7.64 (m, 4H, CH naphth. and CH aromatic Fmoc), 6.61 (1H, t, J = 5.5 Hz, NH amide Fmoc), 4.95 (s, 2H, CH_2 linker), 4.66 (d, 2H, J = 5.5 Hz, $NH\underline{CH_2}$ naphth.), 4.41 (d, 2H, J = 4.6 Hz, $\underline{CH_2}NH_2$), 4.35 (d, 2H, J = 6.7 Hz, CH_2 Fmoc), 4.24 (t, 1H, J = 6.6 Hz, CH Fmoc), 4.08 (s, 2H, $\underline{CH_2}COOMe$), 3.84 (bs, 2H, NH_2), 3.62 (3H, s, OMe), 3.44 (t, 2H, J = 5.9 Hz, $CH_2\underline{CH_2}N$), 3.32 – 3.23 (m, 2H, $\underline{CH_2}CH_2N$). ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ(ppm): 169.9, 167.5, 164.1, 162.3, 158.7, 156.8, 152.6, 150.5, 144.3, 141.2, 139.3, 137.5, 133.3, 131.6, 128.3, 128.1, 127.5, 126.1, 125.9, 125.6, 125.2, 120.6, 105.0, 66.0, 52.3, 49.2, 49.0, 48.3, 47.2, 43.6, 42.7, 38.8. NMR signals are relative to major rotamer. MS (ESI+) m/z: $[M + H]^+$ calcd. 719.3, found 719.3. FT-IR ν (cm^{-1}) 3314, 1686, 1613, 1543, 1466, 1450, 1380, 1340, 1237, 1203, 1180, 1144, 1006. Decomposes at 235 – 236 °C.

2-(N-(2-((((9H-fluoren-9-yl) methoxy) carbonyl) amino) ethyl)-2-(5-(((7-(azidomethyl) naphthalen-2-yl)methyl) carbamoyl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl) acetamido) acetic acid (19).

Compound **18** (185.4 mg, 0.24 mmol) was dissolved in DCM (6 ml) and then TFA (2 ml) was added to the solution at 0 °C. Solution was stirred for 30 min and for further 150 min at room temperature. Reaction was quenched with MeOH (10 ml) and solvent evaporated under vacuum to afford pure product product (82%, 172.2 mg).

(19) ^{1}H-NMR (400 MHz, 25 °C, d_6-DMSO) δ(ppm): 11.96 (s, 1H, OH), 9.20 (s, 1H, NH), 8.45 (s, 1H, CH uracil), 7.90 (d, 2H, J = 8.00 Hz, CH napht.), 7.87 (s, 2H, CH Fmoc), 7.68 (s, 2H, CH napth.), 7.64 (s, 2H, CH Fmoc), 7.38 (d, 2H, J = 1.65 Hz, CH napth.), 7.37 (s, 2H, CH Fmoc), 7.29 (s, 2H, CH Fmoc), 4.90 (s, 2H, NCOOH), 4.71 (s, 2H, CH_2 Fmoc), 4.64 (t, 1H, J = 4.00 Hz, CH Fmoc), 4.56 (d, 2H, J = 5.9 Hz, NH$\underline{CH_2}$), 4.16 (s, 2H, N$\underline{CH_2}$), 3.38 (t, 2H, J = 4.00 Hz, N$\underline{CH_2}CH_2$NH), 3.23 (q, 2H, J = 4.00 Hz, NCH$_2\underline{CH_2}$NH), 2.92 (s, 2H, N$\underline{CH_2}$CO), 2.28 (s, 2H, $\underline{CH_2}$N3); ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ(ppm): 170.7, 167.2, 164.1, 156.8, 150.5, 144.3, 141.2, 138.0, 133.9, 133.1, 128.7, 128.4, 128.0, 127.5, 126.7, 125.8, 125.6, 120.5, 79.6, 65.9, 54.2, 49.1, 48.9, 48.2, 47.3, 47.1, 29.4. NMR signals are relative to major rotamer. MS (ESI+/-) m/z: $[M - H]^-$ calcd. 729.25, found 729.04, $[M + H]^+$ calcd. 731.26, found 731.23, $[M + Na]^+$ calcd. 753.24, found 753.16, $[M + K]^+$ calcd. 769.21, found 769.13. FT-IR ν (cm^{-1}) 2948, 2360, 1682, 1611, 1540, 1465, 1450, 1416, 1381, 1339, 1180, 1136, 1007. m.p. 208 – 210 °C.

PNA oligomers synthesis. PNA oligomers were synthetized on solid phase using Fmoc strategy. The synthesis was performed manually using plastic vessel equipped with polyethylene frit, cap and valve. Rink amide MBHA resin was downloaded before use to 0.2 mmol/g of active sites using Fmoc-Gly-OH. Unreacted active sites were capped with acetic anhydride (DMF/Ac_2O 1:1) and exact loading determined by measuring UV absorbance at 290 nm of small amounts of resin dispersed in a 20% piperidine in DMF solution. Loading, expressed in mmol/g, was given from following formula:

$$Loading = \frac{1.68}{Abs(290nm)\ X\ mg(resin)}$$

Chain elongation occurred with the following protocol: 1) Fmoc deprotection with 20% piperidine solution in DMF for 8 min 2 times; 2) activation of 5 eq of PNA monomer (respect resin active sites) with 4.9 eq HBTU, 10eq DIEA for 2 min; 3) Coupling 30 min for standard monomers and overnight for monomer **20**; 4) Capping with DMF/DIEA/Ac_2O 89/6/5 1 min for 2 times; 5% DIEA in DMF washings 2 min for 2 times. PNA was cleaved with 10% m-cresol in TFA, 1.5 h (twice). The PNA was precipitated from the TFA solution with ethyl ether (10 times TFA volume) in freezer for 2 h and washed three times. Terminal Fmoc group was removed in 20% piperidine solution after first cycle of purification.

Oligomers purification. Crude PNA were purified using HPLC (Agilent Technologies 1260 Infinity, G1311C 1260 Quat Pump VL, G1316A 1260 TCC thermostat, G1314B 1260 VWD VL UV detector). Purification protocol and column changed depending on the PNA. Water +0.1% TFA (A) and acetonitrile +0.1% TFA (B) were used as eluent. Flow was set to 4 ml/min.

Unmodified PNA (NM1, NM2) were purified on reversed phase semi preparative C18 column (Xterra® Prep RP18, 10 μm, 7.8 x 300 mm, 300Å), at 35 °C using the gradient reported in Appendix 8.4.

PNA bearing modifications were separated on reversed phase semi preparative C18 column (Phenomenex® Jupiter C18 300Å, 5 μm, 10 x 250 mm), at 50 °C using gradients described in Appendix 8.4.

Characterization of pure products occurred through UPLC -ESI-MS (Acquity Ultra Performance LC) on reversed phase column (Waters UPLC BEH 300 C18, 1.7 μm, 50 x 2.1 mm), flow 0.25 ml/min, water +0.2% HCOOH (A) and acetonitrile +0.2% HCOOH (B) as eluents (gradient is reported in Appendix 8.5).

NM1. (1650.63) MS (ESI+) m/z: $[M+2H]^{2+}$ Calcd. 826.32 Found 825.92; $[M+3H]^{3+}$ Calcd. 551.21 Found 551.05; $[M+4H]^{4+}$ Calcd. 413.66 Found 413.68. ε = 52700. Yield = 12%.

NM2. (1990.93) MS (ESI+) m/z: $[M+3H]^{3+}$ Calcd. 664.64 Found 664.50; $[M+4H]^{4+}$ Calcd. 498.73 Found 498.70; $[M+5H]^{5+}$ Calcd. 399.19 Found 399.20. ε = 74600. Yield = 20%.

PNA1. (2000.92) MS (ESI+) m/z: $[M+2H]^{2+}$ Calcd. 994.46 Found 995.02; $[M+3H]^{3+}$ Calcd. 663.31 Found 663.69; $[M+4H]^{4+}$ Calcd. 497.73 Found 497.80. ε = 63774. Yield = 0.5%

PNA2. (1986.92) MS (ESI+) m/z: $[M+2H]^{2+}$ Calcd. 1001.46 Found 1001.42; $[M+3H]^{3+}$ Calcd. 667.97 Found 667.76; $[M+4H]^{4+}$ Calcd. 501.23 Found 501.16. ε = 63774. Yield = 1.1 %

PNA3. (2341.22) MS (ESI+) m/z: $[M+2H]^{2+}$ Calcd. 1171.61 Found 1172.18; $[M+3H]^{3+}$ Calcd. 781.41 Found 781.57; $[M+4H]^{4+}$ Calcd. 586.31 Found 586.57. ε = 85674. Yield = 0.2%

Melting experiments. Concentration of stock PNA solutions were determined by UV measurement at 260 nm using extinction coefficients reported above for each PNA. Sample were prepared in order to have a final PNA concentration of 5µM, in PBS buffer, for a total volume of 1 ml. Thermal denaturation profiles (Abs vs T) for single strand were recorded using UV/Vis Thermo Scientific Evolution 260 BIO spectrophotometer equipped with a Peltier temperature controlling system (Thermo Fisher Scientific SPE8W). Heating rate was set to 1°C/min. Melting curves were smoothed using Savitsky Golay algorithm. T_m values were obtained from minima of first derivative of curves.

5. Virtual screening of dimeric base monomers and building blocks synthesis

5.1. Introduction

As discussed in chapter 1, classic approach for finding new modified PNA structures with improved RNA binding properties is based on the information inferred from solid-state structures available in literature and then synthesis of new compounds in a trial-and-error strategy. Synthesis of new modified compounds is time and money consuming, generally requiring many reaction steps and conditions, and each of these must be optimized in order to obtain acceptable yields. However, the crystal structures, though very useful for getting insights into the molecular properties, represent a static model, and do not account for all possible conformations of the molecules involved; furthermore, the solid-state structure can be different from the preferred conformation in solution. Therefore, a computational approach able to predict effects of introduction of modifications on systems would be extremely useful in order to select more promising compounds and focus synthesis on these. In chapter 3 we have presented a robust methodology to study properties of PNA both as single strand and duplex. We wanted then to apply Molecular Dynamics for a virtual screening of modified PNA monomers, in order to take into account the possible evolution of the system upon thermal motion, and all possible conformations of modified monomers and their effect on duplex stability. Moreover such approach, which was made simpler by the availability of the modified NAB developed by us, allowed to directly study the effects on a specific sequence of applicative interest of the PNA:RNA duplex. In particular for this study we have chosen a PNA complementary to the seed region of miR 221 and 222 (H-ATG**T**AGC-Gly-NH_2, **T** is the modified dimeric base). Simulations described in this chapter were carried out in collaboration with the group of Prof. Michele Parrinello, at the Department of Chemistry and Applied Biosciences ETH Zurich, USI Campus, Lugano, using the computing facilities there available.

5.2. Molecular Dynamics for base-modified PNA Model 1 (Mod-1)

In Chapter 4 were presented some interesting properties concerning a modified PNA monomer containing two uracil moieties. We wanted to simulate a PNA containing this dimeric base (Figure

4.2), which will here be called Mod-1, in order to better understand experimental data. Using the modified NAB (section 2.7) and manual refinements, where needed, we constructed duplexes of the PNA with both DNA and RNA (Figure 5.1). These systems were energy-minimized and 200 ns long simulations were performed on both systems (DNA: cubic box 55 X 51 X 64 $Å^3$, 17700 atoms; RNA cubic box 55 X 51 X 62 $Å^3$, 17300), using the same conditions reported in section 2.5.

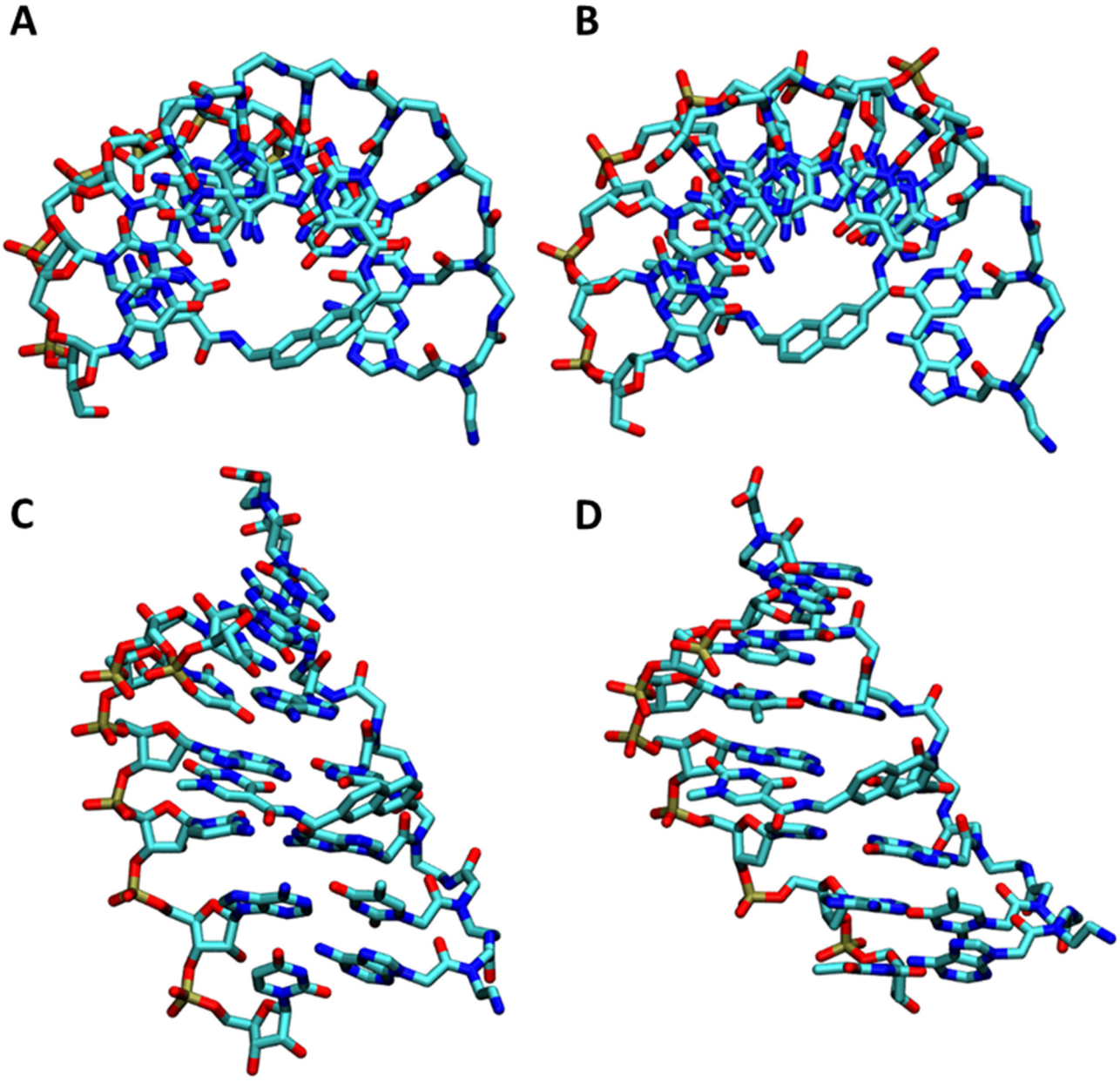


Figure 5.1: Structures of the systems containing Mod-1 monomer. **A**) top view and **C**) side view of PNA:RNA duplex. **B**) top view and **D**) side view of PNA:DNA duplex.

The first uracil (U1) was perfectly inserted in duplex and second uracil (U2) was able to interact with target adenine. However, U2 was not perfectly paired via Hoogsteen hydrogen bonds, due to sterically tensioned conformation, thus inducing only a limited interaction. It had also the possibility

to pair with another adenine present in the sequence and it was actually less tensioned in this conformation. U2 was also free to move in solvent and then to interact with other bases through stacking interactions. These two limit conformations are shown in Figure 5.2 in the case of PNA:RNA duplex.

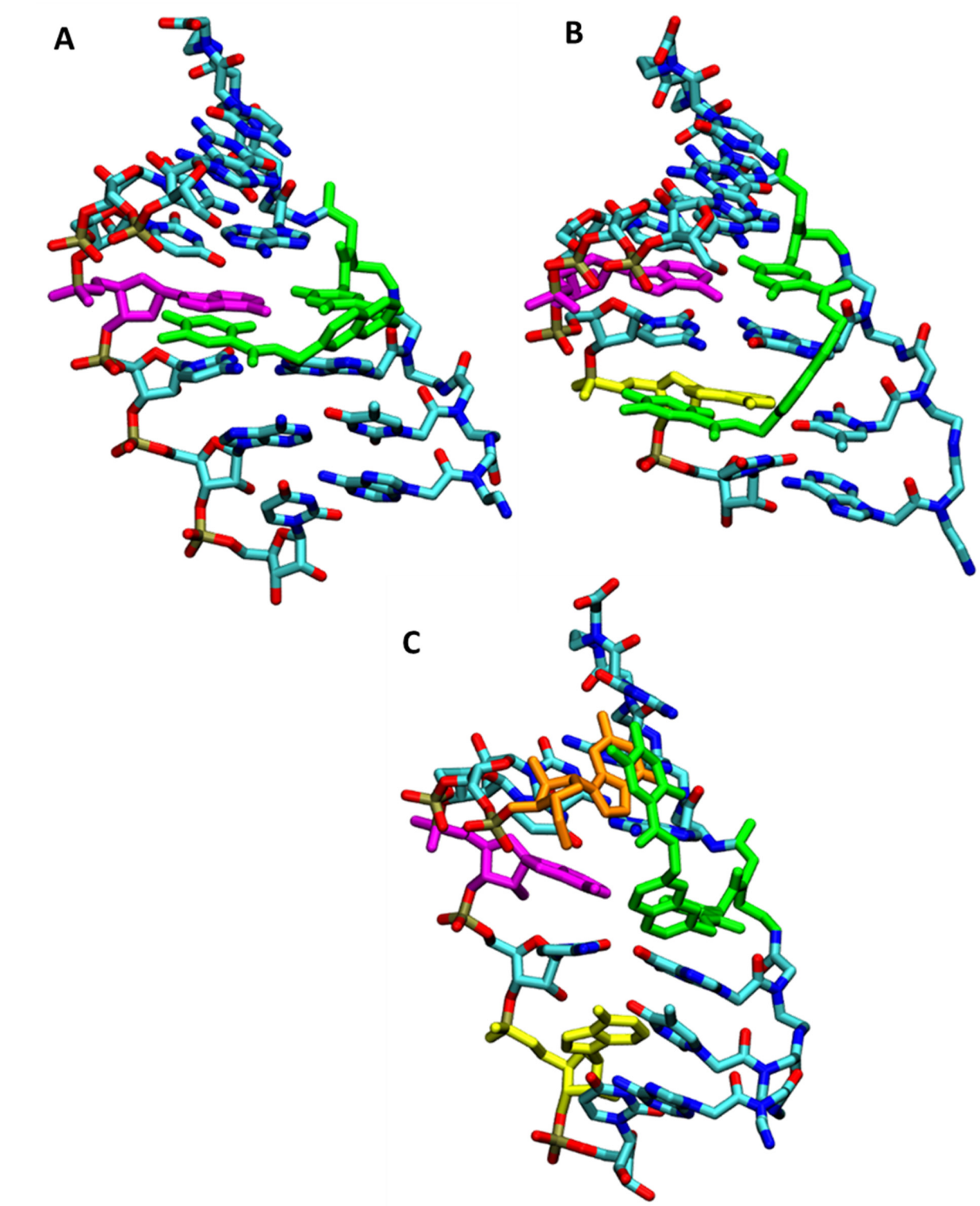


Figure 5.2: Snapshots of the MD simulation for Mod-1 pairing with RNA: **A**) interaction of modified base (green) with target adenine (magenta); **B**) interaction with another adenine of the sequence (yellow); **C**) stacking interaction with other base (orange).

These results are in accordance with experimental data. As seen in previous chapter, study on PNA:DNA heteroduplex revealed that insertion of modified base only slightly increased stability of the duplex and previous study on **PR212**[270] pointed out more interesting properties of selectivity. When target adenine is present in the sequence, partial interaction of second uracil provides some stabilizing effect, while when a mismatch occurs, pedant uracil is completely free to move in solvent, interfering with other bases pairing and thus inducing destabilization. However, for adenine, though some attractive interactions are observed, the effect of the second uracil is less than that expected, because of the relative mobility of this moiety. Furthermore, if the PNA:DNA and PNA:RNA duplex stability is not strongly enhanced, competition of PNA:PNA self-association can be relevant. In fact, experimental melting of PNA:DNA duplexes suggested also a preferred self-aggregation of the system instead of duplex formation in presence of TA or AT doublet in the PNA sequence (section 4.4). Though we have not considered the stability of PNA:PNA duplex at the computational level, the self-pairing of the AT pair would be favored by the flexibility of the two PNA involved, which allows the nucleobases to maximize their mutual interactions.

5.3. Molecular Dynamics for base-modified PNA Model 2 (Mod-2)

Since model described above presented some geometrical constraints which hampered the pairing of the second uracil moiety with the target adenine, we focused our attention on the presence of the rigid amide bonds. We then substituted the amide group of second uracil with an amine, which should presents a greater mobility and consequently adaptability (Figure 5.3).

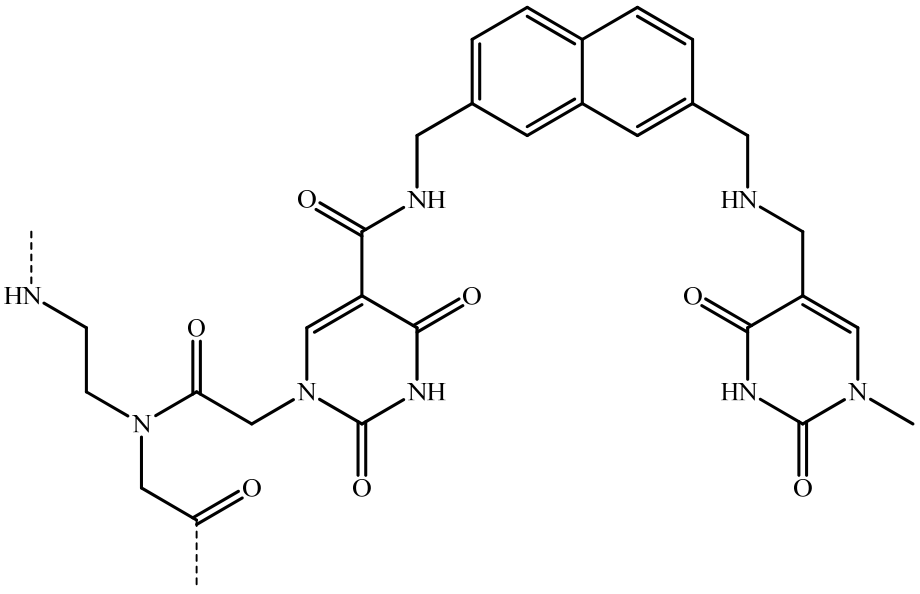


Figure 5.3: Structure of Mod-2 modified base.

Since simulations were time demanding and considering that desired final targets were micro RNAs, here after we performed simulations only on PNA:RNA duplexes. The starting duplex structure was generated using modified NAB, and the new dimeric base was introduced manually by modifying a thymine base (Figure 5.4). We performed 200 ns of standard MD (cubic box of 60 X 55 X 67 $Å^3$, 17300 atoms) in the conditions reported in section 2.5.

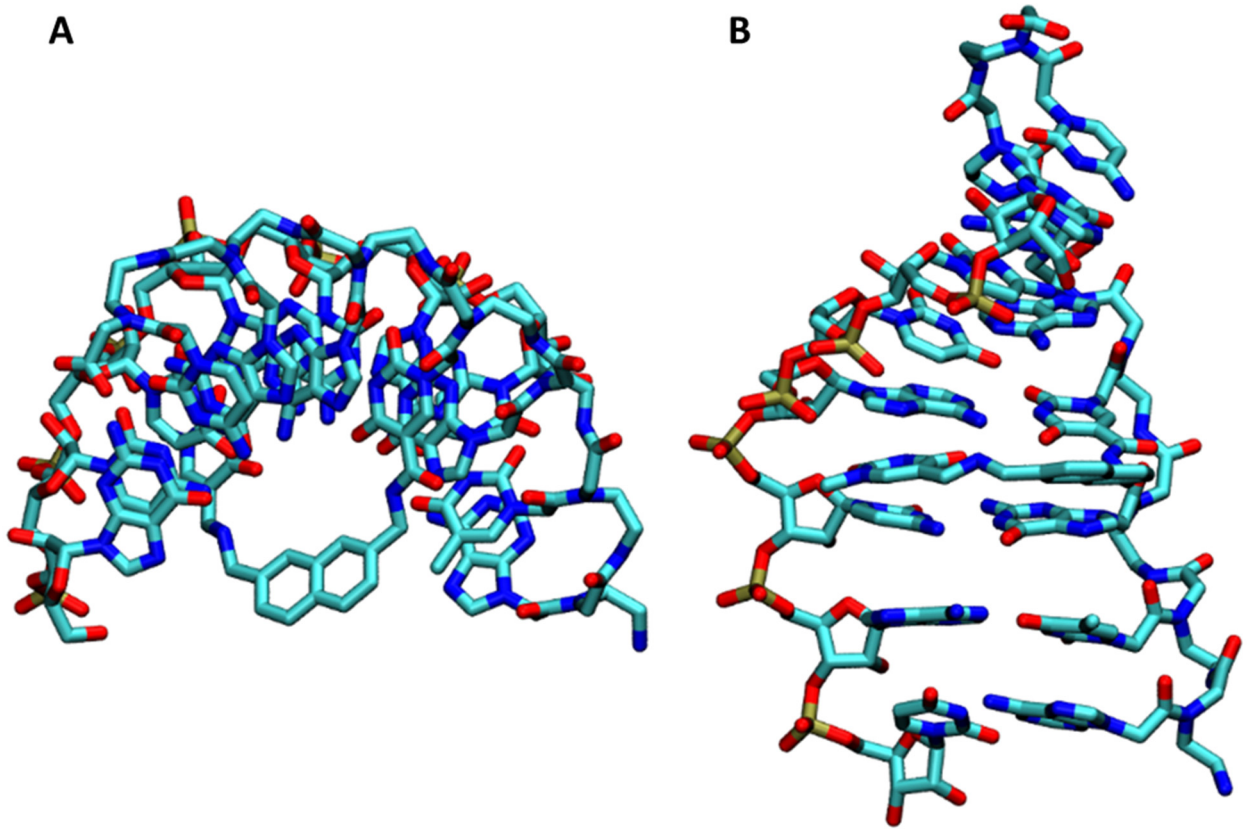


Figure 5.4: Starting structure, containing Mod-2 monomer, used for the MD simulation. **A**) top view **B**) side view.

The simulation showed that the presence of the amino group did not allow the second uracil to reach a stable interaction with the target adenine; thus also this base is not correct for a cooperative interaction with target. Moreover, the increased flexibility of the bridge linking the second uracil unit, determines a favored collocation of U2 away from the target, stabilized by interactions with solvent (Figure 5.5).

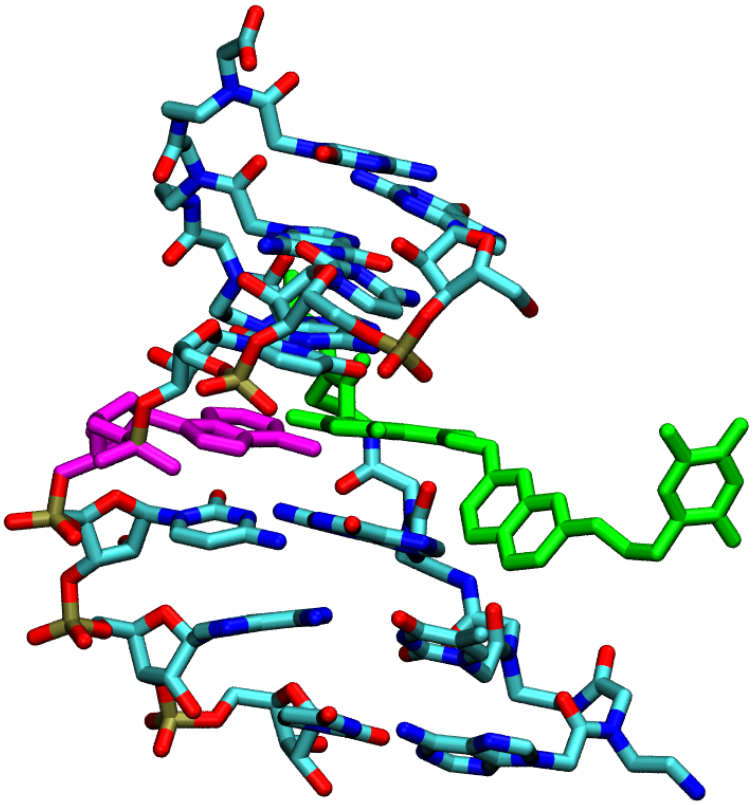


Figure 5.5: Favored disposition of Mod-2 (green). Pedant uracil freely move in the solvent while the first uracil is involved in standard Watson and Crick bonds with target adenine (magenta).

5.4. Molecular Dynamics for base-modified PNA Model 3 (Mod-3)

Opposite to the previous model, we tested if a rigid linker could allow cooperative interactions with the target adenine still avoiding the collapse of the two uracils. We thus kept the naphthalene linker, and looked for synthetically accessible structures. Therefore, we proposed Model 3 as possible new candidate for binding adenine (Figure 5.6).

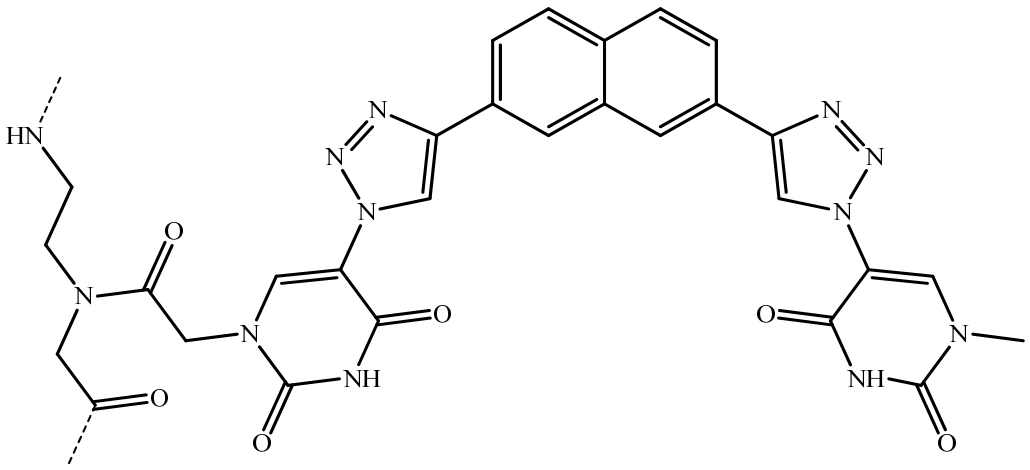


Figure 5.6: Structure of Mod-3 central fragment.

The triazole ring in this monomer can be obtained by a copper (I) catalyzed click reaction between 2,7-diethynylnaphthalene and a 5-azido uracil derivative. This compound presents a major distance between two uracils compared to Mod-1 previously described (10 Å compared to 9.5 Å) and then should reduce steric constrains. The starting PNA:RNA duplex structure was generated using same procedure used for the previous model. We performed a 40 ns simulation in a cubic box filled with TIP3P water molecules (58 X 54 X 64 $Å^3$, 17750 atoms), under the same conditions used before (Figure 5.7).

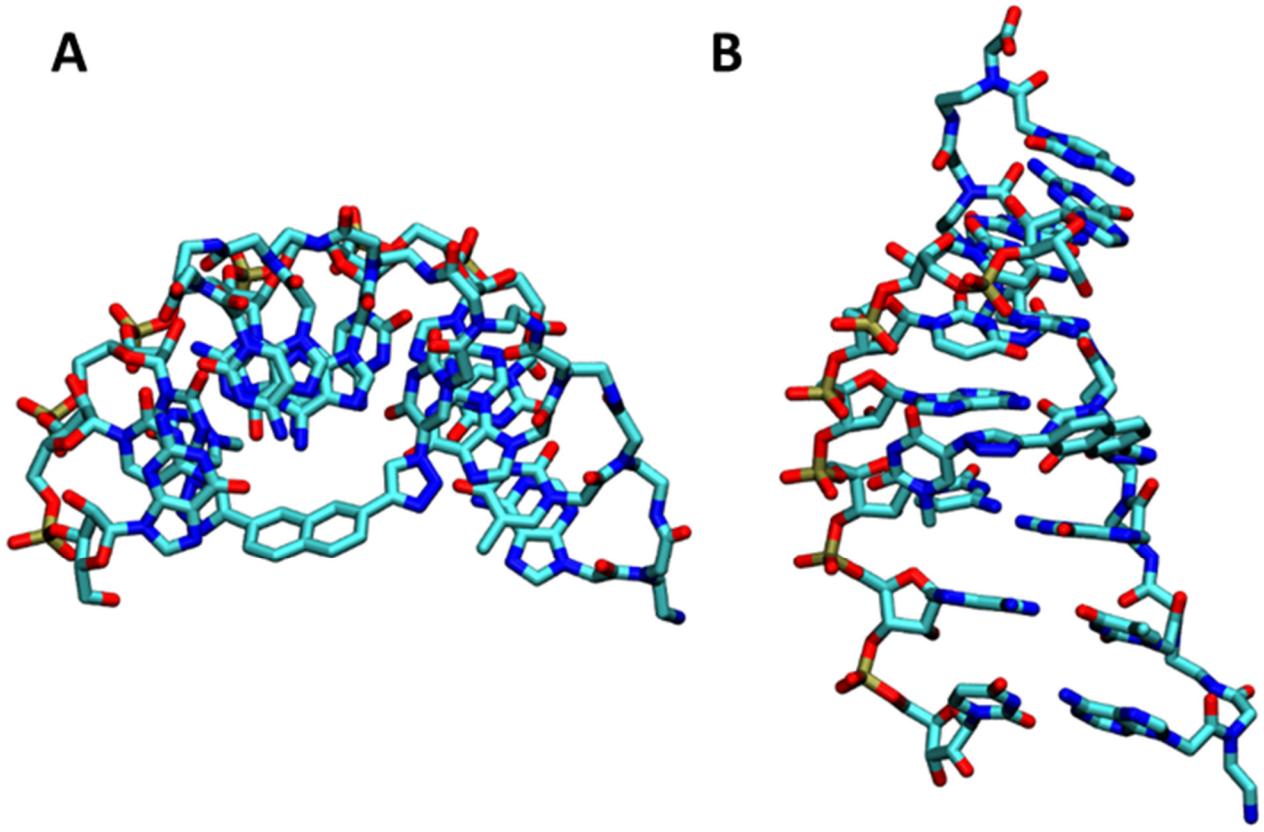


Figure 5.7: Starting structure, containing Mod-3 monomer, used for the MD simulation. **A**) Top view of the system; **B**) Side view of the system.

The simulation showed that also for Mod-3 the second uracil didn't present the correct geometry to interact with adenine. However, in this case, unlike with Mod-2, the second uracil, being part of a rigid moiety, gave rise to destabilizing steric interactions which interfered also with base pairing of the first (Watson-Crick) uracil, causing weakening in the central part of the duplex and consequent duplex disruption. Indeed a little breach was generated during MD simulation, and this allowed water entering destabilizing the systems until complete opening (Figure 5.8).

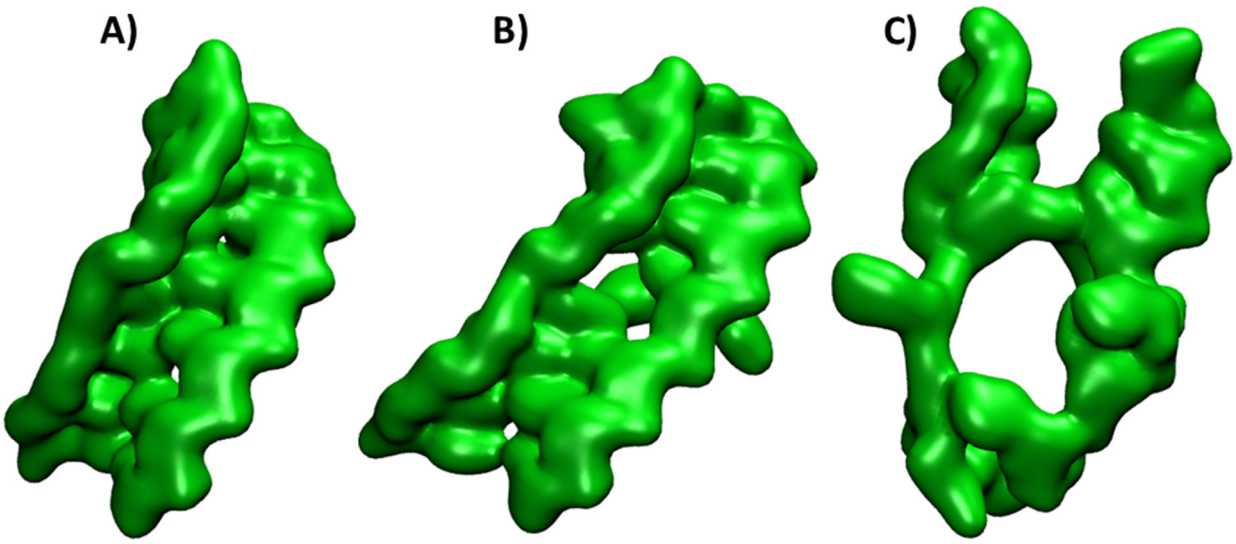


Figure 5.8: Snapshots of MD simulation for Mod-3 PNA:RNA. Duplex surface representation. **A**) view of duplex minor groove at the beginning of the simulation; **B**) in the central part of the duplex was possible to recognize a breach that allowed water entering destabilizing the system; **C**) duplex was completely opened in the central part.

Behavior of Mod-3 is a good example to show difficulty in finding new effective modified systems. Compounds with similar geometries exhibited completely different properties and what was expected to be an advantage, greater uracil distance, was indeed found to be a drawback.

5.5. *Ab initio* calculations and Molecular Dynamics for base-modified PNA Model 4 (Mod-4)

Finally, we studied another modified base, analogue to those shown before (Figure 5.9), but with removal of any sp^3 atoms from the bridge, thus allowing a completely flat geometry to be obtained. Contrary to Mod-3, the additional sp^2 atoms were not forced to be part of a cycle, thus allowing a different geometry.

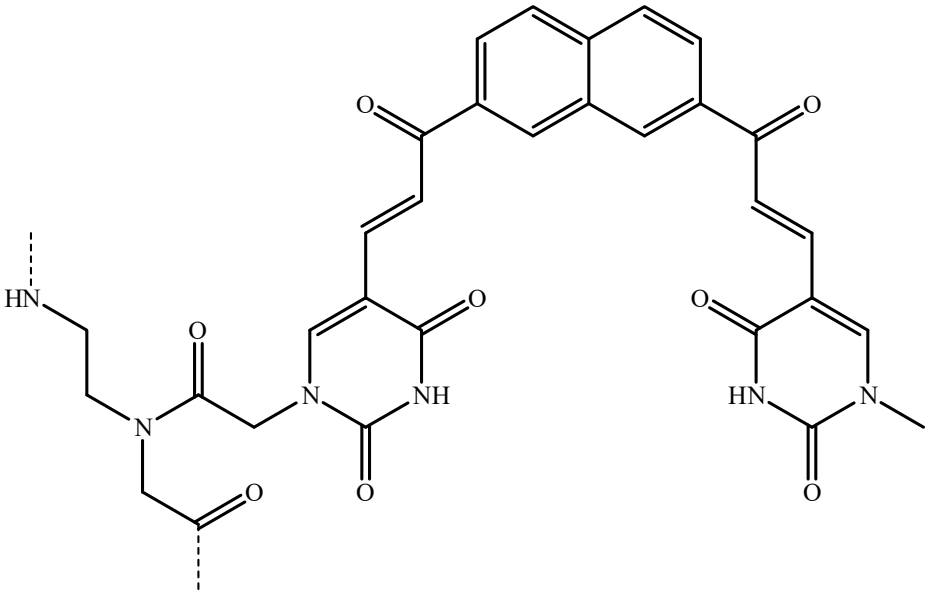


Figure 5.9: Structure of Mod-4 central fragment.

This compound presented an α,β-unsaturated ketone as linker between naphthalene and uracil moieties. This modification was designed taking account of the excessive mobility of pedant uracil of previous models. Conjugation between bonds should favor rigidity and planarity of the molecule, maintaining the Hoogsteen interactions once they are formed. The synthesis of this monomers is more challenging compared to those reported before. Double bond requires *E* geometry and also carbonyl should have the correct orientation in respect to the naphthalene. We then evaluated which could be the preferred orientation by calculating the difference in energy between model compounds reported in Figure 5.10 by *ab initio* energy calculation.

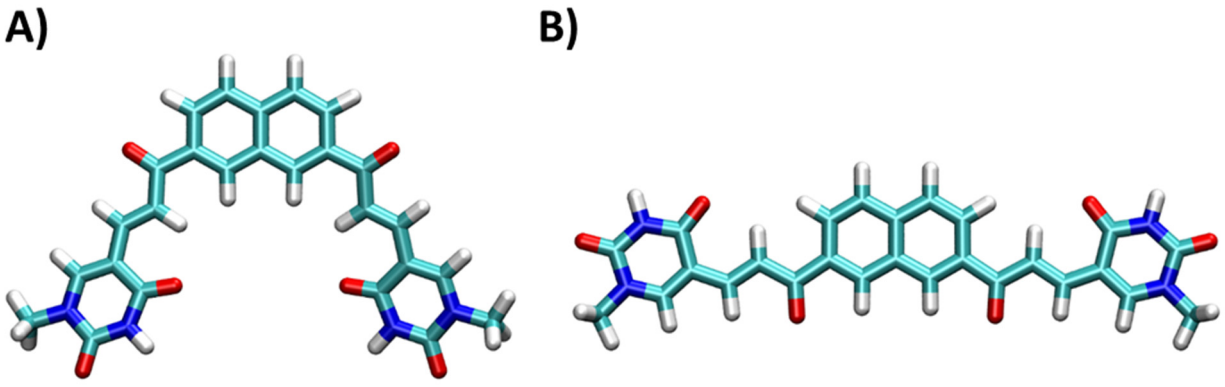


Figure 5.10: Optimized geometry for: **A**) monomer with convergent disposition of uracils which allows adenine clamping; **B**) monomer with conformation with rotated uracils (180° flip of the naphthalene-C=O bonds).

Since this molecule has at least two flat conformations, we performed an optimization of geometry of the cleft-like (Figure 5.10 A) and extended (Figure 5.10 B) structures with Gaussian 09 D1[246] using B3LYP[281] DFT theory and 6-31G* as basis set and then we calculated frequencies in the same way. The most stable conformation was the cleft-like one (Figure 5.10 A), but energy differences of these two compounds in vacuum were small, about 0.5 kcal/mol, and so both are expected to be populated at r.t.. However, once the duplex is formed, it is reasonable to expect that equilibrium can be shifted toward the cleft-like conformation by favorable Hoogsteen interactions. Moreover, the interconversion between these two structures is possible only if the rotational energy barrier is sufficiently low. We then scanned the energy variation as function of the dihedral angle (Figure 5.11) and we found an energy barrier of 6.35 kcal/mol, which should allow the interconversion between these two forms to occur. Optimized structures and PES scan are reported in Appendix 8.7.

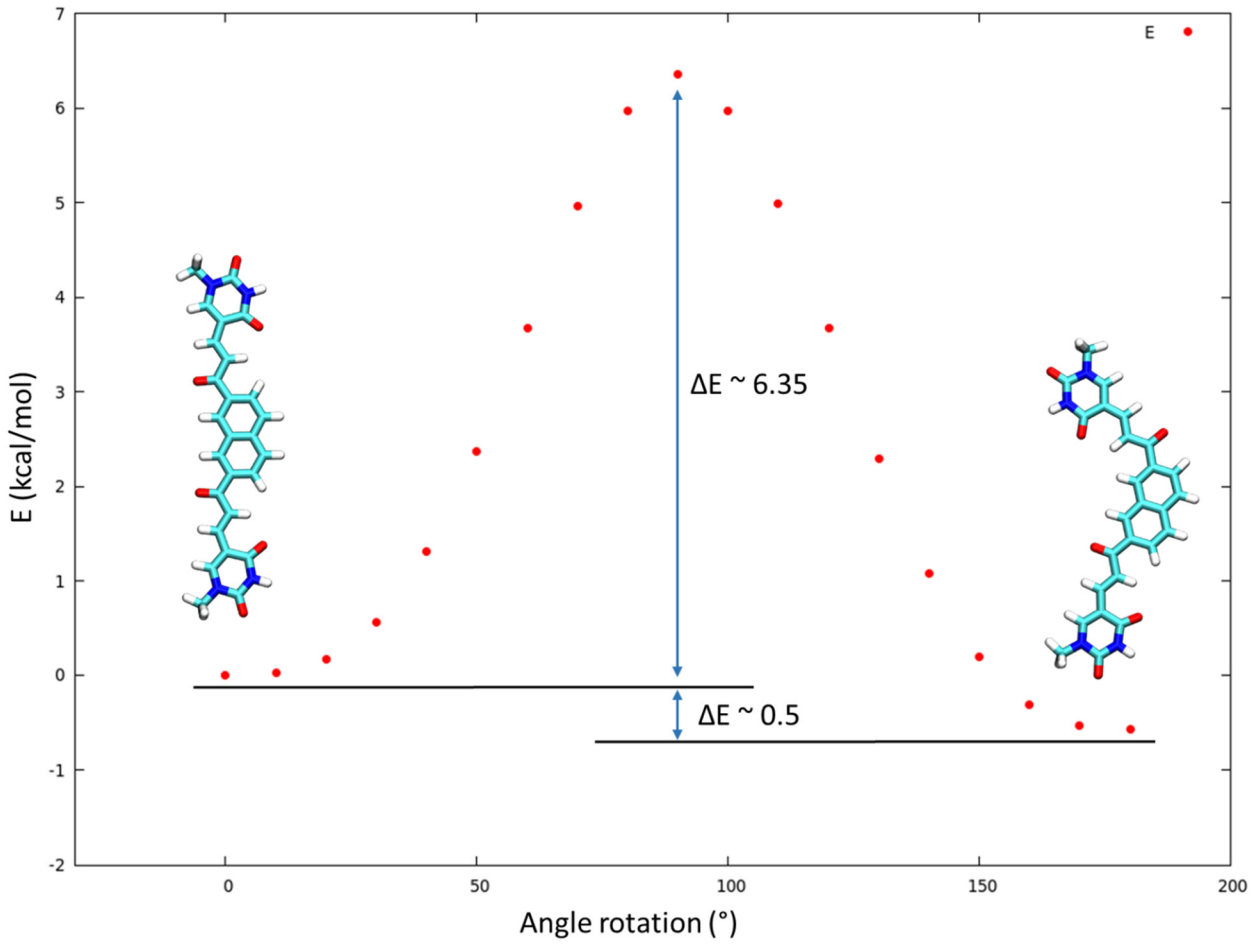


Figure 5.11: Energy plot as function of rotation on naphthalene – CO dihedral angle.

Using this nucleobase, in its cleft-like conformation, we built a PNA:RNA duplex (Figure 5.12) and then performed 200 ns Molecular Dynamics simulation in the condition reported in section 2.5 in a cubic box of 56 X 51 X 62 $Å^3$ filled with TIP3P water molecules for a total of 17350 atoms.

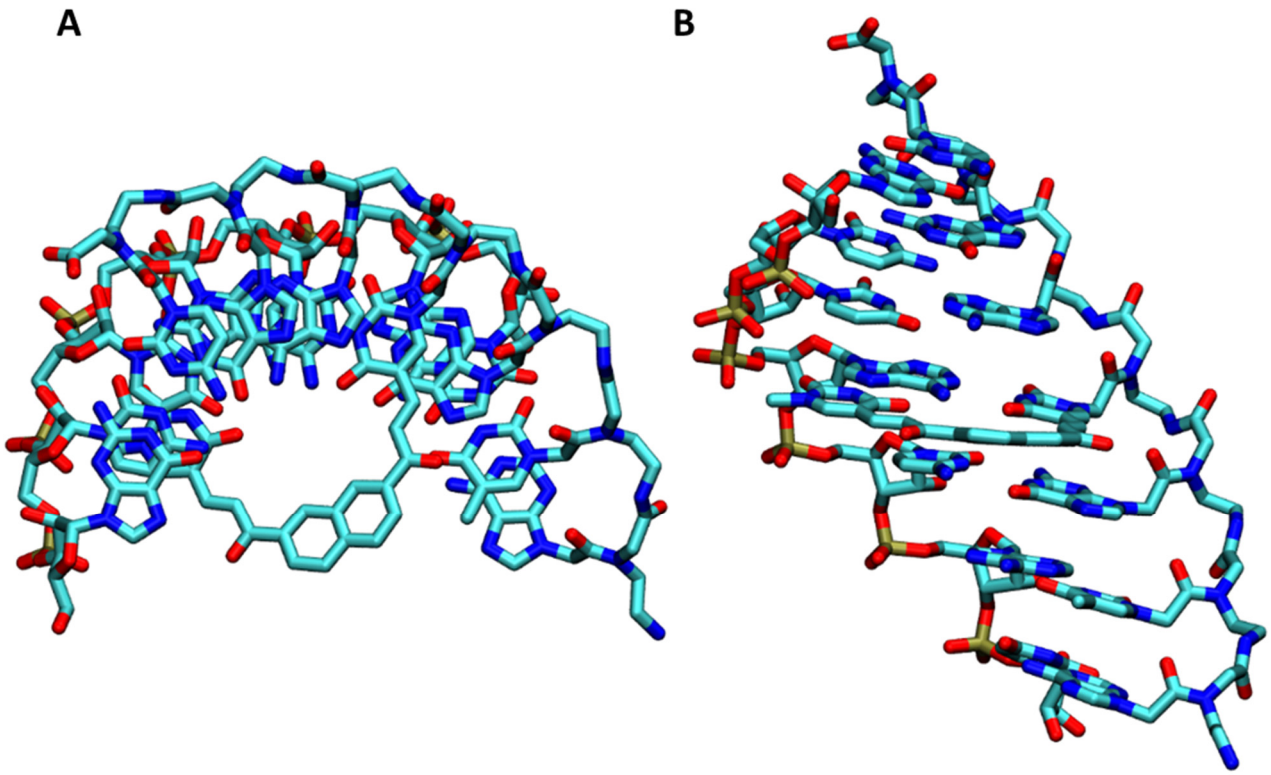


Figure 5.12: PNA containing Mod-4 modified base. **A**) Top view of the system; **B**) Side view of the system.

In this case the duplex remained intact, and the second uracil unit persistently interacted with the target adenine. Checking the average distance between atoms doing H-Bonding of second uracil and target adenine was found that interactions were more consistent that in the previous cases. Base still preserved a certain mobility, as proven by region between 20-50 ns and 60-90 ns, but it retained pairing for about 70% of the simulation (Figure 5.13). Therefore, this compound seemed a more promising candidate compared to other models and we decided to proceed with its synthesis.

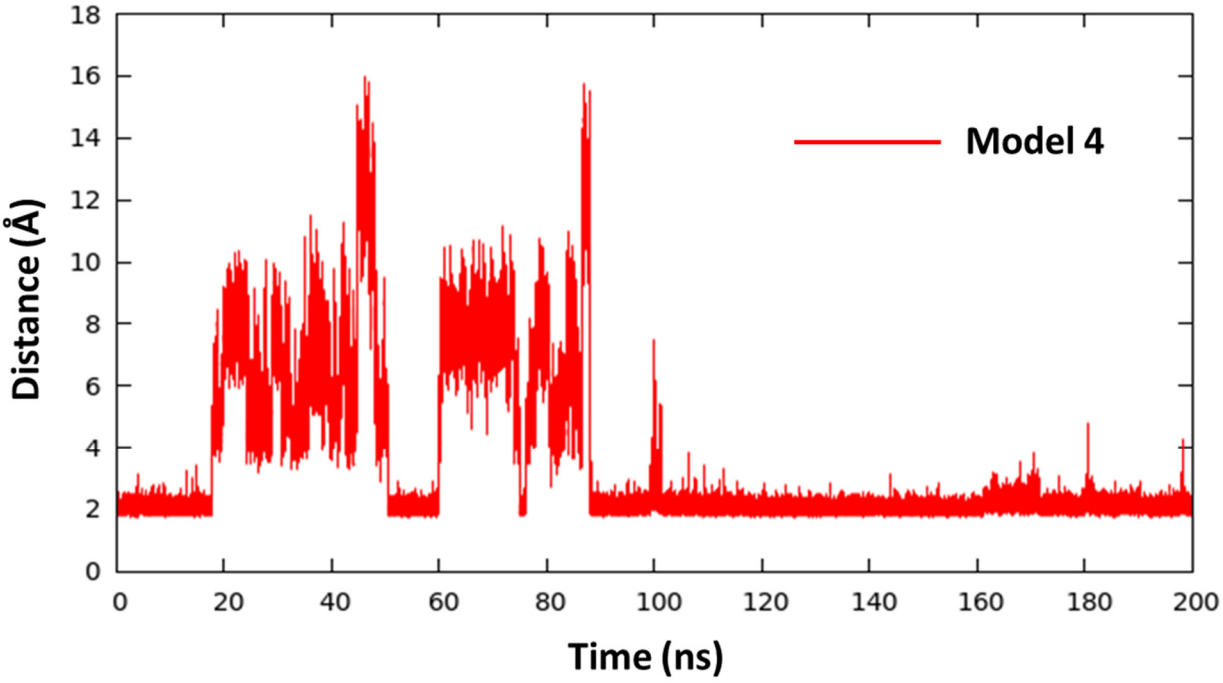


Figure 5.13: Average distance between atoms involved in Hoogsteen bonds (H3 and O4 of modified base and respectively H6 and N7 of target adenine) as function of simulation time.

5.6. Building blocks synthesis

5.6.1. Retrosynthesis

The different synthons composing the Mod-4 nucleobase are depicted in Figure 5.14 with different colors.

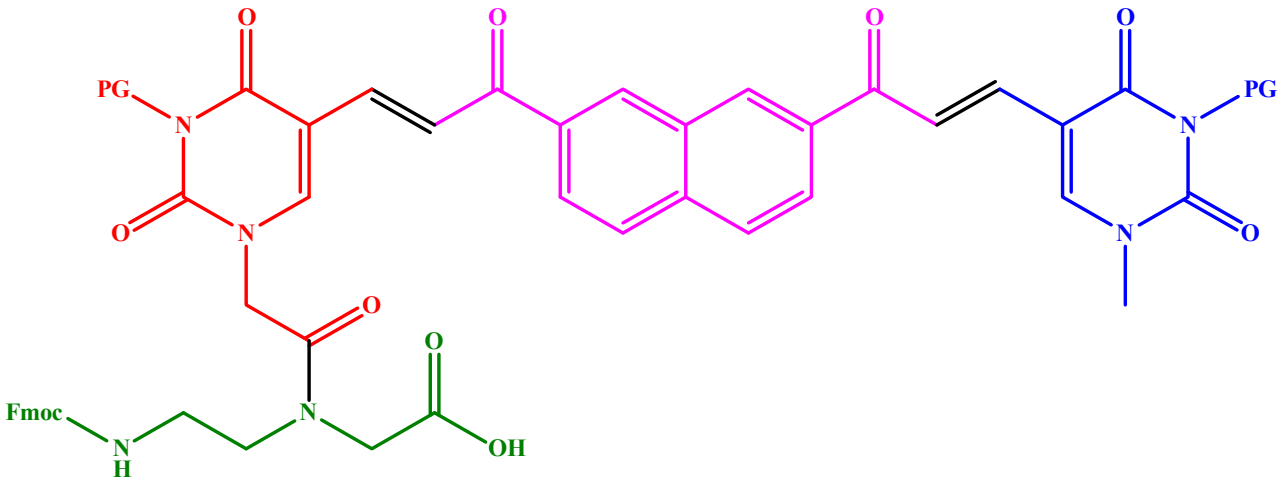
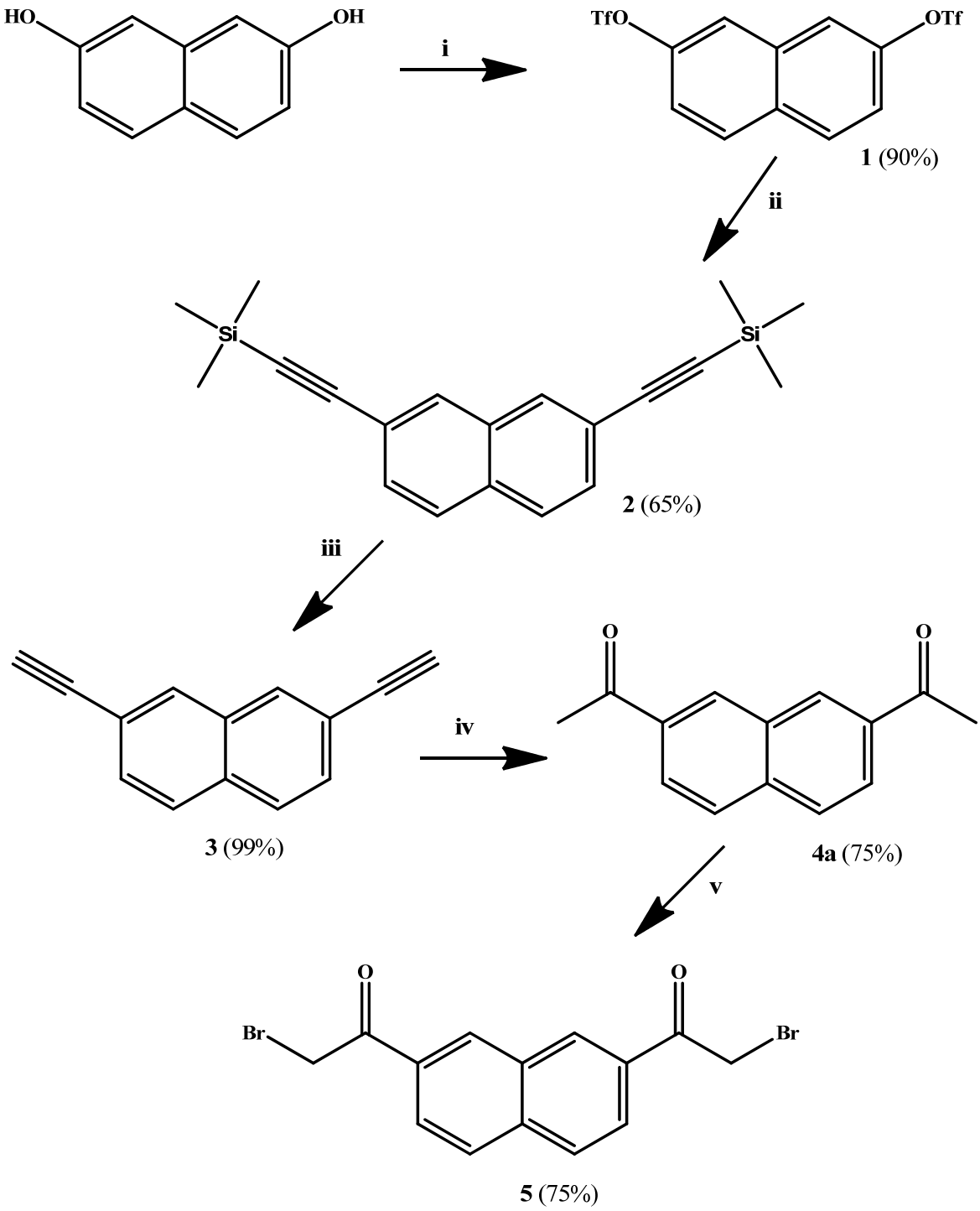


Figure 5.14: Structure of the final monomer derived from Mod-4. Constituent units are marked with different colors. PG is a protective group.

The PNA backbone (green) is the same of the previous monomer and can be synthetized as reported for compound **3** in Chapter 4. The α,β-unsaturated carbonyl unit can be obtained using different strategies such as aldol condensation, carbonylation of the alkene or Wittig reaction.

5.6.1. Synthesis of central linker precursors

The central linker was synthetized starting from 2,7-dihydroxynaphthalene as cheap starting material (Scheme 5.1); this was triflated using $(CF_3SO_2)_2NPh$. Product **2** was obtained from **1** and ethynyltrimethylsilane using a Sonogashira coupling in presence of $Pd(PPh_3)_2Cl_2$ and CuI. Silyl groups were then removed under basic conditions to yield **3**. Hydration of alkyne was achieved in glacial acetic acid using iron sulfate as catalyst.[282] By controlling the reaction time was possible to obtain the mono-hydrated product **4b** (Figure 5.15) or force the reaction toward compound **4a**. After 2 h the main product was **4b**, but still with a large amount of unreacted **3**. After 5 h, reagent conversion was almost complete and mono (**4b**) and di-hydrated (**4a**) products were in about 1:1 ratio. Leaving the reaction overnight, was possible to obtain an almost complete conversion to product **4a**. We tried both base- and acid-catalyzed aldol condensation with 5-formyluracil derivative **7** (Scheme 5.2) and the diketone **4a**, but the condensation reaction didn't occur under the conditions tested (basic: LDA -78 °C; acid: H_2SO_4 40 °C). This was probably due to a reduced reactivity of 5-formyluracil derivative compared to other aldehydes. Thus, a Wittig-type reaction seems to be more suitable for obtaining the final product (Scheme 5.5). In order to obtain the synthon for this reaction, the diketone was then selectively halogenated in alpha position with NBS to obtain product **5**.
Since compound **5** is symmetric, introduction of two different uracil moieties could results difficult. One possible pathway is to form ylide on both side and then slowly add strictly one equivalent of aldehyde followed by one equivalent of the second aldehyde. Another possibility is to convert one bromine atom into iodine using Finkelstein reaction[283]. Iodine should form more readily the ylide allowing to perform sequential Wittig reaction. Another possibility is to use compound **4b** allowing introduction of a first uracil moiety and then of a second unit after hydration of alkyne. All this strategies will be tested in future in order to find best conditions for synthesis.



Scheme 5.1: Synthesis of 2,7-bis(bromoaceto)naphthalene. (i) [$(CF_3SO_2)_2NPh$, NEt_3, DCM, 3h, 0°C to r.t.]; (ii) [Me_3SiCCH, $Pd(PPh3)_2Cl_2$, CuI, NEt_3, dry DMF, N_2, o.n., 40 °C]; (iii) [1M NaOH/THF 1:1, 2 h, r.t.]; (iv) [$Fe_2(SO_4)_3$, AcOH, o.n., 95 °C]; (v) [NBS, DCM, 2 h, reflux].

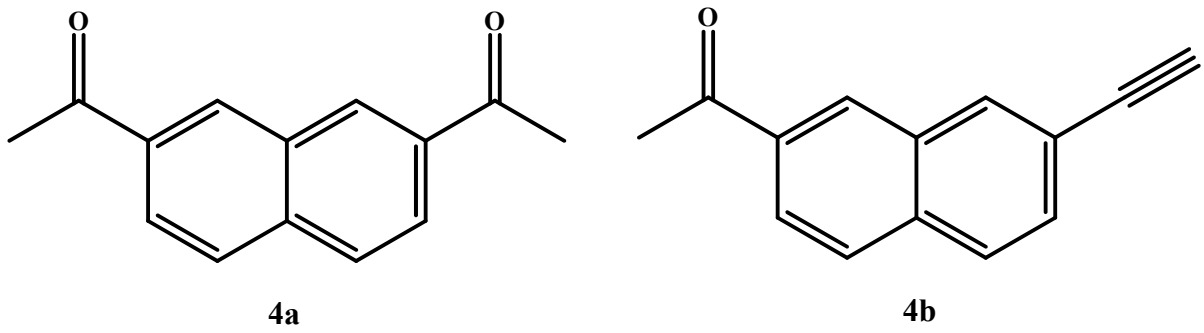
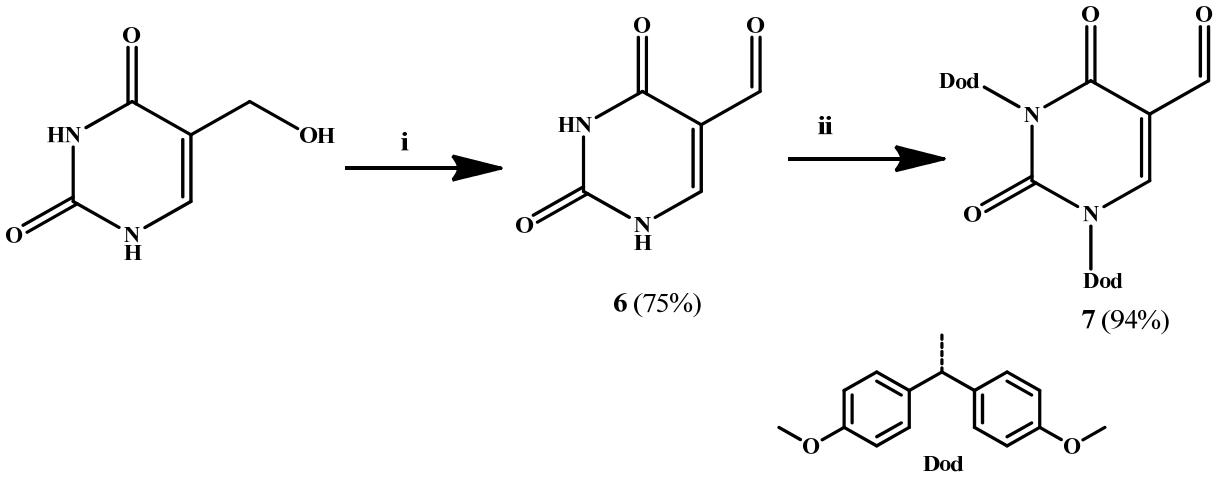


Figure 5.15: Products obtained from hydration of compound **3**.

5.6.2. Uracil derivatives

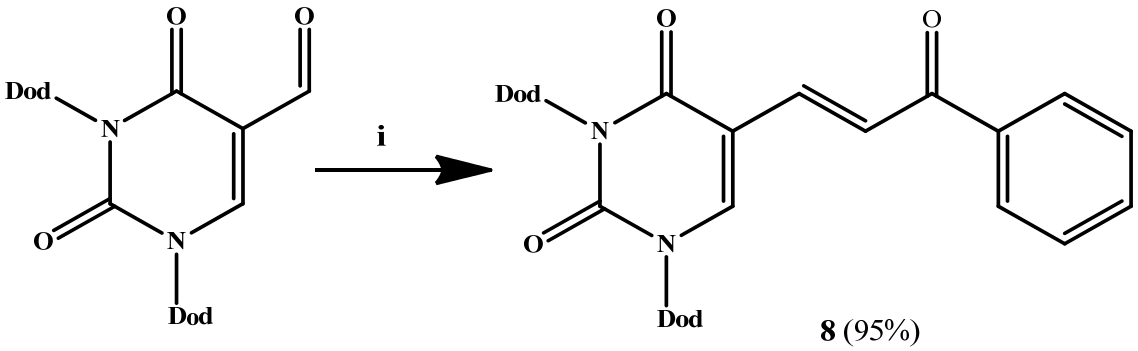
The Wittig reaction involves the use of two differently substituted uracils. In particular we designed a single precursor for both. In order to test the Wittig reaction conditions, we initially synthetized a protected derivative of the aldehyde (Scheme 5.2).



Scheme 5.2: Schematic synthesis of 1,3- di Dod- 5- formyl uracil. (i) [CAN, H_2O, 1h, 90 °C]; (ii) [Dod – Cl, K_2CO_3, dry DMF, 4 h, r.t.].

5 – hydroxymethyl uracil was synthetized according procedure reported in literature.[284] Oxidation to the aldehyde **6** occurred in presence of mild oxidizing condition using cerium ammonium nitrate (CAN) $[(NH_4)_2Ce(NO_3)_6]$.[285] Aldehyde was then protected with Dod chloride in both N(1) and N(3) position.

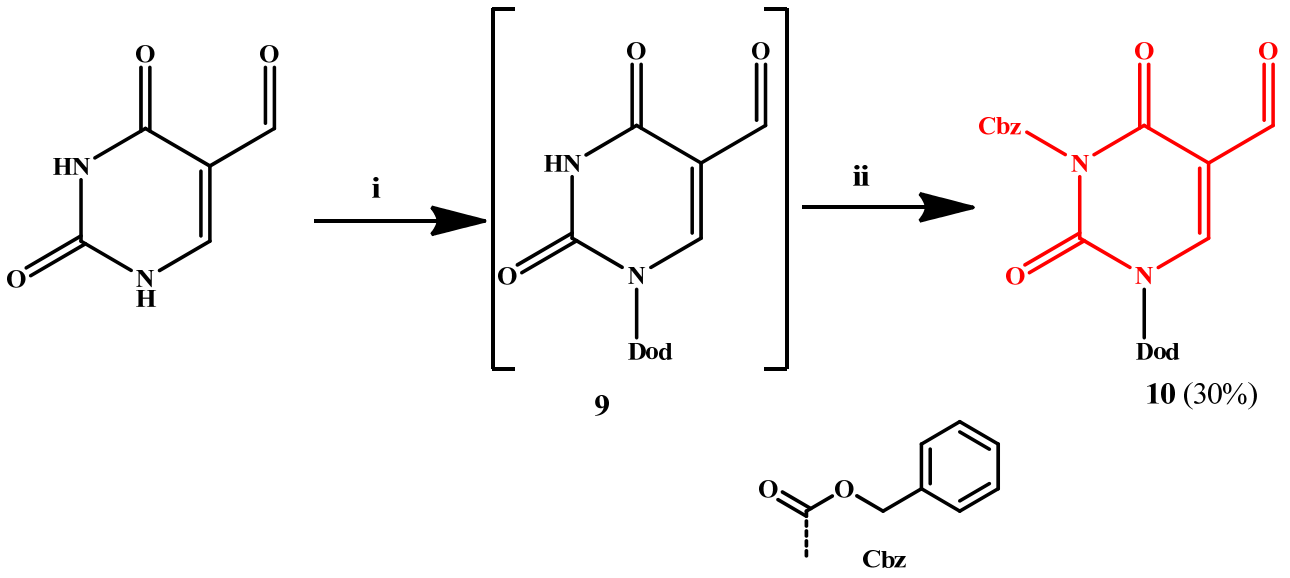
Compound **7** was then used in a model reaction with bromoacetophenone. In particular, since required geometry of the carbon-carbon double bond is *E*, we worked under Schlosser conditions, in which *trans* isomer is favored by heating and by an excess of base (Scheme 5.3).[286]



Scheme 5.3: Synthesis of (E)-1,3-bis(bis(4-methoxyphenyl)methyl)-5-(3-oxo-3-phenylprop-1-en-1-yl)pyrimidine-2,4(1H,3H)-dione. (i) [$BrCH_2COPh$, PPh_3, Na_2CO_3, o.n., 100 °C].

J coupling between alkene hydrogens in the product **8** is 15.4 Hz, that is in the range 12 – 18 Hz that is typical for *E* geometry.[287]

The two different types of uracil to be coupled to the central linker through Wittig reaction were then obtained from a doubly protected 5-formyluracil derivative, using Dod-protection for N(1) and Cbz-protection for N(3); this can be used as a precursor for both units.



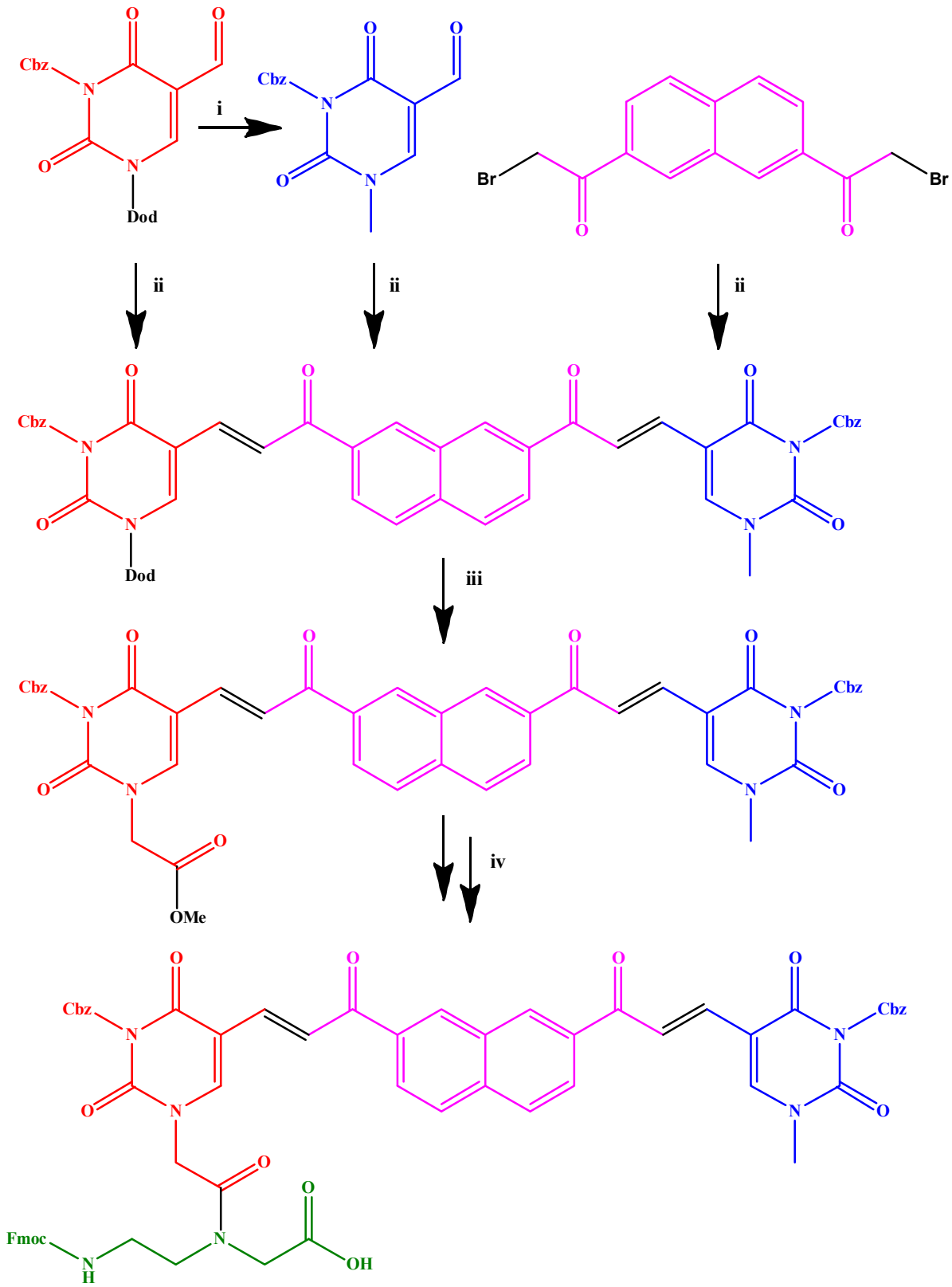
Scheme 5.4: Synthesis of 1-Dod-3-Cbz-5-formyl uracil. (i) [Dod-Cl, K_2CO_3, dry DMF, 1 h, r.t.]; (ii) [Cbz-Cl, Dod-Cl, K_2CO_3, dry DMF].

Compound **6** was protected with Dod chloride to obtain product **9**. Attack in N(1) position was favored since Dod is a bulky group. Using NMR it was possible to confirm the presence of desired product thanks to disappearance of the coupling between hydrogen in N(1) position and that one in C(6). The crude product was then protected with Cbz chloride to obtain final product **10**. Yield was particularly low probably due to oxidation of the aldehyde during purification by flash chromatography. Cbz and Dod are both cleaved in strong acidic media, but Cbz is removed under harsh conditions. Therefore is possible to deprotect N(1) position using a solution of TFA and DCM without cleaving Cbz group and thus introducing alkyl group of interest. In this way it will be possible to initially obtain uracil with methylated N1 position and then, after Wittig reactions, to insert methylencarboxylic linker for coupling between base and backbone (Scheme 5.5).

5.7. Conclusions and future work

We have applied Molecular Dynamics simulations for a virtual screening of modified dimeric bases analogues. This is a novel approach in which rational design is used together a dynamic testing of system from a molecular point of view. Examples explored showed difficulty in predicting modification effects, since even small modifications led to completely different results. With this novel strategy was possible to find a promising candidate for adenine binding, saving time and money in experimental trials. Certainly also simulations have an intrinsic cost,[288] but the price has rapidly decreased during the years due to progressive increased computational power. We believe that this may lead, in future, to a more routinely use of simulations as support to organic synthesis design and understanding.

We have also synthetized building blocks to be used for obtaining final monomer. Synthetic pathway to be carried starting from building blocks produced in the present thesis is reported in the scheme 1.1. Completion of synthesis will be carried in the host laboratory, and represent one of the major goals of European Project ULTRAPLACAD.[289]



Scheme 5.5: Proposed synthetic pathway to obtain Mod-4 monomer. (i) Deprotection of Dod moiety, with 40% TFA solution in DCM, followed by N(1) alkylation with iodo methane to obtain second uracil; (ii) Wittig reaction between central linker **5** and two different uracil units; (iii) Dod removal in the same conditions of i and alkylation with bromo acetate; (iv) hydrolysis of the linker ester with 1M NaOH, coupling with the backbone (compound **3** of chapter 4) and hydrolysis of the final monomer with $Ba(OH)_2$.

5.8. Experimental Section

General. Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, LINK and were used without further purifications. TLC were run on Supelco 56524-25EA silica gel on aluminium foils with fluorescence indicator (254 nm). Column chromatography was performed as flash chromatography on Merck silica 60 (0.040 – 0.063) under 1.1-1.2 bar of air pressure. NMR spectra were obtained on Bruker Avance 300 MHz or 400 MHz spectrometers. Chemical shifts are presented as δ values and expressed in ppm units. FT-IR spectra were recorder on Thermo Nicolet 5700. ESI-MS spectra were recorded on Acquity Ultra Performance LC. GC-MS spectra were recorded on HP instrument (oven 6890N, mass detector 5973 and column SE-52). Meting points were measured with GALLENKAMP melting point apparatus.

Ab initio and Molecular Dynamics simulations. Calculations were performed on Brutus cluster, the central high-performance cluster of ETH Zurich,[290] or on CSCS servers (Dora[291] and Mönch)[292]. Ab initio calculation were performed by means of Gaussian G09 D1 using 8 cores. Molecular Dynamics simulation were performed using 32 – 64 cores and were conducted in the conditions reported in section 2.5.

Naphthalene-2,7-diyl bis(trifluoromethanesulfonate) (1). 2,7-dihydroxy naphthalene (500.5 mg, 3.13 mmol) and $(CF_3SO_2)_2NPh$ (2.344 g, 6.56 mmol) were dispersed in DCM (5 ml) and then NEt_3 (913 µl, 6.56 mmol) was slowly added at 0°C. Temperature was raised to room temperature and mixture stirred for 3 h. Reaction mixture was diluted with DCM (35 ml) and then exctracted with $KHSO_4$. The water phase was extracted with DCM (3 x 40 ml). The organic layers were combined, dried over sodium sulfate and evaporated under vacuum to afford pure product (90 %, 1.19 g).

(**1**) ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 8.03 (d, 2H, J = 9.2 Hz, CH naphth.), 7.84 (d, 2H, J = 2.4 Hz, CH naphth.), 7.50 (dd, 2H, J_1 = 9.1 Hz, J_2 = 2.4 Hz, CH naphth.). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ (ppm): 148.3, 132.1, 131.0, 130.8, 130.0, 121.2, 119.5. FT-IR ν (cm^{-1}) 1413, 1211, 1187, 1135, 1111. Full characterization is reported in literature.[293]

2,7-bis((trimethylsilyl)ethynyl)naphthalene (2). Vacuum-nitrogen cycles were performed on Schlenk tube containing compound **1** (814.6 mg, 1.91 mmol) followed by addition of dry DMF and

solution was stirred at room temperature. Then, in the following order, trimethylsilylacetylene (641.1 µl, 4.78 mmol), $Pd(PPh_3)_2Cl_2$ (63.7 mg, 0.10 mmol), CuI (17.3 mg, 0.10 mmol), NEt_3 (1.14 ml, 8.60 mmol) were added to the solution. When palladium was added to solution it became yellow, then with copper it turned orange and finally with triethylamine it went back to the yellow color. Temperature was raised to 40°C and solution stirred overnight. At the end of the reaction, mixture was diluted with AcOEt (200 ml) and washed with saturated $KHSO_4$ (2 x 200 ml), $NaHCO_3$ (2 x 200 ml), and brine (200 ml). The organic layer was dried over sodium sulfate and evaporated under vacuum. The crude was purified by flash column chromatography on silica gel with hexane as eluent (R_f = 0.1) to afford pure product (65%, 115 mg).

(2) ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 7.93 (s, 2H, CH naphth.), 7.74 (d, 2H, J = 8.4 Hz, CH naphth.) 7.52 (d, 2H, J = 8.6 Hz, CH naphth.), 0.31 (s, 18H, CH_3-Si). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ (ppm): 132.3, 131.7, 129.5, 127.7, 121.2, 105.0, 95.2, 0.0. MS (EI, 70 eV) m/z: 320.20 (M^+, 50), 321.20 (18), 306.2 (35), 305.20 (100). Spectra were compared with those reported in literature.[294]

2,7-diethynylnaphthalene (3). Compound **2** (380 mg, 1.19 mmol) was dissolved in THF (25 ml) followed by addition of 1M solution of NaOH (25 ml). The mixture was vigorously stirred for 2 h and then THF was removed under vacuum. Product was extracted with DCM from water phase to afford, after drying and evaporation, pure product (99 %, 206 mg). TLC (Silica gel, hexane-AcOEt 9:1) R_f = 0.57.

(3) ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 7.99 (s, 2H, CH naphth.) 7.79 (d, 2H, J = 8.6 Hz, CH naphth.), 7.56 (dd, 2H, J_1 = 8.5 Hz, J_2 = 1.5 Hz, CH naphth.), 3.20 (s, 2H, CH alkyne). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ (ppm): 132.6, 132.3, 132.0, 129.6, 127.9, 120.3, 83.6, 78.1. MS (EI, 70 eV) m/z: 176.10 (M^+,100), 177.10 (17), 150.10 (12). FT-IR ν (cm^{-1}) 2987, 2889, 1687, 1629, 1429. Decomposes above 250 °C.

1,1'-(naphthalene-2,7-diyl)diethanone (4a). Compound **3** (70 mg, 0.397 mmol) and $Fe_2(SO_4)_3$ x $5H_2O$ (16.7 mg, 0.03 mmol) were dissolved in glacial acetic acid (3 ml) and stirred at 95°C for 24 h. Reaction was followed by GC-MS. At the end of reaction, the mixture was diluted with AcOEt (150 ml) and washed with saturated $NaHCO_3$ solution (150 ml) and brine (3 x 100 ml). Organic layer was dried over sodium sulfate and evacuated under vacuum. Crude product was purified by flash column

chromatography on silica gel with hexane-AcOEt 9:1 as eluent (R_f = 0.24) to afford pure product (75%, 63 mg).

(4a) ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 8.61 (bs, 2H, CH naphth.), 8.17, (dd, 2H, J_1 = 8.7, J_2 = 1.6 Hz, CH naphth.), 7.96 (d, 2H, J = 8.7 Hz), 2.77 (s, 6H, Me). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ (ppm): 197.6, 137.7, 135.3, 131.8, 131.5, 128.4, 126.5, 26.7. MS (EI, 70 eV) m/z: 212.17 (M^+, 50), 197.10 (100), 169.10 (25), 154.10 (35), 126.10 (24). FT-IR ν (cm^{-1}) 2921, 2849, 1685, 1665, 1625, 1460, 1420, 1353, 1284, 1273, 1249, 1225, 1185, 1132. m.p. 108 – 110 °C.

1-(7-ethynylnaphthalen-2-yl)ethanone (4b). The procedure was the same as one used to synthesize product **4a**, except that reaction time was 5 h. After column chromatography pure product was obtained (42%, 47 mg). Additionally, product **4a** was also collected as a byproduct (51 %, 62 mg).

(4b) ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 8.43 (s, 1H, CH naphth.), 8.16 (s, 1H, CH naphth), 8.07 (dd, 1H, J_1 = 8.6 Hz, J_2 = 1.7 Hz, CH naphth.), 7.90 (d, 1H, J = 8.6 Hz, CH naphth.), 7.85 (d, 1H, J = 8.5 Hz, CH naphth.), 7.66 (dd, 1H, J_1 = 8.5 Hz, J_2 = 1.6 Hz, CH naphth.), 3.22 (s, 1H, CH alkyne), 2.75 (s, 3H, Me). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ (ppm): 197.8, 135.2, 135.1, 133.7, 132.1, 131.1, 129.8, 128.4, 127.9, 125.0, 120.6, 83.3, 78.3, 26.7. MS (EI, 70 eV) m/z: 194.10 (M^+, 64), 179.10 (100), 151.10 (80) 150.10 (36). FT-IR ν (cm^{-1}) 3225, 2922, 2850, 1677, 1623, 1595, 1563, 1456, 1413, 1366, 1351, 1274, 1259, 1224, 1192, 1164. m.p. 121 – 122 °C.

1,1'-(naphthalene-2,7-diyl)bis(2-bromoethanone) (5). Compound **4a** (30 mg, 0.13 mmol) was dissolved in CH_3CN (10 ml) in together with p-toluenesulfonic acid (80.5 mg, 0.39 mmol). NBS (50.2 mg, 0.26 mmol) was dissolved in CH_3CN (600 µl) and added to the solution and stirred for 2 h. Mixture was diluted with DCM and washed with brine two times. Crude product was purified by flash column chromatography on silica gel with hexane-DCM 5:95 as eluent to afford pure product (75%, 39.2 mg).

(5). ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 8.68 (s, 2H, CH naphth.), 8.21 (dd, 2H, J_1 = 8.7 Hz, J_2 = 1.7 Hz, CH naphth.), 8.02 (d, 2H, J = 8.7 Hz, CH naphth.), 4.60 (s, 4H, CH_2). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ (ppm): 190.9, 138.2, 132.4, 132.7, 131.7, 128.9, 127.3, 30.4. FT-IR ν (cm^{-1}) 2991, 2944, 1692,

1670, 1623, 1458, 1418, 1389, 1335, 1279, 1235, 1196, 1169, 1151, 1132, 1100, 1023, 1012. m.p. 155 – 177.

5-formyl uracil (6). 5-hydroxymethyl uracil (500 mg, 3.52 mmol) was dispersed in water (13 ml) and reaction carried at 70°C. Cerium ammonium nitrate was then added (4.8 g, 2.5 eq) and temperature increased to 90°C. Reaction was followed by color change from orange/yellow to transparent and after 1 h reaction was complete. Reaction mixture was cooled and kept in the refrigerator for 4 h to obtain a white precipitate that was filtered and washed with cold water to afford pure product (74%, 363 mg).

(6) ^{1}H-NMR (300 MHz, 25°C, d_6-DMSO) δ(ppm): 11.89 (bs, 1H, NH uracil), 11.50 (s, 1H, NH uracil), 9.74 (s, 1 H, CHO), 8.13 (d, 1H, J = 6.1 Hz, CH uracil). ^{13}C-NMR (75 MHz, 25°C, d_6-DMSO) δ (ppm): 186.9, 162.9, 150.9, 149.8, 110.5. FT-IR ν (cm^{-1}) 3031, 2844, 1762, 1685, 1625, 1598, 1494, 1443, 1417, 1332, 1260, 1227, 1171. m.p. decomposes at 270-280 °C.

1,3-di Dod- 5- formyl uracil (7). Compound **6** (34.3 mg, 0.25 mmol) was dissolved in dry DMF (2 ml) with 4,4'-(chloromethylene)bis(methoxybenzene) (Dod chloride, 160.9 mg, 0.61 mmol) and K_2CO_3 (84.5 mg, 0.61 mmol) and stirred at room temperature for 4 h. Mixture was then diluted with AcOEt (100 ml) and washed with brine (2 x 100 ml). Organic phase was dried over sodium sulfate and evaporated under vacuum. Crude product was purified by column chromatography on silica gel with hexane-AcOEt 8:2 as eluent (Rf = 0.2) to afford pure product (94%, 137.2 mg).

(7) ^{1}H-NMR (300 Mhz, 25°C, $CDCl_3$) δ(ppm): 9.99 (s, 1H, CHO), 7.95 (s, 1H, CH uracil), 7.31 (s, 1H, CH Dod), 7.27 (d, 4H, J = 8 Hz, CH aromatic Dod), 7.04 (d, 4H, J = 8.7 Hz, CH aromatic Dod), 6.95 (s, 1 H, CH Dod), 6.92 (d, 4H, J = 8.7 Hz, CH aromatic Dod), 6.84 (d, 4H, J = 8.7 Hz, CH aromatic Dod), 3.84 (s, 6H, OMe), 3.81 (s, 6H, OMe).

(E)-1,3-bis(bis(4-methoxyphenyl)methyl)-5-(3-oxo-3-phenylprop-1-en-1-yl)pyrimidine-2,4(1H,3H)-dione (8). Bromo acetophenone (51 mg, 0.25 mmol) was dissolved in toluene and stirred with PPh_3 (132 mg, 0.50 mmol) for 30 min at 50°C. Aldehyde **7** (99.4 mg, 0.17 mmol) was then added together with Na_2CO_3 (33.4 mg, 1.88 eq), temperature raised to 100°C and reaction carried overnight. Toluene was removed under vacuum and solid dissolved in AcOEt (100 ml) and washed with brine

(2 x 100 ml). Organic layer was dried over sodium sulfate and evaporated under vacuum. Crude product was purified by gradient flash column chromatography on silica gel (hexane, hexane-AcOEt 8:2, 7:3) to afford pure product (95%, 115 mg).

(8) ^{1}H-NMR (300 MHz, 25°C, d_6-DMSO) δ(ppm): 8.13 (s, 1H, CH uracil), 8.08 (d, 1H, J = 15.6 Hz, CH alkene), 7.91 (d, 2H, J = 7.2 Hz, CH aromatic), 7.64 (t, 1H, J = 7.2 Hz, CH aromatic), 7.55 (t, 2H, J = 7.9 Hz, CH aromatic), 7.50 (d, 1H, J = 15.1 Hz, CH alkene), 7.22 (s, 1H, CH Dod), 7.19 (d, 4H, J = 8.8 Hz, CH aromatic Dod), 7.18 (d, 4H, J = 8.8 Hz, CH aromatic Dod), 6.98 (d, 4H, J = 8.8 Hz, CH aromatic Dod), 6.88 (s, 1H, CH Dod), 6.87 (d, 4H, J = 8.7 Hz, CH aromatic Dod), 3.77 (s, 6H, OMe), 3.74 (s, 6H, OMe). ^{13}C-NMR (75 MHz, 25°C, $CDCl_3$) δ (ppm): 190.8, 159.7, 158.8, 138.1, 137.3, 132.8, 130.4, 129.9, 129.6, 129.2, 128.6, 128.5, 122.4, 114.6, 113.6, 63.3, 55.4, 55.2. FT-IR ν (cm^{-1}) 1663, 1654, 1510, 1442, 1247, 1174. m.p. 99-101 °C.

1-(bis (4-methoxyphenyl) methyl) -2,4- dioxo- 1,2,3,4- tetrahydropyrimidine -5- carbaldehyde (9). Compound **6** (140 mg, 1.00 mmol) was dissolved in dry DMF (7 ml) and stirred under nitrogen atmosphere, at room temperature, for 1 h, with Dod chloride (289 mg, 1.10 mmol) and potassium carbonate (151.8 mg, 1.10 mmol). At the end of reaction, crude product was diluted with AcOEt (125 ml) and washed with brine (2 x 100 ml). Organic layer was dried over sodium sulfate and evacuated under vacuum. Product was obtained by re-crystallization in hot acetate (30%, 110 mg). TLC (Silica gel, hexane/AcOEt 5:5) R_f = 0.45.

(9) ^{1}H-NMR (400 MHz, 25°C, d_6-DMSO) δ(ppm): 11.89 (s, 1H, NH uracil), 9.72 (s, 1H, CHO), 7.79 (s, 1H, CH uracil), 7.14 (d, 4H, J = 8.7 Hz, CH aromatic Dod), 6.98 (d, 4H, J = 8.8 Hz, CH aromatic Dod), 6.84 (s, 1H, CH Dod), 3.77 (s, 6H, OMe). ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ (ppm): 186.7, 162.2, 159.6, 150.4, 148.0, 130.2, 129.7, 114.8, 110.9, 62.9, 55.7. FT-IR ν (cm^{-1}) 3012, 2840, 1729, 1689, 1608, 1597, 1583, 1514, 1488, 1462, 1442, 1396, 1370, 1324, 1301, 1251, 1180, 1118, 1077, 1031. m.p. 198 – 199 °C.

1-Dod-3-Cbz-5-formyl uracil (10). Compound **6** (560.4 mg, 4.0 mmol) was dissolved in dry DMF (10 ml) and stirred under nitrogen atmosphere, at room temperature, for 1 h, with Dod chloride (1.1561 g, 4.4 mmol) and potassium carbonate (607.2 mg, 4.4 mmol). At the end of the reaction Cbz chloride (1.42 ml, 10.0 mmol) was added and solution stirred overnight. Mixture was diluted with AcOEt (150

ml) and washed with brine (3 x 100 ml). Organic layer was dried over sodium sulfate and evacuated under vacuum. Crude product was purified by gradient flash column chromatography on silica gel (hexane-AcOEt 9:1, 8:2, 7:3, 6:4, 0:1) to afford pure product (30%, 600 mg).

10) ^{1}H-NMR (400 MHz, 25°C, d_6-DMSO) δ(ppm): 9.78 (s, 1H, CHO), 7.84 (s, 1H, CH uracil), 7.33 – 7.20 (m, 5H, CH aromatic Cbz), 7.16 (d, 4H, J = 8.3 Hz, CH aromatic Dod), 6.98 (d, 4H, J = 8.3 Hz, CH aromatic Dod), 6.91 (s, 1H, CH Dod), 5.02 (s, 2H, CH_2 Cbz), 3.77 (s, 6H, OMe). ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ (ppm): 186.90, 161.19, 159.64, 150.74, 146.61, 136.86, 130.24, 129.56, 128.83, 128.06, 127.78, 114.84, 110.18, 64.29, 55.67, 44.46. FT-IR ν (cm^{-1}) 3005, 2961, 2839, 1718, 1698, 1663, 1602, 1586, 1511, 1496, 1462, 1450, 1443, 1420, 1380, 1353, 1330, 1310, 1300, 1279, 1248, 1229, 1177, 1146, 1105, 1022. m.p. 170 – 172 °C.

6. Modelling and synthesis of hydrolytic unit for reactive PNAs.

6.1. Introduction

As briefly explained in chapter 1, the introduction of hydrolytic units on systems able to discriminate target DNA or RNA sequences would be extremely useful in order to obtain artificial nucleases. First trials in this direction have been done modifying oligonucleotides.[295] Catalytic moieties have been inserted both on bases or backbone. An example of modified base is reported in Figure 6.1 A, in which terpyridine unit is bound to uracil through a spacer.[296] In Figure 6.1 B a second example in which a lutetium catalytic complex was bound at the 5' backbone end of a 15 mer oligonucleotide.[297]

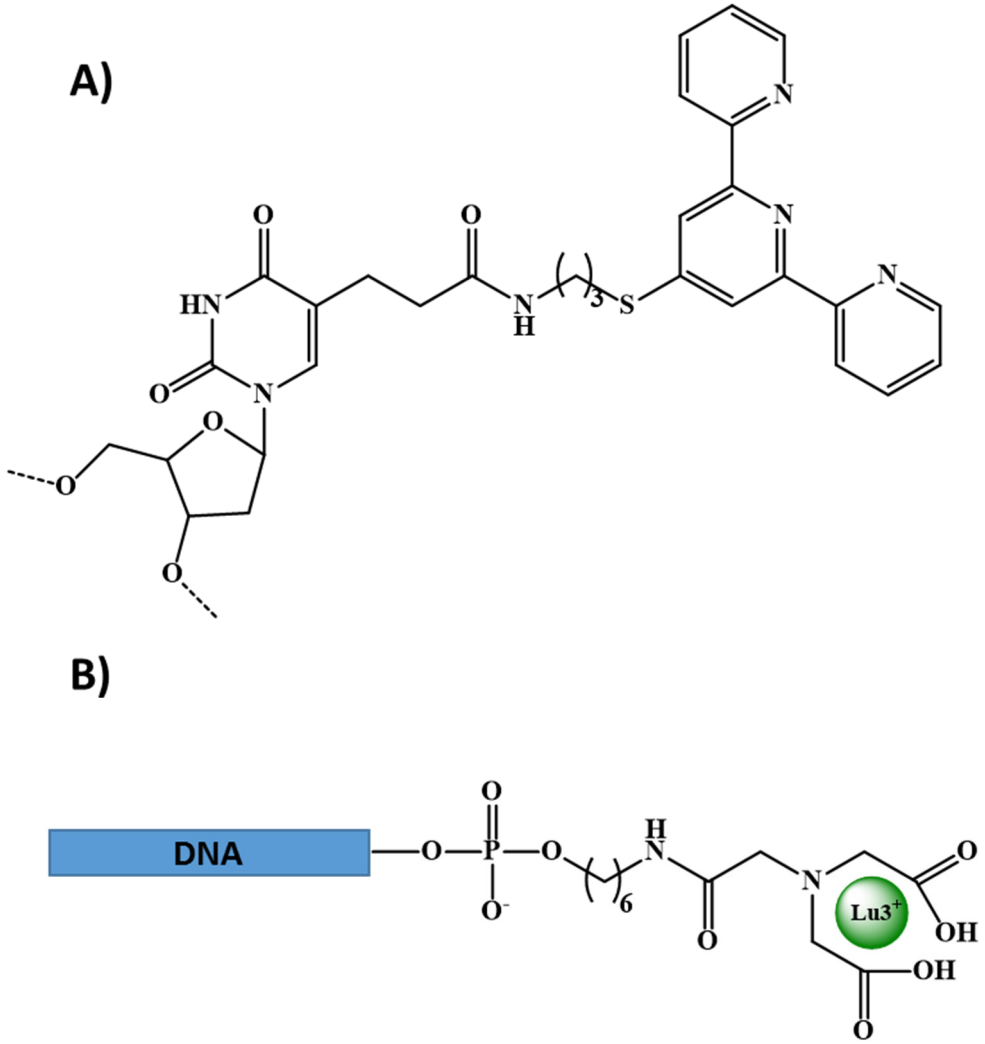


Figure 6.1: Oligonucleotides bearing hydrolytic units. A) example of base modification; B) example of backbone modification.

In a similar manner, PNAs have been also used as artificial nucleases. Stromberg and coworkers reported several examples of PNAs able to cleave RNA (Figure 6.2 A) in presence of a bulge due to

mispairing of several nucleotides on the RNA. PNA bearing phenanthroline derivative complex with Zn^{2+} and Cu^{2+} were able to hydrolyze RNA, preserving binding and selectivity typical of PNA.[153,152] They have also synthetized a metal free nuclease (Figure 6.2 B) coupling a tris(2-aminobenzimidazole) derivative to PNA.[298] Hydrolysis occurred in this case without necessity of a bulge.

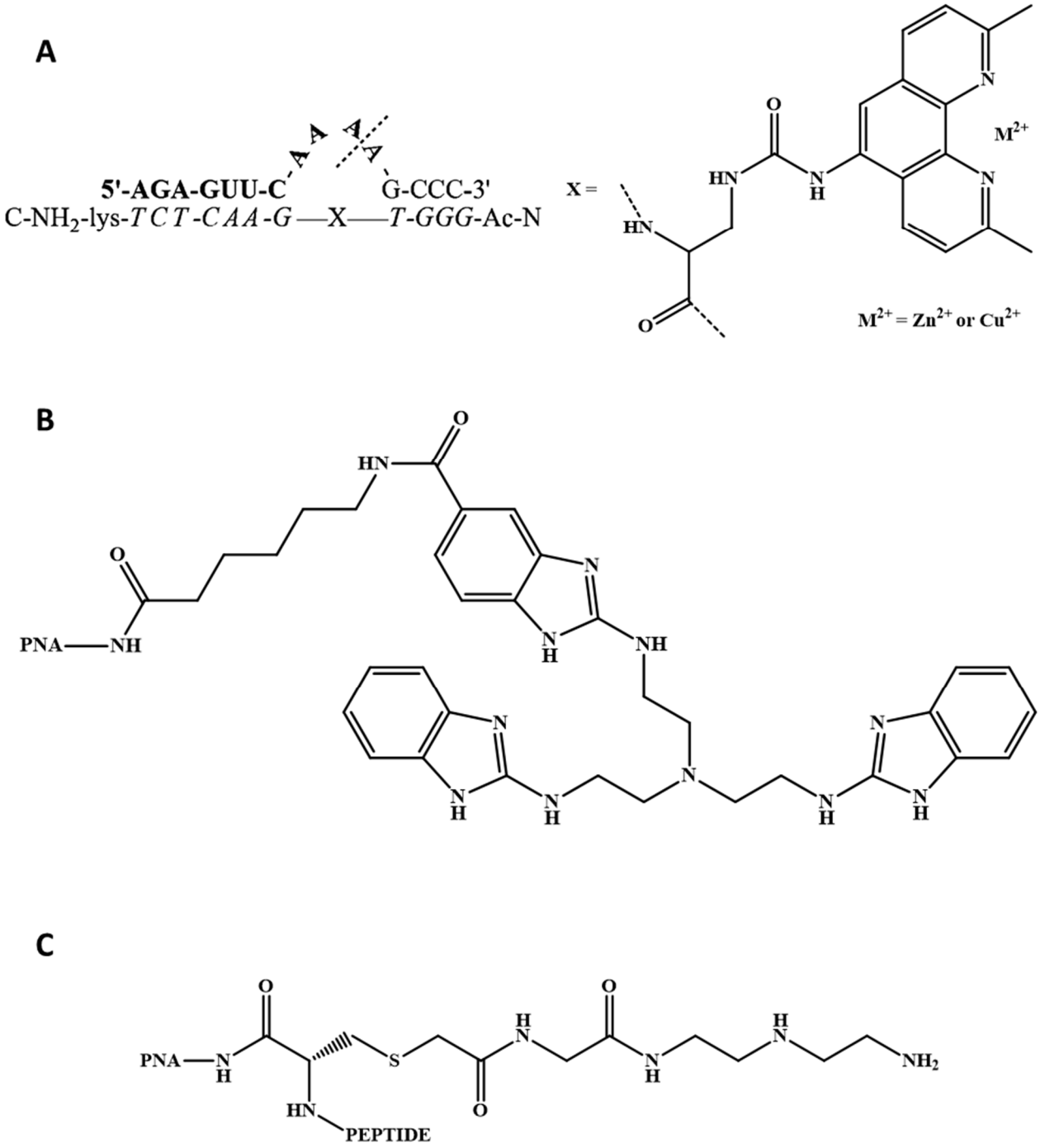


Figure 6.2: **A**) PNA bearing phenanthroline derivative interacting with RNA bulge; **B**) PNA bearing tris(2-aminobenzimidazole) derivative; **C**) example of DETA modified PNA.

Other examples of PNAs bearing hydrolytic moieties have been reported by Boom and coworkers. In particular, they attached diethylenetriamine (DETA) to N terminus of PNA backbone using different spacers (Figure 6.2 C).[299,300]

One of the most studied and effective catalytic unit used for artificial metallo-nucleases is 1,4,7-triazacyclononane (TACN) which has been used in combination with Cu^{2+} and Zn^{2+}; binding to this ligand is very strong, thus preventing metal release; the hydrolytic properties of metal-, in particular of zinc-complexes have been well studied and this unit has been coupled to different molecules retaining its activity.[301,302] Here we report the modelling studies and the synthesis of a modified triazacyclononane (Figure 6.3), with a carboxylic moiety that allows linkage to PNA (1,4,7-triazacylononane-N-acetate, TACNA).

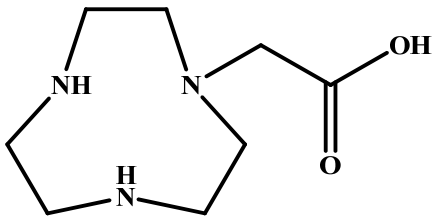


Figure 6.3: Chemical structure of the TACN derivative TACNA.

6.2. *Ab initio* calculation

TACNA is less studied complex compared to TACN,[303,304] but its hydrolytic properties have been tested with copper[305] and manganese.[306] We performed an explorative study on this compound using *ab initio* DFT calculation on zinc complex and we compared results with TACN. Initial starting geometries were generated with Avogadro, a freeware software for input preparation.[307] In order to mimic binding with a spacer linked to PNA, we have simulated TACNA capped with an *N*-ethylamide moiety. The introduction of this group is also important to remove complex extra stabilization due to carboxylic acid which could bias the results.

Calculations were performed using Gaussian 09 D1 (G09),[246] at B3LYP level of theory. Structures were optimized, in vacuum, initially with 3-21G* basis set, then refined with 6-31G and 6-31G*. For the Zn atom the LANL2DZ core potential was used. To obtain more reliable results, the final optimized structures were refined with implicit solvent (Polarizable Continuum Model, PCM)[308], implemented in G09, using 6-31G* as basis set. Finally in order to fill vacant coordination position

we inserted an explicit molecule of water. RESP charges (Figure 6.4) were calculated with procedure explained in chapter 2.

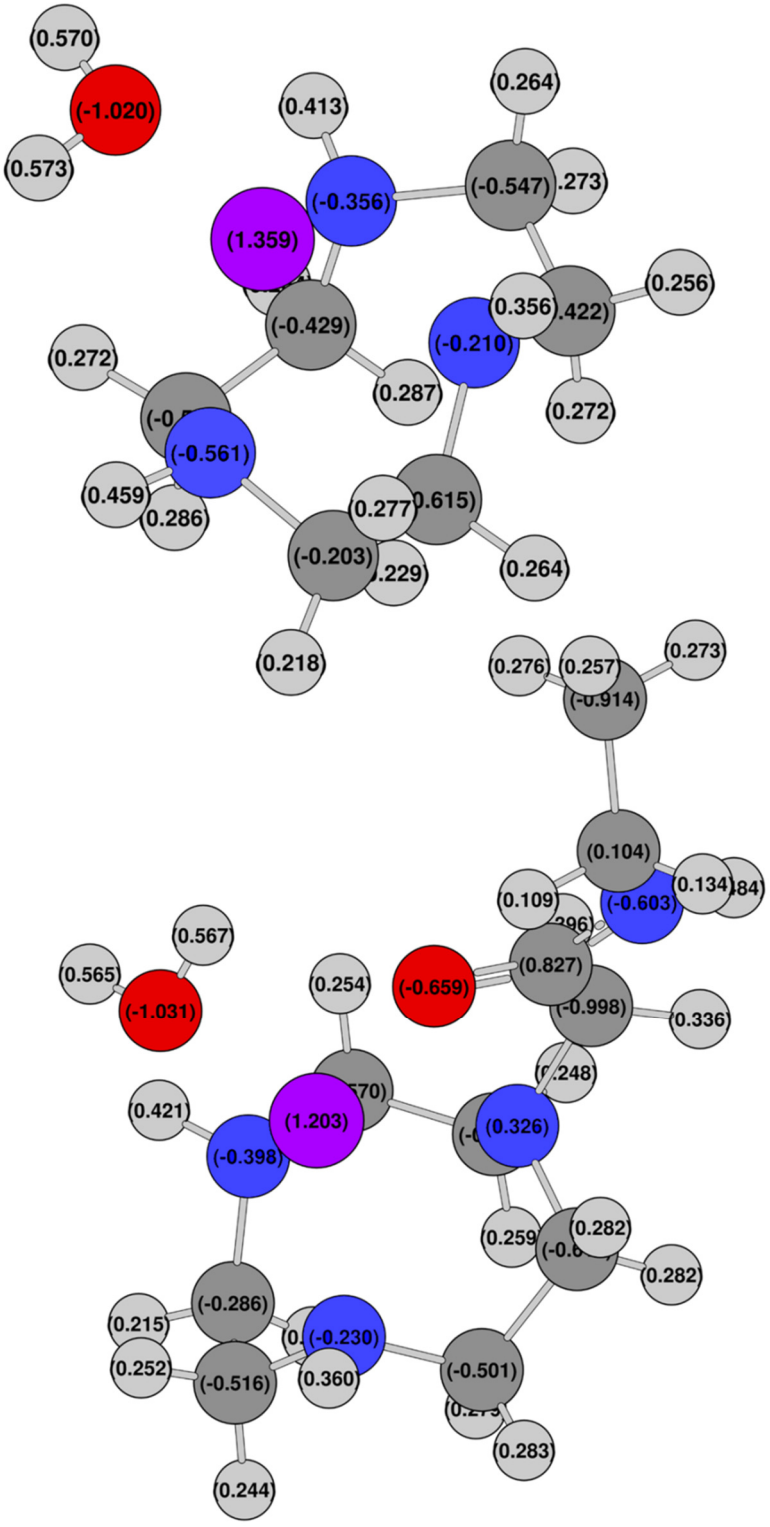


Figure 6.4: RESP charges for TACN (top) and TACNA (bottom) complexes with ZINC (violet).

Acetamide group on TACNA affects the charge distribution of the cycle, changing in particular that of bonded nitrogen (from -0.210 to +0.326). This may slightly reduce binding properties compared

to TACN, but the effect is counterbalanced by an additional interaction with carbonyl group. Partial charge on zinc was reduced (from 1.359 to 1.203), but fortunately this poorly affected polarization of coordinated water molecule. We also considered stability of the complexes by evaluation of free energy variation. For Zinc we decided to consider an octahedral complex $[Zn(H_2O)_6]^{2+}$, since in water it should be preferred to the tetrahedral one[309,310].

Reactions involved are then:

$$Zn^{2+} + 6\ H_2O \rightarrow [Zn(H_2O)_6]^{2+}$$

and consequently

$$[Zn(H_2O)_6]^{2+} + TACN \rightarrow [Zn(TACN)(H_2O)]^{2+} + 5\ H_2O$$

and

$$[Zn(H_2O)_6]^{2+} + TACNA \rightarrow [Zn(TACNA)(H_2O)]^{2+} + 5\ H_2O$$

From ΔG^0 value it is possible to calculate stability constants **β** of complexes (log $\boldsymbol{\beta_{TACNA}}$ = 42.03, log $\boldsymbol{\beta_{TACN}}$ = 46.4). These values in themselves are not fully meaningful since to have more accurate and realistic results equilibrium between species, second sphere coordination and other factors should be considered. However, assuming that approximations are equal in both cases, their ratio represents relative stability of the complexes:

$$\frac{log\ \beta_{TACNA}}{log\ \beta_{TACN}} = 0.92$$

From these results we deduced that complex with TACNA has a stability comparable to TACN, and considering that experimental log $\boldsymbol{\beta_{TACN}}$ is 7.18,[311] we expect a log $\boldsymbol{\beta_{TACNA}}$ of 6.61, making it suitable for zinc complexation. Moreover coordinated water molecule exhibited polarization similar to that one found in unmodified cycle, suggesting retention of hydrolytic properties. Optimized structures obtained from this study are reported in Appendix 8.8.

6.3. Synthesis of modified TACNA

In Figure 6.5 it is reported structure of modified TACNA synthetized during PhD activity. Protection of amino groups was necessary in order to avoid side reactions during PNA synthesis.

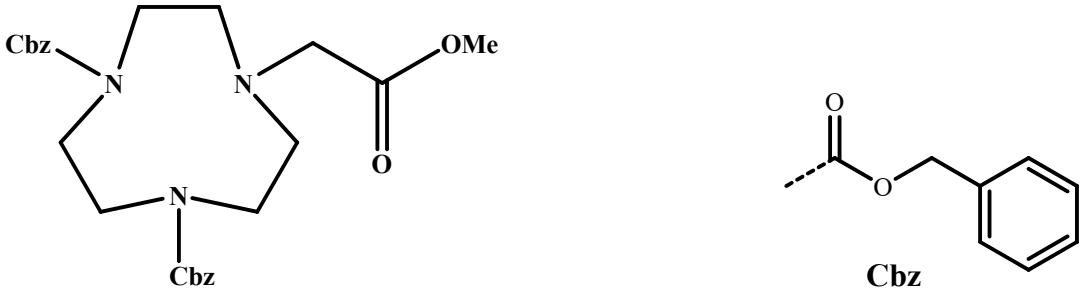
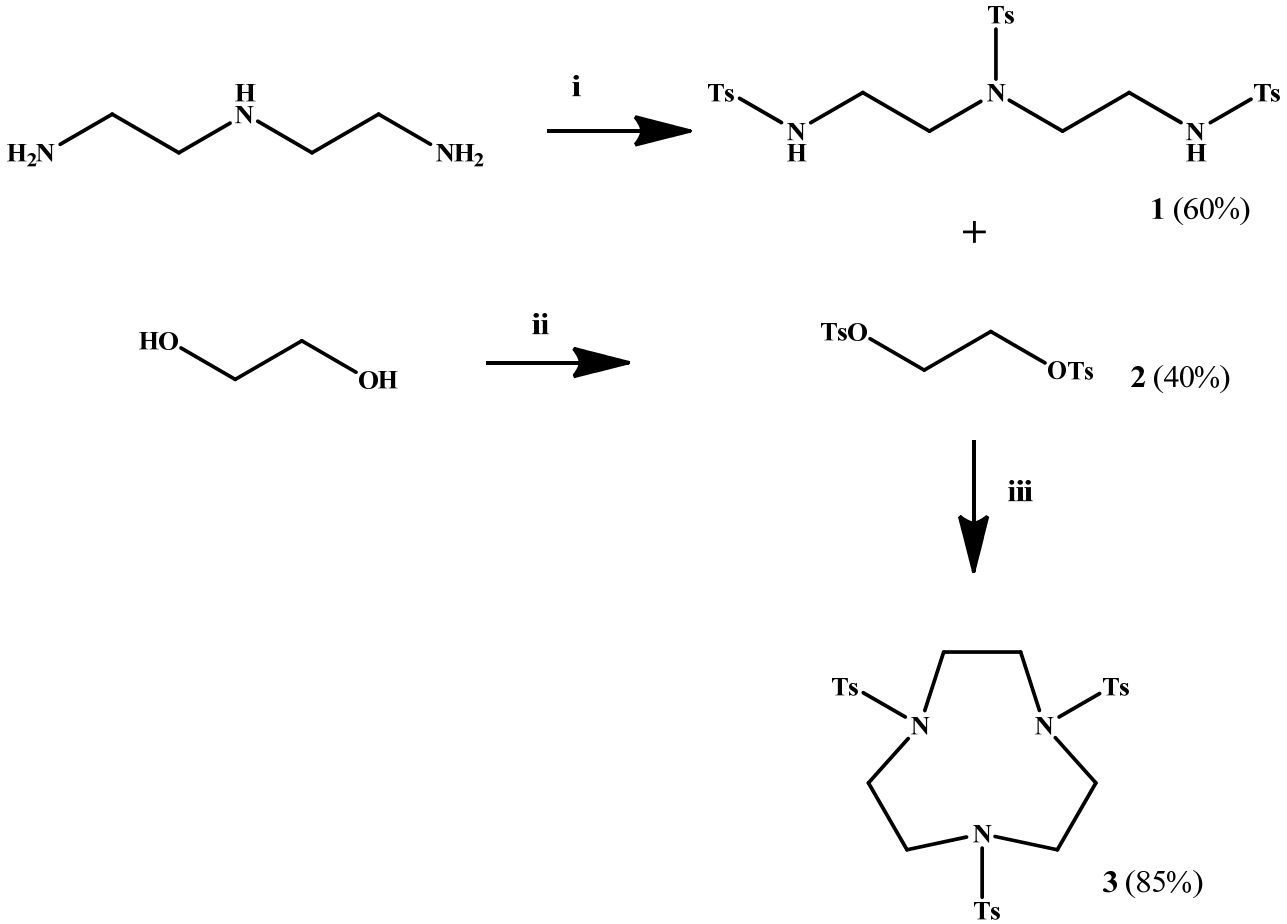


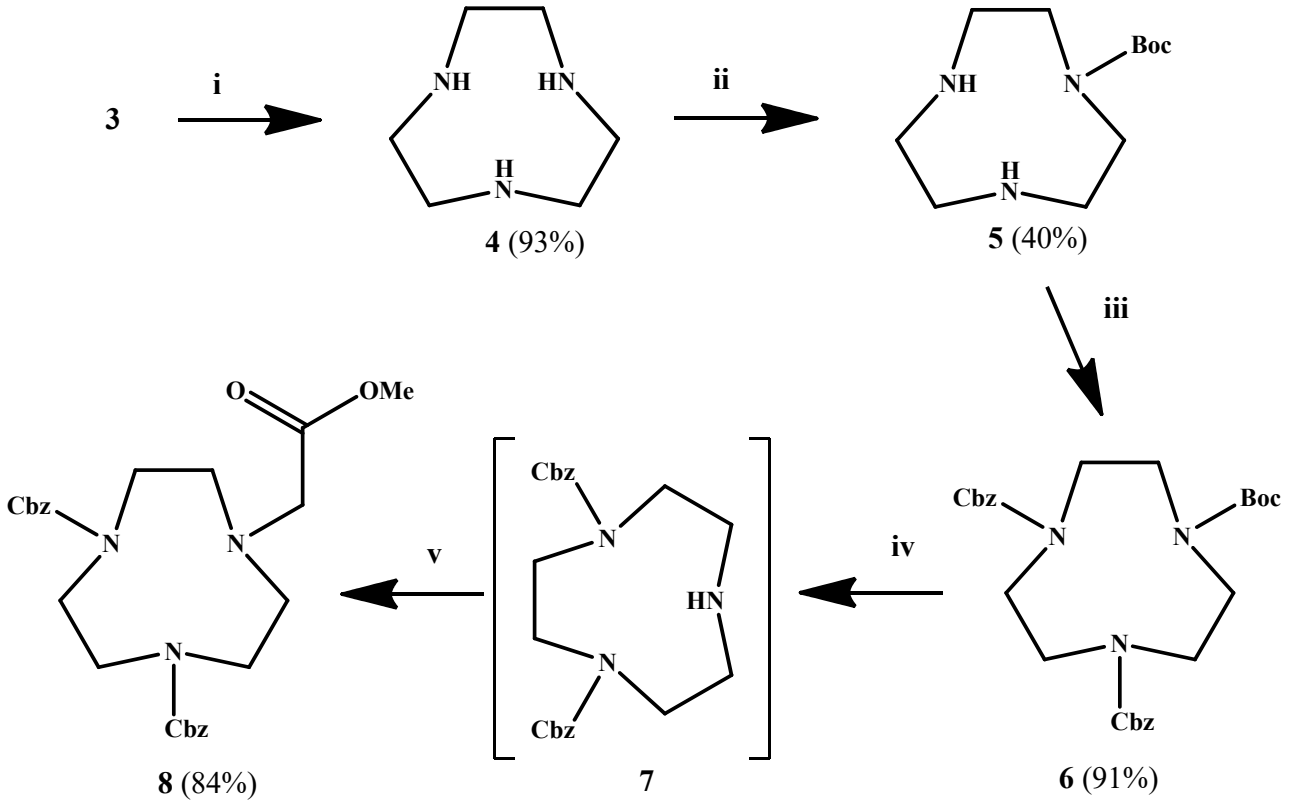
Figure 6.5: Designed hydrolytic unit for coupling with PNA. Original TACNA structure was protected with Cbz in order to avoid side reaction during PNA chain growth.



Scheme 6.1: Synthesis of TACN 3Ts. (i) [K_2CO_3, TsCl, dry DMF/DCM 1:1, o.n., r.t.]; (ii) [(a) TsCl, K_2CO_3, pestle 40 min; (b) NaOH, pestle 10 min]; (iii) [N_2, CsCO3, dry DMF, o.n. r.t.].

Synthesis of TACN was performed starting from two building blocks: diethylenetriamine tritosilate (DETA 3Ts) and glycol ditosilate.[312] DETA 3Ts was obtained from DETA protecting amino groups with tosyl chloride and product **2** was obtained from reaction of ethylene glycol with tosyl chloride, with

a solid phase reaction. Low yields obtained in this case were probably due to presence of water in the starting reagent. In order to favor cyclization of compounds **2** and **3** over polymerization, reaction was conducted in diluted solutions. Cesium carbonate was also necessary to generate reactive nucleophile (loose ion pair).[312]



Scheme 6.2: Synthesis of modified TACNA. (i) [(a) conc. H_2SO_4, N_2, 30 min, 160°C; (b) HCl 37%, EtOH; (c) H_2O/NaOH, continuously extracted with CH_3Cl]; (ii) [Boc_2O, Phosphate buffer (pH = 6.8), o.n., r.t.]; (iii) [Cbz chloride, K_2CO_3, CH_3CN, o.n., r.t.]; (iv) [40% TFA/DCM, 1h, r.t.]; (v) [$BrCH_2COOMe$, K_2CO_3, CH_3CN, o.n., r.t.].

Compound **3** was deprotected using hot concentrated sulfuric acid and the crude was precipitated as hydrochloric salt. This was then dissolved in 1M NaOH solution and continuously extracted with chloroform. Product **4** was mono protected with $(Boc)_2O$ in presence of phosphate buffer (pH = 6.8). Without pH control, major products were the macrocycle bearing three Boc-groups and unprotected TACN, while mono- and di-protection occurred in traces; with pH control it was possible to obtain about 40% of **5** with 10% of three protected compound and traces of di protected. About 50% remained unreacted TACN. At the end of the reaction it was possible to extract with chloroform a first fraction containing Boc products and then, rising pH of the solution, to collect unreacted TACN. Product **6** was obtained protecting **5** with Cbz chloride. Treating **6** with a 40% solution of TFA

in DCM allowed selective deprotection of Boc moiety since Cbz protecting groups require harsher acidic conditions. Intermediate **7** was then made react with bromoacetate to obtain product **8**.

6.4. Molecular Dynamics simulation on Model 1 with coupled TACNA

Positioning of hydrolytic moieties on PNA is not trivial, in particular if selective cut is needed. The model developed in the previous chapter could be very useful for properly directing TACNA hydrolytic moiety towards the target phosphate groups. The second uracil of the dimeric base described in Chapter 5 could allow this catalytic unit to be precisely localized in the PNA:DNA or PNA:RNA duplex in an inner base, unlike many derivatives reported in the literature, thanks to the vicinity of the latter to the DNA/RNA strand. In order to better focus on this possibility we have performed a Molecular Dynamics simulation after introduction of the TACNA unit. The PNA:DNA duplex simulated using the Model 1 structure of Chapter 5, which was then modified to bear the hydrolytic unit linked in N1 position, in place of methyl group, through an ethyl amino linker (Figure 6.6).

From the structures obtained, the hydrolytic moiety was found to be in a favorable position and could effectively interact with target phosphate even if, as shown before, Model1 was not so efficient in adenine targeting. From simulation TACNA and Zn^{2+} presented average distances from target phosphate of respectively 3.47 ± 0.10 Å and 1.83 ± 0.11 Å.

Interaction with the target was quite stable and durable even if one considers that Molecular Dynamics tend to overestimate coulombic interactions between charged groups.[313]

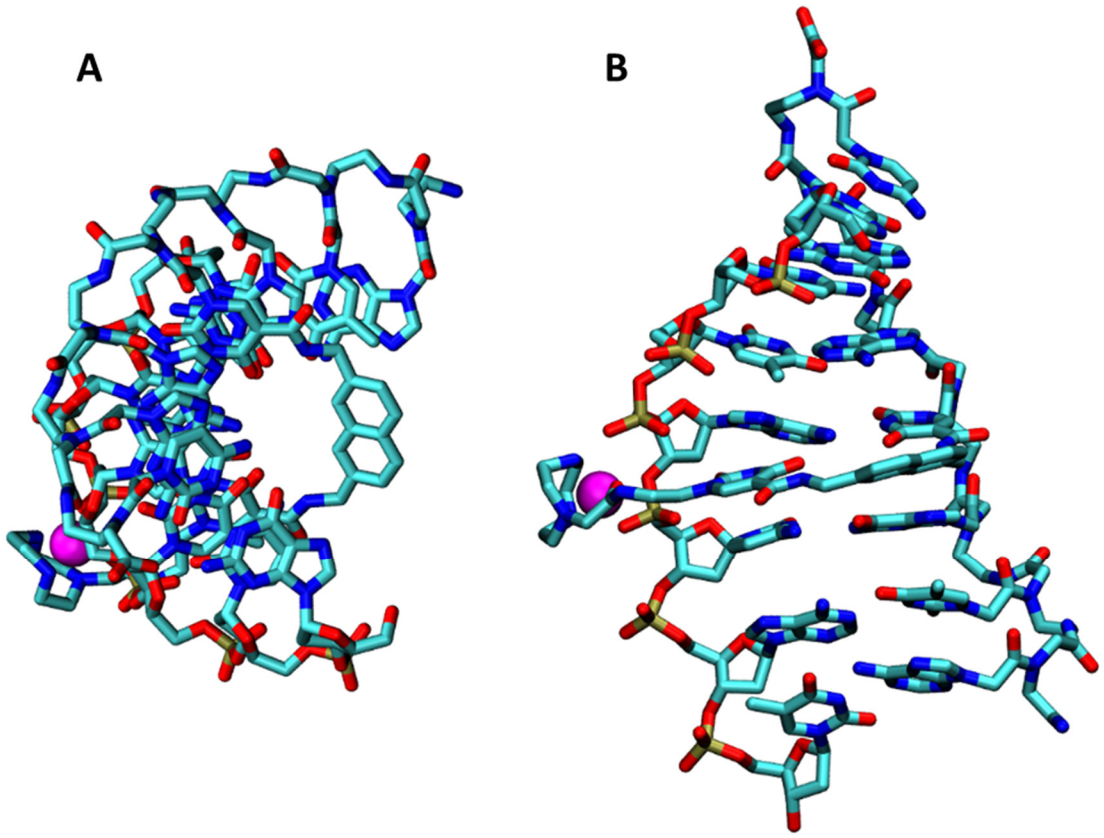
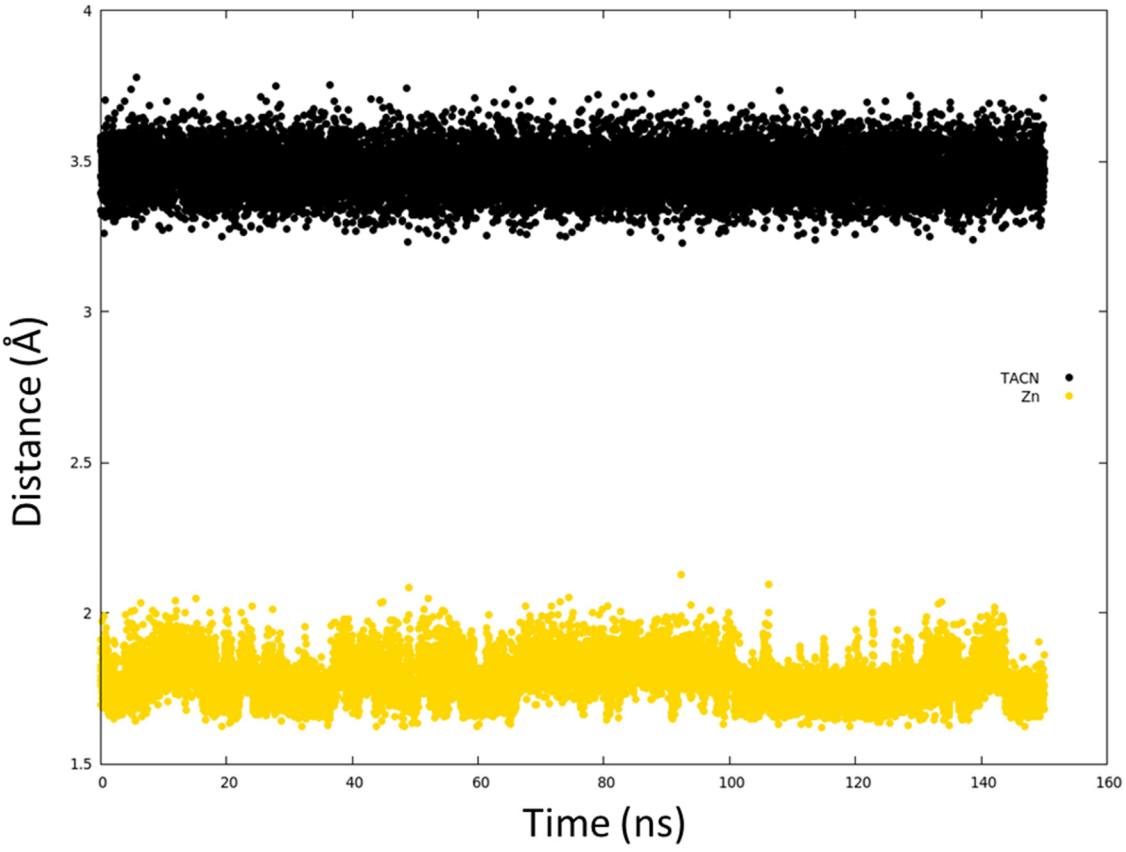


Figure 6.6: Structure of PNA:DNA duplex containing TACNA bonded to Model 1 dimeric base. **A**) top view; **B**) side view. Zn ion is shown in magenta. The system was solvated with TIP3P water in a cubic box of 66 X 63 X 74 $Å^3$ for a total of 30100 atoms.



Graph 6.1: Distance between phosphate and TACN (black) or Zinc ion (gold).

6.5. Conclusions and future works

Ab initio calculations on TACNA complex and MD simulation on a duplex containing dimeric base linked to TACNA suggested a possible combination of these elements to obtain a favorable disposition of the catalytic unit with respect to phosphate groups of complementary DNA. This should be a starting point for the design of new artificial nucleases. The vicinity effect was evaluated using Model1 modified base (chapter 5) already tested in our lab, though this presents a constrained interaction with adenine counterparts; further studies will be performed also on Model4 nucleobase, once the synthesis of the latter is completed.

For the final experiments, product **8**, after ester deprotection, will be coupled to the second uracil base of PNA, in order to test the hydrolytic properties of the final engineered molecule.

The aim of this work is therefore, in perspective, to endow PNA not only with DNA/RNA binding properties, but also with the ability to cleave the target sequence and release it, thus acting as a real artificial nuclease. Application of a similar system can have an impact both in the production of more effective toolkits for molecular biologists, and in the ability to selectively an efficiently deactivate unwanted gene expression, as in the case of genetic diseases or cancer.

6.6. Experimental section

General. Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, LINK and used without further purifications. TLC were run on Supelco 56524-25EA silica gel on aluminium foils with fluorescence indicator 254 nm. Column chromatography were performed as flash chromatography on Merck silica 60 (0.040 – 0.063) under 0.1-0.2 bar of air overpressure. NMR spectra were obtained on Bruker Avance 300 MHz or 400 MHz spectrometers. δ values are expressed in ppm. FT-IR were recorder on Thermo Nicolet 5700. ESI-MS spectra were recorded on a Acquity Ultra Performance LC. Meting points were measured with GALLENKAMP melting point apparatus.

Ab initio and Molecular Dynamics simulations. Calculations were performed on Brutus cluster, the central high-performance cluster of ETH Zurich.[290] Ab initio calculation were performed by means of Gaussian G09 D1 using 8 cores. Molecular Dynamics simulation were performed using 32 cores and were conducted in the conditions reported in section 2.5.

4-methyl-N,N-bis(2-(4-methylphenylsulfonamido)ethyl)benzenesulfonamide (DETA 3Ts, 1). DETA (2g, 19.385 mmol) was vigorously stirred in dry DMF (10 ml) with K_2CO_3 (9.36 g, 67.846 mmol), at room temperature. Tosyl chloride (12.82 g, 63.971 mmol) was dissolved in DCM (30 ml) and slowly added to the mixture over a period of 2 h. Solution was kept reacting overnight. Solvent was evacuated under vacuum and white solid partitioned between brine (250 ml) and DCM (250 ml). Aqueous phase was extracted several times with DCM. Compound **1** was purified by crystallization in DCM-MeOH 1:3 (60%, 6.5 g). TLC (Silica gel, hexane-AcOEt 5:5, R_f = 0.49).

(1) ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 7.78 (d, J = 8.1 Hz, 4H, CH aromatic of external Ts), 7.64 (d, J = 8.1 Hz, 2H, CH aromatic internal Ts), 7.33 (d, J = 8.1 Hz, 4H, CH aromatic of external Ts), 7.31 (d, J = 8.1 Hz, 2H, CH aromatic internal Ts), 5.24 (t, J = 4.7, 2H, NH), 3.25 – 3.11 (m, 8H, CH_2), 2.45 (s, 9H, Me Ts). ^{13}C-NMR (100 MHz, 25°C, d_6-DMSO) δ(ppm): 143.9, 143.2, 137.8, 135.8, 130.4, 130.2, 127.3, 127.0, 48.8, 42.0, 21.4. FT-IR ν (cm^{-1}) 3285, 2910, 1598, 1495, 1438, 1423, 1372, 1335, 1322, 1307, 1256, 1154, 1092, 1077. Full characterization is reported in literature.[314]

Ethylene glycol di tosylate (2). Ethylene glycol (1g, 16.111 mmol), K_2CO_3 (13.4 g, 96.666 mmol) and tosyl chloride (9.2 g, 48.333 mmol) were placed in a mortar. Reagents were ground together for about 40 min (when reaction started becoming very exothermic). Excess of TsCl was quenched with NaOH (3.22g, 80.555 mmol). The resulting solid was transferred into a flask, AcOEt was added (150 ml) and kept stirring for 10 min. Organic phase was filtered on sintered glass filter (G4) and concentrated under vacuum. Solid was crystallized from hot methanol to afford pure product (40%, 2.35 g). TLC (Silica gel, hexane-AcOEt 3:7, R_f = 0.62).

(2) ^{1}H-NMR (300 MHz, 25°C, $CDCl_3$) δ(ppm): 7.75 (d, 4H, J= 8 Hz, CH aromatic), 7.36 (d, 4H, J= 8 Hz, CH aromatic), 4.20 (s, 4H, CH_2), 2.48 (s, 6H, Me). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ(ppm): 145.3, 132.3, 130.0, 128.0, 66.7, 21.7. Full characterization is reported in literature.[315,316]

1,4,7-Triazacylononane tritosylate (3). Compound **1** (1.3 g, 2.310 mmol) was dissolved in dry DMF (15 ml), under nitrogen. Cs_2CO_3 (1.582 g, 4.851 mmol) was added and mixture kept stirring for 1 h. Compound **2** (865.2 mg, 2.333 mmol) was dissolved in dry DMF (6 ml), slowly added over a period of 2 h and reaction conducted overnight. Crude product was divided into four centrifuge tubes (50

ml). They were filled with water and ice and kept at 4 °C for 4 h. Solid was separated by centrifuge and washed with ethanol and ethyl ether (85%, 1.2 g). TLC (Silica gel, hexane-AcOEt 5:5, R_f = 0.66).

(3) ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 7.72 (d, J = 8.2 Hz, 6H, CH aromatic), 7.35 (d, J = 8.2 Hz, 6H, CH aromatic), 3.44 (s, 12H, CH_2), 2.46 (s, 9H, Me). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ(ppm): 143.9, 134.6, 129.9, 127.5, 51.9, 21.5. MS (ESI+) m/z:$[M+Cs]^+$ calcd. 724.06, found 724.24. FT-IR ν (cm^{-1}) 1336, 1320, 1308, 1161, 1149, 1089. m.p. 218 – 219 °C.

1,4,7-Triazacyclononane (4). Compound **3** (3.4 g, 5.74 mmol) was introduced in a tube and stirred with concentrated H_2SO_4 (6 ml) at 160°C for 30 min, under nitrogen. Tube was cooled to room temperature and cold ethanol was slowly added (3 ml), followed by addition of ethyl ether (45 ml). Mixture was kept in freezer for 4 h. Precipitate was separated by centrifugation and washed with diethyl ether. Concentrated HCl was then added (5 ml) and solid sonicated till dissolved. Ethanol and diethyl ether were added again and mixture kept in freezer overnight. White solid was separated by centrifugation and washed with diethyl ether. Solid was dissolved in 1 M NaOH (pH > 12) and then continuously extracted with chloroform for 48 h to afford pure product (93%, 693.7 mg, viscous liquid).

(4) ^{1}H-NMR (300 MHz, 25°C, $CDCl_3$) δ(ppm): 2.80 (s, 12H, CH_2), 2.38 (bs, 3H, NH). ^{13}C-NMR (75 MHz, 25°C, $CDCl_3$) δ(ppm): 47.2. MS (ESI+) m/z: $[M+H]^+$ calcd. 130.13, found. 130.23. FT-IR ν (cm^{-1}) 3339, 2919, 1646, 1449, 1402, 1350, 1282, 1149.

tert-butyl 1,4,7-triazonane-1-carboxylate (5). TACN (350 mg, 2.709 mmol) was dissolved in water (150 ml) containing a phosphate buffer (pH = 6.8, 186 mM). Boc anhydride (1.182 g, 5.418 mmol) was dissolved in water/methanol 1:1 mixture (40 ml) and slowly added in the flask with buffered solution of TACN. Mixture was stirred overnight. Water solution was extracted with chloroform (3 x 200 ml) to separate the first fraction containing mainly three Boc protected byproduct. Water solution pH was adjusted to ~10 and product continuously extracted with $CHCl_3$ overnight to collect crude mono Boc product. Concentrated organic layer was purified using gradient flash chromatography (hexane-AcOEt 7:3, 5:5, 4:6 + 1% NH_3, 2:8 + 1% NH_3, AcOEt +1 % NH_3) to afford pure product (40%, 250 mg, viscous liquid). TLC (Silica gel, hexane-AcOEt 5:5, Rf = 0.33). Rising pH of water solution above 12 was possible to collect unreacted TACN.

(5) ^{1}H-NMR (300 MHz, 25°C, $CDCl_3$) δ(ppm): 5.73 (bs, 2H, NH), 3.75 – 2.65 (m, 12H, CH_2), 1.49 (s, 9H, ^{t}Bu). MS (ESI+) m/z: $[M+H]^+$ calcd. 230.19, found 230.05, $[2M+H]^+$ calcd. 459.37, found 459.52, $[M-{}^{t}Bu+H]^+$ calcd. 174.12, found 173.99. FT-IR ν (cm^{-1}) 3298, 3166, 2975, 2917, 2848, 1659, 1458, 1397, 1364, 1248, 1149.

1,4-dibenzyl 7-tert-butyl 1,4,7-triazonane-1,4,7-tricarboxylate (6). Compound **5** (147.6 mg, 0.643 mmol) was dissolved in CH_3CN (15 ml) in the presence of K_2CO_3 (266.6 mg, 1.929 mmol). Benzyl chloroformate (218 µl, 1.929 mmol) was added and mixture stirred overnight. Solvent was evacuated under vacuum and rinsed with AcOEt (150 ml). Organic phase was washed with brine (3 x 125 ml), dried over Na_2SO_4 and concentrated under vacuum. The crude product was purified by gradient flash chromatography (hexane-AcOEt 75:25, 7:3, 6:4, 5:5) to afford the pure product (91%, 290 mg, viscous liquid). TLC (Silica gel, hexane-AcOEt 7:3, Rf = 0.3, with ninhydrin yielded blue spot).

(6) ^{1}H-NMR (300 MHz, 25°C, $CDCl_3$) δ(ppm): 7.40 – 7.29 (m, 10H, CH Cbz), 5.20 – 4.95(m, 4H, CH_2 Cbz), 3.60 – 3.30 (m, 12H, CH_2), 1.49 – 1.43 (m, 9H, ^{t}Bu). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ(ppm): 156.4, 155.7, 136.6, 128.5, 128.1, 128.0, 80.1, 67.3, 50.7, 50.4, 50.1, 50.0, 49.9, 49.7, 49.6, 49.5, 49.4, 49.3, 49.1, 49.0, 48.9, 48.8, 48.7, 28.4. MS (ESI+) m/z: $[M+H]^+$ calcd. 498.26, found 498.40, $[M+Na]^+$ calcd. 520.24, found 520.44, $[M+K]^+$ calcd. 536.22, found 536.38, $[2M+Na]^+$ calcd. 1017.49, found 1017.75, $[M-{}^{t}Bu+H]^+$ calcd. 442.20, found 442.35, $[M-Boc+H]^+$ calcd. 398.21, found 398.35. FT-IR ν (cm^{-1}) 2973, 2924, 1688, 1497, 1466, 1419, 1408, 1361, 1317, 1231, 1167, 1135, 1090, 1034.

Dibenzyl 7-(2-methoxy-2-oxoethyl)-1,4,7-triazonane-1,4-dicarboxylate, TACNA (8). Compound **6** (200 mg, 0.402 mmol) was dissolved in 40% TFA solution in DCM (5 ml) and stirred for 1 h. Solvent was removed under vacuum and crude product dissolved in CH_3CN (5 ml). K_2CO_3 (500 mg, 3.618 mmol) was added together with 2-bromoacetate (92 ul, 0.965 mmol) and reaction conducted overnight. At the end of reaction, suspension was filtered on G5 sintered-glass filter to remove KBr and solvent. The crude product was dissolved in water and continuously extracted with chloroform overnight. The extract was purified by preparative HPLC (Phenomenex Jupiter RP C18, 5µM, 300 Å, 250 X 15 mm), under isocratic conditions (water-CH_3CN 50:50) affording to pure **8** (84%, 158.8 mg).

(**8**) ^{1}H-NMR (400 MHz, 25°C, $CDCl_3$) δ(ppm): 7.45 – 7.30 (m, 10H, CH aromatic), 5.2 (M), 5.19 (m), 5.15 (m) (s, 4H, $\underline{CH_2}$ Cbz), 4.12 (M), 4.03 (m) (s, 2H, $\underline{CH_2}$COOMe), 3.79 (s, 3H, OMe), 3.76 – 3.17 (m, 12H, CH cycle). ^{13}C-NMR (100 MHz, 25°C, $CDCl_3$) δ(ppm): 166.9, 156.3, 155.8, 136.0, 128.7, 128.5, 68.1, 55.2, 53.2, 52.8, 49.2, 48.5, 46.6,46.3, 45.4. MS (ESI+) m/z: $[M+H]^+$ = 471.06, $[2M+H]^+$ = 939.67. FT-IR v (cm^{-1}) 2953, 1752, 1696, 1674, 1499, 1466, 1423, 1413, 1372, 1298, 1237, 1174, 1155, 1129, 1089. m.p. 152 – 153 °C.

7. References

(1) Demain, A. L.; Sanchez, S. *J. Antibiot.* **2009**, *62* (1), 5–16.

(2) Nantajit, D.; Lin, D.; Li, J. J. *J. Cancer Res. Clin. Oncol.* **2014**, 1697–1713.

(3) Noy, R.; Pollard, J. W. *Immunity* **2014**, *41* (1), 49–61.

(4) Sarnpitak, P.; Mujumdar, P.; Taylor, P.; Cross, M.; Coster, M. J.; Gorse, A.-D.; Krasavin, M.; Hofmann, A. *Biotechnol. Adv.* **2015**, *33* (6), 941–947.

(5) Osborn, M. J.; Freeman, M.; Huennekens, F. M. *Exp. Biol. Med.* **1958**, *97* (2), 429–431.

(6) Longley, D. B.; Harkin, D. P.; Johnston, P. G. *Nat. Rev. Cancer* **2003**, *3* (5), 330–338.

(7) Wang, W. C.; Ware, R. E.; Miller, S. T.; Iyer, R. V; Casella, J. F.; Minniti, C. P.; Rana, S.; Thornburg, C. D.; Rogers, Z. R.; Kalpatthi, R. V; Barredo, J. C.; Brown, R. C.; Sarnaik, S. A.; Howard, T. H.; Wynn, L. W.; Kutlar, A.; Armstrong, F. D.; Files, B. A.; Goldsmith, J. C.; Waclawiw, M. A.; Huang, X.; Thompson, B. W. *Lancet (London, England)* **2011**, *377* (9778), 1663–1672.

(8) Bouhnik, Y.; Scemama, G.; Taï, R.; Matuchansky, C.; Rambaud, J.-C.; Lémann, M.; Modigliani, R.; Mary, J.-Y. *Lancet* **1996**, *347* (8996), 215–219.

(9) Büchner, T.; Urbanitz, D.; Emmerich, B.; Fischer, J. T.; Fülle, H.-H.; Heinecke, A.; Hossfeld, D. K.; Koeppen, K. M.; Labedzki, L.; Löffler, H.; Nowrousian, M. R.; Pfreundschuh, M.; Pralle, H.; Rühl, H.; Wendt, F.-C. *Leuk. Res.* **1982**, *6* (6), 827–831.

(10) William P. Mcguire, M.D., William J. Hoskins, M.D., Mark F. Brady, B.S., Paul R. Kucera, M.D., Edward E. Partridge, M.D., Katherine Y. Look, M.D., Daniel L. Clarke-Pearson, M.D., And Martin Davidson, M. . *N. Engl. J. Med.* **1996**, *334* (1), 1–6.

(11) Weller, M.; Fisher, B.; Taphoorn, M. J. B.; Belanger, K.; Brandes, A. a; Marosi, C.; Bogdahn, U.; Curschmann, J.; Janzer, R. C. *Cancer/Radiothérapie* **2005**, *9* (3), 196–197.

(12) Hiddemann, W.; Kneba, M.; Dreyling, M.; Schmitz, N.; Lengfelder, E.; Schmits, R.; Reiser, M.; Metzner, B.; Harder, H.; Hegewisch-Becker, S.; Fischer, T.; Kropff, M.; Reis, H.-E.; Freund, M.; Wörmann, B.; Fuchs, R.; Planker, M.; Schimke, J.; Eimermacher, H.; Trümper, L.; Aldaoud, A.; Parwaresch, R.; Unterhalt, M. *Blood* **2005**, *106* (12), 3725–3732.

(13) Sandoe, J.; Eggan, K. *Nat. Neurosci.* **2013**, *16* (7), 780–789.

(14) Bonnamain, V.; Neveu, I.; Naveilhan, P. *Front. Cell. Neurosci.* **2012**, *6*, 17.

(15) Blurton-Jones, M.; Kitazawa, M.; Martinez-Coria, H.; Castello, N. A.; Müller, F.-J.; Loring, J. F.; Yamasaki, T. R.; Poon, W. W.; Green, K. N.; LaFerla, F. M. *Proc. Natl. Acad. Sci. U. S. A.* **2009**,

106 (32), 13594–13599.

(16) June, C. H. *J. Clin. Invest.* **2007**, *117* (6), 1466–1476.

(17) Lamers, C. H. J.; Sleijfer, S.; Vulto, A. G.; Kruit, W. H. J.; Kliffen, M.; Debets, R.; Gratama, J. W.; Stoter, G.; Oosterwijk, E. *J. Clin. Oncol.* **2006**, *24* (13), e20–e22.

(18) Rosenberg, S. A.; Packard, B. S.; Aebersold, P. M.; Solomon, D.; Topalian, S. L.; Toy, S. T.; Simon, P.; Lotze, M. T.; Yang, J. C.; Seipp, C. A. *N. Engl. J. Med.* **1988**, *319* (25), 1676–1680.

(19) Morgan, R. A.; Dudley, M. E.; Wunderlich, J. R.; Hughes, M. S.; Yang, J. C.; Sherry, R. M.; Royal, R. E.; Topalian, S. L.; Kammula, U. S.; Restifo, N. P.; Zheng, Z.; Nahvi, A.; de Vries, C. R.; Rogers-Freezer, L. J.; Mavroukakis, S. A.; Rosenberg, S. A. *Science* **2006**, *314* (5796), 126–129.

(20) Moll, R.; Löwe, A.; Laufer, J.; Franke, W. W. *Am. J. Pathol.* **1992**, *140* (2), 427–447.

(21) Rettig, W. J.; Old, L. J. *Annu. Rev. Immunol.* **1989**, *7*, 481–511.

(22) Elvin, J. G.; Couston, R. G.; van der Walle, C. F. *Int. J. Pharm.* **2013**, *440* (1), 83–98.

(23) Chames, P.; Van Regenmortel, M.; Weiss, E.; Baty, D. *Br. J. Pharmacol.* **2009**, *157* (2), 220–233.

(24) Scott, A. M.; Wolchok, J. D.; Old, L. J. *Nat. Rev. Cancer* **2012**, *12* (4), 278–287.

(25) Bertrand Coiffier, M.D., Eric Lepage, M.D., Ph.D., Josette Brière, M.D., Raoul Herbrecht, M.D., Hervé Tilly, M.D., Reda Bouabdallah, M.D., Pierre Morel, M.D., Eric Van Den Neste, M.D., Gilles Salles, M.D., Ph.D., Philippe Gaulard, M.D., Felix Reyes, M.D., M. D. *N Engl J Med* **2002**, *346* (4), 235–242.

(26) McLaughlin, P.; Grillo-Lopez, A.; Link, B.; Levy, R.; Czuczman, M.; Williams, M.; Heyman, M.; Bence-Bruckler, I.; White, C.; Cabanillas, F.; Jain, V.; Ho, A.; Lister, J.; Wey, K.; Shen, D.; Dallaire, B. *J. Clin. Oncol.* **1998**, *16* (8), 2825–2833.

(27) Boisselier, E.; Astruc, D. *Chem. Soc. Rev.* **2009**, *38* (6), 1759–1782.

(28) Farokhzad, O. C.; Langer, R. *Adv. Drug Deliv. Rev.* **2006**, *58* (14), 1456–1459.

(29) Cho, K.; Wang, X.; Nie, S.; Chen, Z. G.; Shin, D. M. *Clin. Cancer Res.* **2008**, *14* (5), 1310–1316.

(30) Hu, C.-M.; Zhang, L. *Curr. Drug Metab.* **2009**, *10* (8), 836–841.

(31) Janib, S. M.; Moses, A. S.; MacKay, J. A. *Adv. Drug Deliv. Rev.* **2010**, *62* (11), 1052–1063.

(32) Riehemann, K.; Schneider, S. W.; Luger, T. A.; Godin, B.; Ferrari, M.; Fuchs, H. *Angew. Chem. Int. Ed. Engl.* **2009**, *48* (5), 872–897.

(33) Lian, T.; Ho, R. J. *J. Pharm. Sci.* **2001**, *90* (6), 667–680.

(34) Gradishar, W. J. *Expert Opin. Pharmacother.* **2006**.

(35) Eck, W.; Craig, G.; Sigdel, A.; Ritter, G.; Old, L. J.; Tang, L.; Brennan, M. F.; Allen, P. J.; Mason,

M. D. *ACS Nano* **2008**, *2* (11), 2263–2272.

(36) Rao, K. S.; Reddy, M. K.; Horning, J. L.; Labhasetwar, V. *Biomaterials* **2008**, *29* (33), 4429–4438.

(37) Tkachenko, A. G.; Xie, H.; Coleman, D.; Glomm, W.; Ryan, J.; Anderson, M. F.; Franzen, S.; Feldheim, D. L. *J. Am. Chem. Soc.* **2003**, *125* (16), 4700–4701.

(38) Zhao, J.; Mi, Y.; Liu, Y.; Feng, S.-S. *Biomaterials* **2012**, *33* (6), 1948–1958.

(39) Liu, Y.; Li, K.; Pan, J.; Liu, B.; Feng, S.-S. *Biomaterials* **2010**, *31* (2), 330–338.

(40) Katz, E.; Willner, I. *Angew. Chemie Int. Ed.* **2004**, *43* (45), 6042–6108.

(41) Rossi, F.; Bedogni, E.; Bigi, F.; Rimoldi, T.; Cristofolini, L.; Pinelli, S.; Alinovi, R.; Negri, M.; Dhanabalan, S. C.; Attolini, G.; Fabbri, F.; Goldoni, M.; Mutti, A.; Benecchi, G.; Ghetti, C.; Iannotta, S.; Salviati, G. *Sci. Rep.* **2015**, *5*, 7606.

(42) Bertucci, A.; Lülf, H.; Septiadi, D.; Manicardi, A.; Corradini, R.; De Cola, L. *Adv. Healthc. Mater.* **2014**, *3* (11), 1812–1817.

(43) Yah, C. S.; Simate, G. S.; Iyuke, S. E. *Pak. J. Pharm. Sci.* **2012**, *25* (2), 477–491.

(44) Jaenisch, R.; Bird, A. *Nat. Genet.* **2003**, *33*, 245–254.

(45) Vogelstein, B.; Kinzler, K. W. *Nat. Med.* **2004**, *10* (8), 789–799.

(46) Sabrina Buchini, C. J. L. *Curr. Opin. Chem. Biol.* **2003**, *7* (6), 717–726.

(47) E. Nielsen, P. *Curr. Med. Chem.* **2001**, *8* (5), 545–550.

(48) Yamayoshi, A.; Yasuhara, M.; Galande, S.; Kobori, A.; Murakami, A. *Oligonucleotides 21* (2), 115–121.

(49) Mann, M. J.; Dzau, V. J. *J. Clin. Invest.* **2000**, *106* (9), 1071–1075.

(50) Deuss, P. J.; Arzumanov, A. A.; Williams, D. L.; Gait, M. J. *Org. Biomol. Chem.* **2013**, *11* (43), 7621–7630.

(51) Sazani, P.; Kang, S.-H.; Maier, M. A.; Wei, C.; Dillman, J.; Summerton, J.; Manoharan, M.; Kole, R. *Nucleic Acids Res.* **2001**, *29* (19), 3965–3974.

(52) Gleave, M. E.; Monia, B. P. *Nat. Rev. Cancer* **2005**, *5* (6), 468–479.

(53) CROOKE, S. T. *Antisense Nucleic Acid Drug Dev.* **1998**, *8* (2), 133–134.

(54) Lennox, K. A.; Behlke, M. A. *Pharm. Res.* **2010**, *27* (9), 1788–1799.

(55) Lu, Y.; Xiao, J.; Lin, H.; Bai, Y.; Luo, X.; Wang, Z.; Yang, B. *Nucleic Acids Res.* **2009**, *37* (3), e24.

(56) Eckstein, F. *Nucleic Acid Ther.* **2014**, *24* (6), 374–387.

(57) Sinha, N. D.; Michaud, D. P.; Roy, S. K.; Casale, R. A. *Nucleic Acids Res.* **1994**, *22* (15), 3119–3123.

(58) Thiviyanathan, V.; Vyazovkina, K. V; Gozansky, E. K.; Bichenchova, E.; Abramova, T. V; Luxon,

B. A.; Lebedev, A. V; Gorenstein, D. G. *Biochemistry* **2002**, *41* (3), 827–838.

(59) Lesnik, E. A.; Guinosso, C. J.; Kawasaki, A. M.; Sasmor, H.; Zounes, M.; Cummins, L. L.; Ecker, D. J.; Cook, P. D.; Freier, S. M. *Biochemistry* **1993**, *32* (30), 7832–7838.

(60) Majlessi, M.; Nelson, N. C.; Becker, M. M. *Nucleic Acids Res.* **1998**, *26* (9), 2224–2229.

(61) Kawasaki, a M.; Casper, M. D.; Freier, S. M.; Lesnik, E. a; Zounes, M. C.; Cummins, L. L.; Gonzalez, C.; Cook, P. D. *J. Med. Chem.* **1993**, *36* (7), 831–841.

(62) Steffens, R.; Leumann, C. J. *J. Am. Chem. Soc.* **1997**, *119* (47), 11548–11549.

(63) Renneberg, D.; Bouliong, E.; Reber, U.; Schümperli, D.; Leumann, C. J.; Dorte Renneberg, Emilie Bouliong, Ulrich Reber, D. S. and C. J. L. *Nucleic Acids Res.* **2002**, *30* (13), 2751–2757.

(64) Schöning, K.-U.; Scholz, P.; Wu, X.; Guntha, S.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. *Helv. Chim. Acta* **2002**, *85* (12), 4111–4153.

(65) Herdewijn, P. *Angew. Chem. Int. Ed. Engl.* **2001**, *40* (12), 2249–2251.

(66) Yu, H.; Zhang, S.; Dunn, M. R.; Chaput, J. C. *J. Am. Chem. Soc.* **2013**, *135* (9), 3583–3591.

(67) Singh, S. K.; Koshkin, A. A.; Wengel, J.; Nielsen, P. *Chem. Commun.* **1998**, No. 4, 455–456.

(68) Petersen, M.; Wengel, J. *Trends Biotechnol.* **2003**, *21* (2), 74–81.

(69) Heasman, J. *Dev. Biol.* **2002**, *243* (2), 209–214.

(70) Ekker, S. C.; Larson, J. D. *Genesis* **2001**, *30* (3), 89–93.

(71) Summerton, J. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1999**, *1489* (1), 141–158.

(72) Varizhuk, A. M.; Kaluzhny, D. N.; Novikov, R. A.; Chizhov, A. O.; Smirnov, I. P.; Chuvilin, A. N.; Tatarinova, O. N.; Fisunov, G. Y.; Pozmogova, G. E.; Florentiev, V. L. *J. Org. Chem.* **2013**, *78* (12), 5964–5969.

(73) Isobe, H.; Fujino, T. *Chem. Rec.* **2014**, *14* (1), 41–51.

(74) He, L.; Hannon, G. J.; O'Connell, R. M.; Rao, D. S.; Chaudhuri, A. a.; Baltimore, D. *Nat. Rev. Immunol.* **2010**, *10* (2), 111–122.

(75) Mencía, A.; Modamio-Høybjør, S.; Redshaw, N.; Morín, M.; Mayo-Merino, F.; Olavarrieta, L.; Aguirre, L. A.; del Castillo, I.; Steel, K. P.; Dalmay, T.; Moreno, F.; Moreno-Pelayo, M. A. *Nat. Genet.* **2009**, *41* (5), 609–613.

(76) Tan, K. S.; Armugam, A.; Sepramaniam, S.; Lim, K. Y.; Setyowati, K. D.; Wang, C. W.; Jeyaseelan, K. *PLoS One* **2009**, *4* (11), e7689.

(77) van Rooij, E.; Sutherland, L. B.; Liu, N.; Williams, A. H.; McAnally, J.; Gerard, R. D.; Richardson, J. A.; Olson, E. N. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (48), 18255–18260.

(78) Zhao, Y.; Ransom, J. F.; Li, A.; Vedantham, V.; von Drehle, M.; Muth, A. N.; Tsuchihashi, T.;

McManus, M. T.; Schwartz, R. J.; Srivastava, D. *Cell* **2007**, *129* (2), 303–317.

(79) Asangani, I. A.; Rasheed, S. A. K.; Nikolova, D. A.; Leupold, J. H.; Colburn, N. H.; Post, S.; Allgayer, H. *Oncogene* **2008**, *27* (15), 2128–2136.

(80) Murakami, Y.; Yasuda, T.; Saigo, K.; Urashima, T.; Toyoda, H.; Okanoue, T.; Shimotohno, K. *Oncogene* **2006**, *25* (17), 2537–2545.

(81) Song, S. J.; Pandolfi, P. P. *Curr. Opin. Hematol.* **2014**, *21* (4), 276–282.

(82) Calin, G. A.; Liu, C.-G.; Sevignani, C.; Ferracin, M.; Felli, N.; Dumitru, C. D.; Shimizu, M.; Cimmino, A.; Zupo, S.; Dono, M.; Dell'Aquila, M. L.; Alder, H.; Rassenti, L.; Kipps, T. J.; Bullrich, F.; Negrini, M.; Croce, C. M. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (32), 11755–11760.

(83) Calin, G. A.; Dumitru, C. D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; Rassenti, L.; Kipps, T.; Negrini, M.; Bullrich, F.; Croce, C. M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (24), 15524–15529.

(84) Eulalio, A.; Mano, M.; Ferro, M. D.; Zentilin, L.; Sinagra, G.; Zacchigna, S.; Giacca, M. *Nature* **2012**, *492* (7429), 376–381.

(85) Krützfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K. G.; Tuschl, T.; Manoharan, M.; Stoffel, M. *Nature* **2005**, *438* (7068), 685–689.

(86) Grimes, C. S. V. and H. L. **2012**, *928* (3), 53–65.

(87) Fontana, L.; Fiori, M. E.; Albini, S.; Cifaldi, L.; Giovinazzi, S.; Forloni, M.; Boldrini, R.; Donfrancesco, A.; Federici, V.; Giacomini, P.; Peschle, C.; Fruci, D. *PLoS One* **2008**, *3* (5), e2236.

(88) Brock, M.; Samillan, V. J.; Trenkmann, M.; Schwarzwald, C.; Ulrich, S.; Gay, R. E.; Gassmann, M.; Ostergaard, L.; Gay, S.; Speich, R.; Huber, L. C. *Eur. Heart J.* **2014**, *35* (45), 3203–3211.

(89) Scherr, M.; Venturini, L.; Battmer, K.; Schaller-Schoenitz, M.; Schaefer, D.; Dallmann, I.; Ganser, A.; Eder, M. *Nucleic Acids Res.* **2007**, *35* (22), e149.

(90) Gebert, L. F. R.; Rebhan, M. A. E.; Crivelli, S. E. M.; Denzler, R.; Stoffel, M.; Hall, J. *Nucleic Acids Res.* **2013**, *42* (1), 609–621.

(91) Ottosen, S.; Parsley, T. B.; Yang, L.; Zeh, K.; van Doorn, L.-J.; van der Veer, E.; Raney, A. K.; Hodges, M. R.; Patick, A. K. *Antimicrob. Agents Chemother.* **2015**, *59* (1), 599–608.

(92) Nielsen, P.; Egholm, M.; Berg, R.; Buchardt, O. *Science* **1991**, *254* (5037), 1497–1500.

(93) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol.* **1996**, *3* (5), 410–413.

(94) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. **1995**, *270* (December).

(95) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365* (6446), 566–568.

(96) Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskil, M. D.; Egholm, M.; Buchard, O.; Sönnichsen, S. H.; Nlelsen, P. E. *Biochem. Pharmacol.* **1994**, *48* (6), 1310–1313.

(97) Nielsen, P. E. *Peptide Nucleic Acids: Protocols and Applications*; Garland Science, 2004.

(98) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chemie Int. Ed.* **1998**, *37* (20), 2796–2823.

(99) Bertucci, A.; Manicardi, A.; Candiani, A.; Giannetti, S.; Cucinotta, A.; Spoto, G.; Konstantaki, M.; Pissadakis, S.; Selleri, S.; Corradini, R. *Biosens. Bioelectron.* **2015**, *63*, 248–254.

(100) *Detection of Non-Amplified Genomic DNA*; Spoto, G., Corradini, R., Eds.; Soft and Biological Matter; Springer Netherlands: Dordrecht, 2012.

(101) Sforza, S.; Corradini, R.; Tedeschi, T.; Marchelli, R. *Chem. Soc. Rev.* **2011**, *40* (1), 221–232.

(102) Amirkhanov, N. V; Zhang, K.; Aruva, M. R.; Thakur, M. L.; Wickstrom, E. *Bioconjug. Chem.* **2010**, *21* (4), 731–740.

(103) Malic, S.; Hill, K. E.; Hayes, A.; Percival, S. L.; Thomas, D. W.; Williams, D. W. *Microbiology* **2009**, *155* (Pt 8), 2603–2611.

(104) Cheng, C. J.; Bahal, R.; Babar, I. A.; Pincus, Z.; Barrera, F.; Liu, C.; Svoronos, A.; Braddock, D. T.; Glazer, P. M.; Engelman, D. M.; Saltzman, W. M.; Slack, F. J. *Nature* **2014**, *518* (7537), 107–110.

(105) Tonelli, R.; McIntyre, A.; Camerin, C.; Walters, Z. S.; Di Leo, K.; Selfe, J.; Purgato, S.; Missiaglia, E.; Tortori, A.; Renshaw, J.; Astolfi, A.; Taylor, K. R.; Serravalle, S.; Bishop, R.; Nanni, C.; Valentijn, L. J.; Faccini, A.; Leuschner, I.; Formica, S.; Reis-Filho, J. S.; Ambrosini, V.; Thway, K.; Franzoni, M.; Summersgill, B.; Marchelli, R.; Hrelia, P.; Cantelli-Forti, G.; Fanti, S.; Corradini, R.; Pession, A.; Shipley, J. *Clin. Cancer Res.* **2012**, *18* (3), 796–807.

(106) Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *European J. Org. Chem.* **1999**, 197–204.

(107) Crawford, M. J.; Rapireddy, S.; Bahal, R.; Sacui, I.; Ly, D. H. *J. Nucleic Acids* **2011**, *2011*, ID 652702.

(108) Wojciechowski, F.; Hudson, R. H. E. *Curr. Top. Med. Chem.* **2007**, *7* (7), 667–679.

(109) Hudson, R. H. E.; Viirre, R. D.; Liu, Y. H.; Wojciechowski, F.; Dambenieks, a K. *Society* **2004**, *76*, 1591–1598.

(110) Corradini, R.; Sforza, S.; Tedeschi, T.; Totsingan, F.; Manicardi, A.; Marchelli, R. *Curr. Top. Med. Chem.* **2011**, *11* (12), 1535–1554.

(111) Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli, R. *European J. Org. Chem.* **2007**, *2007* (35), 5879–5885.

(112) Tedeschi, T.; Sforza, S.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2005**, *17 Suppl*, S196–S204.

(113) Sforza, S.; Corradini, R.; Ghirardi, S.; Dossena, A.; Marchelli, R. *European J. Org. Chem.* **2000**, *2000* (16), 2905–2913.

(114) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128* (31), 10258–10267.

(115) Dragulescu-Andrasi, A.; Zhou, P.; He, G.; Ly, D. H. *Chem. Commun.* **2005**, No. 2, 244–246.

(116) Kumar, P.; Jain, D. R. *Tetrahedron* **2015**, *71* (21), 3378–3384.

(117) Lebleu, B.; Moulton, H. M.; Abes, R.; Ivanova, G. D.; Abes, S.; Stein, D. A.; Iversen, P. L.; Arzumanov, A. A.; Gait, M. J. *Adv. Drug Deliv. Rev.* **2008**, *60* (4-5), 517–529.

(118) Manicardi, A.; Fabbri, E.; Tedeschi, T.; Sforza, S.; Bianchi, N.; Brognara, E.; Gambari, R.; Marchelli, R.; Corradini, R. *Chembiochem* **2012**, *13* (9), 1327–1337.

(119) Boffa, L. C.; Cutrona, G.; Cilli, M.; Matis, S.; Damonte, G.; Mariani, M. R.; Millo, E.; Moroni, M.; Roncella, S.; Fedeli, F.; Ferrarini, M. *Cancer Gene Ther.* **2006**, *14* (2), 220–226.

(120) Eldrup, A. B.; Nielsen, B. B.; Haaima, G.; Rasmussen, H.; Kastrup, J. S.; Christensen, C.; Nielsen, P. E. *European J. Org. Chem.* **2001**, *2001* (9), 1781–1790.

(121) Eldrup, A. B.; Christensen, C.; Haaima, G.; Nielsen, P. E. *J. Am. Chem. Soc.* **2002**, *124* (13), 3254–3262.

(122) Bethge, L.; Jarikote, D. V.; Seitz, O. *Bioorg. Med. Chem.* **2008**, *16* (1), 114–125.

(123) Manicardi, A.; Guidi, L.; Ghidini, A.; Corradini, R. *Beilstein J. Org. Chem.* **2014**, *10*, 1495–1503.

(124) Flanagan, W. M.; Wolf, J. J.; Olson, P.; Grant, D.; Lin, K.-Y. Y.; Wagner, R. W.; Matteucci, M. D. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (7), 3513–3518.

(125) Ausín, C.; Ortega, J. A.; Robles, J.; Grandas, A.; Pedroso, E. *Org. Lett.* **2002**, *4* (c), 4073–4075.

(126) Wilds, C. J.; Maier, M. A.; Tereshko, V.; Manoharan, M.; Egli, M. *Angew. Chem. Int. Ed. Engl.* **2002**, *41* (1), 115–117.

(127) Wojciechowski, F.; Hudson, R. H. E. *Org. Lett.* **2009**, *11* (21), 4878–4881.

(128) Chenna, V.; Rapireddy, S.; Sahu, B.; Ausin, C.; Pedroso, E.; Ly, D. H. *ChemBioChem* **2008**, *9* (15), 2388–2391.

(129) Lee, S. E.; Sidorov, a; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. a; Dickman, M. J.; Hornby, D. P.; Grasby, J. a; Williams, D. M. *Nucleic Acids Res.* **2001**, *29* (7), 1565–1573.

(130) Giller, G. *Nucleic Acids Res.* **2003**, *31* (10), 2630–2635.

(131) Luyten, I.; Herdewijn, P. *Eur. J. Med. Chem.* **1998**, *33* (7-8), 515–576.

(132) Telser, J.; Cruickshank, K. A.; Morrison, L. E.; Netzel, T. L. *J. Am. Chem. Soc.* **2002**, *111* (18), 6966–6976.

(133) Hudson, R. H.; Liu, Y.; Wojciechowski, F. *Can. J. Chem.* **2007**, *85* (4), 302–312.

(134) Meyer, R. B.; Tabone, J. C.; Hurst, G. D.; Smith, T. M.; Gamper, H.; Meyer Jr., R. B.; Tabone, J. C.; Hurst, G. D.; Smith, T. M.; Gamper, H. *J. Am. Chem. Soc.* **1989**, *111* (22), 8517–8519.

(135) Hudson, R. H. .; Li, G.; Tse, J. *Tetrahedron Lett.* **2002**, *43* (8), 1381–1386.

(136) Manicardi, A.; Accetta, A.; Tedeschi, T.; Sforza, S.; Marchelli, R.; Corradini, R. *Artif. DNA PNA XNA* **2012**, *3* (2), 53–62.

(137) Menchise, V.; De Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (21), 12021–12026.

(138) He, W.; Crawford, M. J.; Rapireddy, S.; Madrid, M.; Gil, R. R.; Ly, D. H.; Achim, C. *Mol. Biosyst.* **2010**, *6* (9), 1619–1629.

(139) Petersson, B.; Nielsen, B. B.; Rasmussen, H.; Larsen, I. K.; Gajhede, M.; Nielsen, P. E.; Kastrup, J. S. *J.Am.Chem.Soc.* **2005**, *127*, 1424–1430.

(140) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270* (5243), 1838–1841.

(141) Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat.Struct.Biol.* **1997**, *4*, 98–101.

(142) Yeh, J. I.; Pohl, E.; Truan, D.; He, W.; Sheldrick, G. M.; Du, S.; Achim, C. *Chemistry (Easton).* **2010**, *16*, 11867–11875.

(143) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777–780.

(144) Hyrup, B.; Egholm, M.; Nielsen, P. E.; Wittung, P.; Norden, B.; Buchardt, O. *J. Am. Chem. Soc.* **1994**, *116* (18), 7964–7970.

(145) Govindaraju, T.; Madhuri, V.; Kumar, V. A.; Ganesh, K. N. *J. Org. Chem.* **2006**, *71* (1), 14–21.

(146) Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38* (5), 404–412.

(147) Oh, S. Y.; Ju, Y.; Park, H. *Mol. Cells* **2009**, *28* (4), 341–345.

(148) Fabani, M. M.; Abreu-Goodger, C.; Williams, D.; Lyons, P. A.; Torres, A. G.; Smith, K. G. C.; Enright, A. J.; Gait, M. J.; Vigorito, E. *Nucleic Acids Res.* **2010**, *38* (13), 4466–4475.

(149) Brognara, E.; Fabbri, E.; Bazzoli, E.; Montagner, G.; Ghimenton, C.; Eccher, A.; Cantù, C.; Manicardi, A.; Bianchi, N.; Finotti, A.; Breveglieri, G.; Borgatti, M.; Corradini, R.; Bezzerri, V.; Cabrini, G.; Gambari, R. *J. Neurooncol.* **2014**, *118* (1), 19–28.

(150) Ryoo, S.-R.; Lee, J.; Yeo, J.; Na, H.-K.; Kim, Y.-K.; Jang, H.; Lee, J. H.; Han, S. W.; Lee, Y.; Kim, V.

N.; Min, D.-H. *ACS Nano* **2013**, *7* (7), 5882–5891.

(151) Jackson, A. L.; Burchard, J.; Schelter, J.; Chau, B. N.; Cleary, M.; Lim, L.; Linsley, P. S. *RNA* **2006**, *12* (7), 1179–1187.

(152) Murtola, M.; Wenska, M.; Strömberg, R.; Stro, R. *J. Am. Chem. Soc.* **2010**, *1* (V), 8984–8990.

(153) Murtola, M.; Str, R.; Strömberg, R. *Org. Biomol. Chem.* **2008**, *6* (20), 3837–3842.

(154) Schroeder, G. K.; Lad, C.; Wyman, P.; Williams, N. H.; Wolfenden, R. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (11), 4052–4055.

(155) Thompson, J. E.; Kutateladze, T. G.; Schuster, M. C.; Venegas, F. D.; Messmore, J. M.; Raines, R. T. *Bioorg. Chem.* **1995**, *23* (4), 471–481.

(156) Horvath, M.; Choi, J.; Kim, Y.; Wilkosz, P.; Rosenberg, J. M. The Crystal Structure Of A Post-Reactive Cognate Dna-Eco Ri Complex At 2.50 A In The Presence Of Mn2+ Ion http://www.rcsb.org/pdb/explore.do?structureId=1qps.

(157) Davies, J. F.; Hostomska, Z.; Hostomsky, Z.; Matthews, D. A.; Jordan, S. R.; Matthews, D. A. *Science* **1991**, *252* (5002), 88–95.

(158) Kim, E. E.; Wyckoff, H. W. *J. Mol. Biol.* **1991**, *218* (2), 449–464.

(159) Cox, R. S.; Schenk, G.; Mitić, N.; Gahan, L. R.; Hengge, A. C. *J. Am. Chem. Soc.* **2007**, *129* (31), 9550–9551.

(160) Pitié, M.; Burrows, C. J.; Meunier, B. *Nucleic Acids Res.* **2000**, *28* (24), 4856–4864.

(161) Sreedhara, A.; Cowan, J. a. *J. Biol. Inorg. Chem.* **2001**, *6* (4), 337–347.

(162) Shahabadi, N.; Kashanian, S.; Darabi, F. *Eur. J. Med. Chem.* **2010**, *45* (9), 4239–4245.

(163) Kong, D.-M.; Wang, J.; Zhu, L.-N.; Jin, Y.-W.; Li, X.-Z.; Shen, H.-X.; Mi, H.-F. *J. Inorg. Biochem.* **2008**, *102* (4), 824–832.

(164) Yamamoto, Y.; Igawa, T.; Sumaoka, J.; Komiyama, M. *Nucleic Acids Symp. Ser.* **2000**, *44* (1), 231–232.

(165) Kitamura, Y.; Komiyama, M. *Nucleic Acids Res.* **2002**, *30* (19), 102 – 107.

(166) Aiba, Y.; Hamano, Y.; Kameshima, W.; Araki, Y.; Wada, T.; Accetta, A.; Sforza, S.; Corradini, R.; Marchelli, R.; Komiyama, M. *Org. Biomol. Chem.* **2013**, *11* (32), 5233–5238.

(167) Hertzberg, R. P.; Dervan, P. B. *Biochemistry* **1984**, *23* (17), 3934–3945.

(168) Desbouis, D.; Troitsky, I. P.; Belousoff, M. J.; Spiccia, L.; Graham, B. *Coord. Chem. Rev.* **2012**, *256* (11-12), 897–937.

(169) Linkletter, B.; Chin, J. *Angew. Chemie Int. Ed. English* **1995**, *34* (4), 472–474.

(170) Koike, T.; Kimura, E. *J. Am. Chem. Soc.* **1991**, *113* (23), 8935–8941.

(171) Burstyn, J. N.; Deal, K. A. *Inorg. Chem.* **1993**, *32* (17), 3585–3586.

(172) Young, M. J.; Chin, J. *J. Am. Chem. Soc.* **1995**, *117* (42), 10577–10578.

(173) Subat, M.; Woinaroschy, K.; Gerstl, C.; Sarkar, B.; Kaim, W.; König, B. *Inorg. Chem.* **2008**, *47* (11), 4661–4668.

(174) Kady, I. O.; Tan, B.; Ho, Z.; Scarborough, T. *J. Chem. Soc. Chem. Commun.* **1995**, No. 11, 1137.

(175) Kimura, E.; Kodama, Y.; Koike, T.; Shiro, M. *J. Am. Chem. Soc.* **1995**, *117* (32), 8304–8311.

(176) Rossi, P.; Felluga, F.; Tecilla, P.; Formaggio, F.; Crisma, M.; Toniolo, C.; Scrimin, P. *Biopolymers* **2000**, *55* (6), 496–501.

(177) Bonomi, R.; Selvestrel, F.; Lombardo, V.; Sissi, C.; Polizzi, S.; Mancin, F.; Tonellato, U.; Scrimin, P. *J. Am. Chem. Soc.* **2008**, *130* (47), 15744–15745.

(178) Mancin, F.; Scrimin, P.; Tecilla, P.; Tonellato, U. *Chem. Commun.* **2005**, No. 20, 2540–2548.

(179) Gavroglou, K.; Simões, A. *Neither Physics Nor Chemistry: A History of Quantum Chemistry*; MIT Press, 2012.

(180) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, *267* (5612), 585–590.

(181) Dokholyan, N. V. *Computational Modeling of Biological Systems: From Molecules to Pathways*; Nikolay V Dokholyan, Ed.; Biological and Medical Physics, Biomedical Engineering; Springer New York, **2012**.

(182) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science Publ; Clarendon Press, **1989**.

(183) de Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1980**, *373* (1752), 27–56.

(184) Boyce, S.; Mobley, D.; Rocklin, G.; Graves, A.; Dill, K.; Shoichet, B. K. *J.Mol.Biol.* **2009**, *394*, 747–763.

(185) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91* , 43–56.

(186) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688.

(187) Brooks, B. R. *et al J. Comput. Chem.* **2009**, *30* (10), 1545–1614.

(188) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

(189) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25* (13), 1656–1676.

(190) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65* (3), 712–725.

(191) MacKerell, A. D. *et al J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.

(192) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1997**, *18* (16), 1955–1970.

(193) Freddolino, P. L.; Park, S.; Roux, B.; Schulten, K. *Biophys. J.* **2009**, *96* (9), 3772–3780.

(194) Ponder, J. W.; Case, D. A. *Protein Simulations*; Advances in Protein Chemistry; Elsevier, 2003; Vol. 66.

(195) Palazzesi, F.; Prakash, M. K.; Bonomi, M.; Barducci, A. *J. Chem. Theory Comput.* **2015**, *11* (1), 2–7.

(196) Dickson, C. J.; Madej, B. D.; Skjevik, A. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. *J. Chem. Theory Comput.* **2014**, *10* (2), 865–879.

(197) Raman, E. P.; Guvench, O.; MacKerell, A. D. *J. Phys. Chem. B* **2010**, *114* (40), 12981–12994.

(198) Ricci, C. G.; de Andrade, A. S. C.; Mottin, M.; Netz, P. A. *J. Phys. Chem. B* **2010**, *114* (30), 9882–9893.

(199) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78* (8), 1950–1958.

(200) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92* (11), 3817–3829.

(201) MacKerell, A. D.; Banavali, N.; Foloppe, N. *Biopolymers* **2000**, *56* (4), 257–265.

(202) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2015**, *11* (7), 3499–3509.

(203) Wickstrom, L.; Okur, A.; Simmerling, C. *Biophys. J.* **2009**, *97* (3), 853–856.

(204) Wolf, M. G.; Groenhof, G. *J. Comput. Chem.* **2012**, *33* (28), 2225–2232.

(205) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.

(206) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *J. Comput. Chem.* **2010**, *31* (4), 671–690.

(207) Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450* (7172), 964–972.

(208) Shastry, M. C. R.; Roder, H. *Nat. Struct. Biol.* **1998**, *5* (5), 385–392.

(209) Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E. *Curr. Opin. Struct. Biol.* **2009**, *19* (2), 120–127.

(210) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281* (1-3), 140–150.

(211) Helms, V.; Wade, R. C. *J. Am. Chem. Soc.* **1998**, *120* (12), 2710–2713.

(212) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314* (1-2), 141–151.

(213) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23* (2), 187–199.

(214) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (20), 12562–12566.

(215) Bussi, G.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2006**, *96* (9), 090601.

(216) Miller, F. P.; Vandome, A. F.; John, M. B. *Metadynamics*; VDM Publishing, 2010.

(217) Bernardi, R. C.; Melo, M. C. R.; Schulten, K. *BBA - Gen. Subj.* **2015**, *1850* (5), 872–877.

(218) Spiwok, V. V.; Sucur, Z.; Hosek, P.; Šućur, Z.; Hošek, P. *Biotechnol. Adv.* **2015**, *33* (6), 1130–1140.

(219) Bonomi, M.; Barducci, A.; Parrinello, M. *J. Comput. Chem.* **2009**, *30* (11), 1615–1621.

(220) Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100* (2), 1–4.

(221) Dama, J. F.; Parrinello, M.; Voth, G. A. *Phys. Rev. Lett.* **2014**, *112* (24), 240602.

(222) Ilott, A. J.; Palucha, S.; Hodgkinson, P.; Wilson, M. R. *J. Phys. Chem. B* **2013**, *117* (40), 12286–12295.

(223) Deighan, M.; Bonomi, M.; Pfaendtner, J. *J. Chem. Theory Comput.* **2012**, *8* (7), 2189–2192.

(224) Cui, H.; Lyman, E.; Voth, G. A. *Biophys. J.* **2011**, *100* (5), 1271–1279.

(225) Limongelli, V.; Bonomi, M.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (16), 6358–6363.

(226) Leone, V.; Marinelli, F.; Carloni, P.; Parrinello, M. *Curr. Opin. Struct. Biol.* **2010**, *20* (2), 148–154.

(227) Salvalaglio, M.; Vetter, T.; Giberti, F.; Mazzotti, M.; Parrinello, M. *J. Am. Chem. Soc.* **2012**, *134* (41), 17221–17233.

(228) Polino, D.; Parrinello, M. *J. Phys. Chem. A* **2015**, *119* (6), 978–989.

(229) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29* (7), 845–854.

(230) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. *Comput. Phys. Commun.* **2009**, *180* (10), 1961–1972.

(231) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185* (2), 604–613.

(232) Do, T. N.; Carloni, P.; Varani, G.; Bussi, G. *J. Chem. Theory Comput.* **2013**, *9* (3), 1720–1730.

(233) Haider, S.; Neidle, S. *Methods Mol. Biol.* **2010**, *608*, 17–37.

(234) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. *Phys. Chem. Chem. Phys.* **2010**, *12* (28), 7821–7839.

(235) REDDB Server http://upjv.q4md-forcefieldtools.org/REDDB/projects/F-93/.

(236) Thomas J. Macke; David A. Case. *Molecular Modeling of Nucleic Acids*; Leontis, N. B., SantaLucia, J., Eds.; ACS Symposium Series; American Chemical Society: Washington, DC, 1997; Vol. 682.

(237) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3* (2), 198–210.

(238) Sousa da Silva, A. W.; Vranken, W. F. *BMC Res. Notes* **2012**, *5* (1), 367.

(239) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81* (8), 3684.

(240) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126* (1), 014101.

(241) Yoneya, M.; Berendsen, H. J. C.; Hirasawa, K. *Mol. Simul.* **1994**, *13* (6), 395–405.

(242) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280.

(243) Hariharan, P. C.; Pople, J. A. *Chem. Phys. Lett.* **1972**, *16* (2), 217–219.

(244) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y. *Nucleic Acids Res.* **2011**, *39* (Web Server issue), W511–W517.

(245) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197.

(246) Frisch, M. J. *et al*, Gaussian 09 D1, Gaussian, Inc.: Wallingford CT **2009**.

(247) RESP ESP charge Derive: R.E.D. version III - R.E.D. III http://upjv.q4md-forcefieldtools.org/REDServer-Development/.

(248) Wang, F.; Becker, J. P.; Cieplack, P.; Dupradeau, F.-Y. R.E.D. Server Development - Performing calculations with the PyRED program: Application to charge derivation, force field library building and force field parameter generation http://upjv.q4md-forcefieldtools.org/Tutorial/Tutorial-4.php.

(249) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I. *Nucleic Acids: Structures, Properties, and Functions*; University Science Books, 2000.

(250) Autiero, I.; Saviano, M.; Langella, E. *Phys. Chem. Chem. Phys.* **2014**, *16* (5), 1868–1874.

(251) Soliva, R.; Sherer, E.; Luque, F. J.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **2000**, *122*

(25), 5997–6008.

(252) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120* (24), 5895–5904.

(253) Gantchev, T. G.; Girouard, S.; Dodd, D. W.; Wojciechowski, F.; Hudson, R. H. E.; Hunting, D. J. *Biochemistry* **2009**, *48* (29), 7032–7044.

(254) Hatcher, E.; Balaeff, A.; Keinan, S.; Venkatramani, R.; Beratan, D. N. *J. Am. Chem. Soc.* **2008**, *130* (35), 11752–11761.

(255) Mansawat, W.; Boonlua, C.; Siriwong, K.; Vilaivan, T. *Tetrahedron* **2012**, *68* (21), 3988–3995.

(256) Sanders, J. M.; Wampole, M. E.; Chen, C.-P.; Sethi, D.; Singh, A.; Dupradeau, F.-Y.; Wang, F.; Gray, B. D.; Thakur, M. L.; Wickstrom, E. *J. Phys. Chem. B* **2013**, *117* (39), 11584–11595.

(257) Autiero, I.; Saviano, M.; Langella, E. *Eur. J. Med. Chem.* **2015**, *91*, 109–117.

(258) Kiliszek, A.; Banaszak, K.; Dauter, Z.; Rypniewski, W. *Nucleic Acids Res.* **2015**, 1513–1519.

(259) Topham, C. M.; Smith, J. C. *J. Mol. Biol.* **1999**, *292* (5), 1017–1038.

(260) Giberti, F.; Salvalaglio, M.; Mazzotti, M.; Parrinello, M. *Chem. Eng. Sci.* **2015**, *121*, 51–59.

(261) Manna, A.; Rapireddy, S.; Sureshkumar, G.; Ly, D. H. *Tetrahedron* **2015**, *71* (21), 3507–3514.

(262) Yeh, J. I.; Shivachev, B.; Rapireddy, S.; Crawford, M. J.; Gil, R. R.; Du, S.; Madrid, M.; Ly, D. H. *J. Am. Chem. Soc.* **2010**, *132* (31), 10717–10727.

(263) Sacui, I.; Hsieh, W.-C.; Manna, A.; Sahu, B.; Ly, D. H. *J. Am. Chem. Soc.* **2015**, *137* (26), 8603–8610.

(264) Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26* (9), 1601–1620.

(265) Ratilainen, T.; Holmén, A.; Tuite, E.; Nielsen, P. E.; Nordén, B. *Biochemistry* **2000**, *39* (26), 7781–7791.

(266) Uhlmann, E.; Will, D. W.; Breipohl, G.; Langner, D.; Ryte, A. *Angew. Chemie* **1996**, *35* (22), 2632–2635.

(267) Sen, A.; Nielsen, P. E. *Biophys. Chem.* **2009**, *141* (1), 29–33.

(268) Igloi, G. L. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (15), 8562–8567.

(269) Dias, N.; Sénamaud-Beaufort, C.; Forestier, E. Le; Auvin, C.; Hélène, C.; Ester Saison-Behmoaras, T. *J. Mol. Biol.* **2002**, *320* (3), 489–501.

(270) Accetta, A. Molecular Engineering of PNA Using Modified Uracil Derivatives and Porphyrins, University of Parma, **2010**.

(271) Tedeschi, T.; Sforza, S.; Maffei, F.; Corradini, R.; Marchelli, R. *Tetrahedron Lett.* **2008**, *49* (33), 4958–4961.

(272) Theodorou, V.; Ragoussis, V.; Strongilos, A.; Zelepos, E.; Eleftheriou, A.; Dimitriou, M.

Tetrahedron Lett. **2005**, *46* (8), 1357–1360.

(273) Micklitz, W.; Lippert, B.; Schöllhorn, H.; Thewalt, U. *J. Heterocycl. Chem.* **1989**, *26* (5), 1499–1500.

(274) Staudinger, H.; Meyer, J. *Helv. Chim. Acta* **1919**, *2* (1), 619–635.

(275) Tian, W. Q.; Wang, Y. A. *J. Org. Chem.* **2004**, *69* (13), 4299–4308.

(276) Rossi, A.; Verona, M. D.; Corradini, R. Sintesi e studi funzionali di PNA contenenti una base modificata, Parma, **2015**.

(277) Corradini, R.; Feriotto, G.; Sforza, S.; Marchelli, R.; Gambari, R. *J. Mol. Recognit. 17* (1), 76–84.

(278) Feriotto, G.; Corradini, R.; Sforza, S.; Bianchi, N.; Mischiati, C.; Marchelli, R.; Gambari, R. *Lab. Investig.* **2001**, *81* (10), 1415–1427.

(279) Garofalo, M.; Di Leva, G.; Romano, G.; Nuovo, G.; Suh, S.-S.; Ngankeu, A.; Taccioli, C.; Pichiorri, F.; Alder, H.; Secchiero, P.; Gasparini, P.; Gonelli, A.; Costinean, S.; Acunzo, M.; Condorelli, G.; Croce, C. M. *Cancer Cell* **2009**, *16* (6), 498–509.

(280) Brognara, E.; Fabbri, E.; Aimi, F.; Manicardi, A.; Bianchi, N.; Finotti, A.; Breveglieri, G.; Borgatti, M.; Corradini, R.; Marchelli, R.; Gambari, R. *Int. J. Oncol.* **2012**, *41* (6), 2119–2127.

(281) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37* (2), 785–789.

(282) Bassetti, M.; Ciceri, S.; Lancia, F.; Pasquini, C. *Tetrahedron Lett.* **2014**, *55* (9), 1608–1612.

(283) Klahn, P.; Erhardt, H.; Kotthaus, A.; Kirsch, S. F. *Angew. Chem. Int. Ed. Engl.* **2014**, *53* (30), 7913–7917.

(284) Manicardi, A.; Accetta, A.; Tedeschi, T.; Sforza, S.; Marchelli, R.; Corradini, R. *Artif. DNA. PNA XNA 3* (2), 53–62.

(285) Chiacchio, U.; Corsaro, A.; Mates, J.; Merino, P.; Piperno, A.; Rescifina, A.; Romeo, G.; Romeo, R.; Tejero, T. *Tetrahedron* **2003**, *59* (26), 4733–4738.

(286) Schlosser, M.; Schaub, B. *Chimia (Aarau).* **1982**, *36* (10), 396–397.

(287) Pretsch, E.; Biemann, K.; Clerc, T.; Seibl, J.; Simon, W. *Tables of Spectral Data for Structure Determination of Organic Compounds*; Chemical Laboratory Practice; Springer Berlin Heidelberg, **2013**.

(288) Gray, J. *Queue* **2008**, *6* (3), 63–68.

(289) ULTRAPLACAD http://ultraplacad.eu/.

(290) Brutus Cluster home page https://www1.ethz.ch/id/services/list/comp_zentral/cluster/index_EN.

(291) Piz Dora http://www.cscs.ch/computers/piz_daint_piz_dora/index.html.

(292) Mönch http://www.cscs.ch/computers/moench/index.html.

(293) Frigoli, M.; Moustrou, C.; Samat, A.; Guglielmetti, R. *European J. Org. Chem.* **2003**, *2003* (15), 2799–2812.

(294) Yamaji, M.; Maeda, H.; Minamida, K.; Maeda, T.; Asai, K.; Konishi, G.-I.; Mizuno, K. *Res. Chem. Intermed.* **2012**, *39* (1), 321–345.

(295) Trawick, B. N.; Daniher, A. T.; Bashkin, J. K. *Chem. Rev.* **1998**, *98* (3), 939–960.

(296) Bashkin, J. K.; Frolova, E. I.; Sampath, U. *J. Am. Chem. Soc.* **1994**, *116* (13), 5981–5982.

(297) Matsumura, K.; Endo, M.; Komiyama, M. *J. Chem. Soc. Chem. Commun.* **1994**, No. 17, 2019.

(298) Danneberg, F.; Ghidini, A.; Dogandzhiyski, P.; Kalden, E.; Strömberg, R.; Göbel, M. W. *Beilstein J. Org. Chem.* **2015**, *11*, 493–498.

(299) Petersen, L.; de Koning, M. C.; van Kuik-Romeijn, P.; Weterings, J.; Pol, C. J.; Platenburg, G.; Overhand, M.; van der Marel, G. A.; van Boom, J. H. *Bioconjug. Chem.* **2004**, *15* (3), 576–582.

(300) Verheijen, J. C.; Deiman, B. A. L. M.; Yeheskiely, E.; van der Marel, G. A.; van Boom, J. H. *Angew. Chemie* **2000**, *112* (2), 377–380.

(301) Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Peracchi, A.; Reinhoudt, D. N.; Salvio, R.; Sartori, A.; Ungaro, R. *J. Am. Chem. Soc.* **2007**, *129* (41), 12512–12520.

(302) Rossi, P.; Felluga, F.; Tecilla, P.; Formaggio, F.; Crisma, M.; Toniolo, C.; Scrimin, P. *J. Am. Chem. Soc.* **1999**, *121* (29), 6948–6949.

(303) Cai, H.-Z.; Kaden, T. A. *Helv. Chim. Acta* **1993**, *76* (1), 557–562.

(304) Studer, M.; Kaden, T. A. *Helv. Chim. Acta* **1986**, *69* (8), 2081–2086.

(305) Qian, J.; Wang, L.-P.; Tian, J.-L.; Xie, C.-Z.; Yan, S.-P. *J. Coord. Chem.* **2012**, *65* (1), 122–130.

(306) Qian, J.; Yu, S.; Wang, W.; Wang, L.; Tian, J.; Yan, S. *Dalton Trans.* **2014**, *43* (6), 2646–2655.

(307) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminform.* **2012**, *4* (1), 17.

(308) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105* (8), 2999–3093.

(309) Hartmann, M.; Clark, T.; van Eldik, R. *J. Am. Chem. Soc.* **1997**, *119* (33), 7843–7850.

(310) Bock, C. W.; Katz, A. K.; Glusker, J. P. *J. Am. Chem. Soc.* **1995**, *117* (13), 3754–3765.

(311) Perinelli, M.; Tegoni, M. Stabilità in soluzione di complessi di Zn(II) con leganti amminici a potenziale attività fosfodiesterasica, University of Parma, **2013**.

(312) Cao, R.; Müller, P.; Lippard, S. J. *J. Am. Chem. Soc.* **2010**, *132* (49), 17366–17369.

(313) Leontyev, I. V; Stuchebrukhov, A. A. *J. Chem. Theory Comput.* **2010**, *6* (5), 1498–1508.

(314) Kohl, S. W.; Kuse, K.; Hummert, M.; Schumann, H.; Mügge, C.; Janek, K.; Weißhoff, H.

Zeitschrift für Naturforsch. B **2007**, *62* (3), 397–406.

(315) Barta, C. A.; Bayly, S. R.; Read, P. W.; Patrick, B. O.; Thompson, R. C.; Orvig, C. *Inorg. Chem.* **2008**, *47* (7), 2280–2293.

(316) Martin, A. E.; Ford, T. M.; Bulkowski, J. E. *J. Org. Chem.* **1982**, *47* (3), 412–415.

(317) Mambriani, C.; Manicardi, A.; Corradini, R.; Sintesi di un PNA anti-miR221 contenente una nucleobase modificata, University of Parma, **2015**.

(318) Cerrurale, L.; Manicardi, A.; Corradini, R.; Nuove strategie sintetiche per pna contenenti dimeri di uracile, University of Parma, **2013**.

8. Appendix

8.1. Standard PNA Force Field

Tripos1: APN

```
@<TRIPOS>MOLECULE
APN
   33    34     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'      3.621054   -1.450373   -0.068677 N     1 APN    -0.4030 ****
  2 H1'1     4.246165   -1.271793    0.682602 H     1 APN     0.2842 ****
  3 C2'      3.457784   -0.420083   -1.077294 CT    1 APN    -0.1241 ****
  4 H2'1     3.628216   -0.847867   -2.056986 H1    1 APN     0.1147 ****
  5 H2'2     4.227242    0.322257   -0.905941 H1    1 APN     0.1147 ****
  6 C3'      2.080030    0.240647   -1.080497 CT    1 APN    -0.0522 ****
  7 H3'1     1.323813   -0.470938   -1.363061 H1    1 APN     0.1033 ****
  8 H3'2     2.071373    1.016457   -1.836440 H1    1 APN     0.1033 ****
  9 N4'      1.741210    0.867720    0.196197 N     1 APN    -0.0941 ****
 10 C5'      2.298716    2.197163    0.430342 CT    1 APN    -0.1112 ****
 11 H5'1     2.391804    2.356515    1.493133 H1    1 APN     0.0678 ****
 12 H5'2     3.275899    2.250678   -0.024827 H1    1 APN     0.0678 ****
 13 C'       1.449604    3.289214   -0.217087 C     1 APN     0.6209 ****
 14 O1'      1.767998    3.770682   -1.273205 O     1 APN    -0.5807 ****
 15 C7'      0.857680    0.402551    1.099946 C     1 APN     0.4391 ****
 16 O7'      0.505502    1.058337    2.054036 O     1 APN    -0.5502 ****
 17 C8'      0.288243   -1.011489    0.977878 CT    1 APN    -0.1214 ****
 18 H8'1     0.638353   -1.548180    1.850536 H1    1 APN     0.1124 ****
 19 H8'2     0.593037   -1.557702    0.104279 H1    1 APN     0.1124 ****
 20 N9      -1.153273   -0.974097    1.005139 N*    1 APN    -0.0578 ****
 21 C8      -1.981619   -0.898844    2.098267 CK    1 APN     0.1298 ****
 22 H8      -1.577259   -0.897330    3.089174 H5    1 APN     0.1677 ****
 23 N7      -3.225847   -0.831369    1.805401 NB    1 APN    -0.5582 ****
 24 C6      -4.278736   -0.782229   -0.521089 CA    1 APN     0.6597 ****
 25 N6      -5.576223   -0.734082   -0.171904 N2    1 APN    -0.8883 ****
 26 H61     -5.821383   -0.567634    0.776793 H     1 APN     0.4048 ****
 27 H62     -6.246165   -0.541126   -0.880495 H     1 APN     0.4048 ****
 28 C5      -3.247873   -0.843054    0.423001 CB    1 APN     0.0300 ****
 29 C4      -1.971958   -0.920512   -0.084086 CB    1 APN     0.4343 ****
 30 N3      -1.628350   -0.929692   -1.366627 NC    1 APN    -0.7127 ****
 31 C2      -2.681040   -0.862225   -2.150429 CQ    1 APN     0.5660 ****
 32 H2      -2.485791   -0.862420   -3.207408 H5    1 APN     0.0560 ****
 33 N1      -3.961295   -0.789586   -1.810556 NC    1 APN    -0.7398 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
```

```
   10     9    15 1
   11    10    11 1
   12    10    12 1
   13    10    13 1
   14    13    14 1
   15    15    16 1
   16    15    17 1
   17    17    18 1
   18    17    19 1
   19    17    20 1
   20    20    21 1
   21    20    29 1
   22    21    22 1
   23    21    23 1
   24    23    28 1
   25    24    25 1
   26    24    28 1
   27    24    33 1
   28    25    26 1
   29    25    27 1
   30    28    29 1
   31    29    30 1
   32    30    31 1
   33    31    32 1
   34    31    33 1
@<TRIPOS>SUBSTRUCTURE
      1  APN              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

Tripos2: NAP

```
@<TRIPOS>MOLECULE
NAP
   35    36     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'     0.000000    0.000000    0.000000 N3    1 NAP     0.0049 ****
  2 H1'1   -0.372154   -0.940539    0.000008 H     1 NAP     0.2531 ****
  3 H1'2   -0.372155    0.470271    0.814530 H     1 NAP     0.2531 ****
  4 H1'3   -0.372159    0.470265   -0.814535 H     1 NAP     0.2531 ****
  5 C2'     1.451021    0.000000    0.000000 CT    1 NAP     0.0723 ****
  6 H2'1    1.809090    0.728963   -0.715711 HP    1 NAP     0.0819 ****
  7 H2'2    1.772427   -0.976835   -0.339099 HP    1 NAP     0.0819 ****
  8 C3'     2.077421    0.318830    1.356743 CT    1 NAP    -0.0606 ****
  9 H3'1    1.853667    1.331448    1.644142 H1    1 NAP     0.0894 ****
 10 H3'2    3.154718    0.255838    1.262557 H1    1 NAP     0.0894 ****
 11 N4'     1.673355   -0.612895    2.408686 N     1 NAP    -0.0615 ****
 12 C5'     2.391832   -1.883556    2.456313 CT    1 NAP    -0.1251 ****
 13 H5'1    1.755750   -2.629775    2.905913 H1    1 NAP     0.0674 ****
 14 H5'2    2.636268   -2.183962    1.448873 H1    1 NAP     0.0674 ****
 15 C'      3.712813   -1.758991    3.212768 C     1 NAP     0.6140 ****
 16 O1'     4.752967   -1.641945    2.618671 O     1 NAP    -0.5754 ****
 17 C7'     0.814276   -0.353206    3.412835 C     1 NAP     0.4500 ****
```

```
 18 O7'     0.656345   -1.115185    4.339559 O     1 NAP    -0.5598 ****
 19 C8'    -0.040831    0.914581    3.399618 CT    1 NAP    -0.1055 ****
 20 H8'1   -1.067893    0.578481    3.333829 H1    1 NAP     0.1129 ****
 21 H8'2    0.144283    1.587359    2.582595 H1    1 NAP     0.1129 ****
 22 N9      0.128970    1.645092    4.631539 N*    1 NAP    -0.0808 ****
 23 C8     -0.484234    1.426978    5.841149 CK    1 NAP     0.1431 ****
 24 H8     -1.217445    0.654840    5.948895 H5    1 NAP     0.1651 ****
 25 N7     -0.092748    2.212538    6.772847 NB    1 NAP    -0.5626 ****
 26 C6      1.677778    4.047283    6.622865 CA    1 NAP     0.6559 ****
 27 N6      1.615237    4.513686    7.882340 N2    1 NAP    -0.8852 ****
 28 H61     1.101561    4.007469    8.566237 H     1 NAP     0.4041 ****
 29 H62     2.320174    5.146225    8.184189 H     1 NAP     0.4041 ****
 30 C5      0.862351    3.006437    6.165260 CB    1 NAP     0.0362 ****
 31 C4      1.016266    2.658603    4.843739 CB    1 NAP     0.4357 ****
 32 N3      1.862590    3.199779    3.975449 NC    1 NAP    -0.7175 ****
 33 C2      2.573771    4.158240    4.525565 CQ    1 NAP     0.5728 ****
 34 H2      3.286378    4.646507    3.885945 H5    1 NAP     0.0549 ****
 35 N1      2.533154    4.605154    5.773949 NC    1 NAP    -0.7416 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5     8 1
    8     8     9 1
    9     8    10 1
   10     8    11 1
   11    11    12 1
   12    11    17 1
   13    12    13 1
   14    12    14 1
   15    12    15 1
   16    15    16 1
   17    17    18 1
   18    17    19 1
   19    19    20 1
   20    19    21 1
   21    19    22 1
   22    22    23 1
   23    22    31 1
   24    23    24 1
   25    23    25 1
   26    25    30 1
   27    26    27 1
   28    26    30 1
   29    26    35 1
   30    27    28 1
   31    27    29 1
   32    30    31 1
   33    31    32 1
   34    32    33 1
   35    33    34 1
   36    33    35 1
@<TRIPOS>SUBSTRUCTURE
      1 NAP              1 ****              0 ****  ****
@<TRIPOS>HEADTAIL
0 0
C' 1
@<TRIPOS>RESIDUECONNECT
1 0 C' 0 0 0 0
```

Tripos3: CAP

```
@<TRIPOS>MOLECULE
CAP
   34    35     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 C'      0.000000    0.000000    0.000000 C     1 CAP     0.7881 ****
  2 O1'    -0.511689   -1.124427    0.000000 O2    1 CAP    -0.8206 ****
  3 OXT    -0.540795    1.108412    0.000000 O2    1 CAP    -0.8206 ****
  4 N1'     4.658960    2.043595    1.148355 N     1 CAP    -0.4302 ****
  5 H1'1    5.197266    1.232456    0.949976 H     1 CAP     0.2931 ****
  6 C2'     3.403976    1.870867    1.855917 CT    1 CAP    -0.0573 ****
  7 H2'1    3.389308    2.519417    2.722526 H1    1 CAP     0.0960 ****
  8 H2'2    3.373611    0.848591    2.211622 H1    1 CAP     0.0960 ****
  9 C3'     2.164235    2.172976    1.015351 CT    1 CAP    -0.0479 ****
 10 H3'1    2.132830    3.216818    0.755607 H1    1 CAP     0.0959 ****
 11 H3'2    1.284270    1.978060    1.616247 H1    1 CAP     0.0959 ****
 12 N4'     2.068693    1.343761   -0.185166 N     1 CAP    -0.0212 ****
 13 C5'     1.527328    0.000000    0.000000 CT    1 CAP    -0.4798 ****
 14 H5'1    1.915655   -0.644423   -0.772983 H1    1 CAP     0.1621 ****
 15 H5'2    1.839381   -0.371939    0.963957 H1    1 CAP     0.1621 ****
 16 C7'     2.293194    1.748333   -1.449955 C     1 CAP     0.4669 ****
 17 O7'     2.032945    1.052610   -2.405262 O     1 CAP    -0.5706 ****
 18 C8'     2.935922    3.107040   -1.731879 CT    1 CAP    -0.1309 ****
 19 H8'1    3.884219    2.892995   -2.208391 H1    1 CAP     0.1143 ****
 20 H8'2    3.125602    3.715935   -0.867188 H1    1 CAP     0.1143 ****
 21 N9      2.119337    3.866285   -2.646673 N*    1 CAP    -0.0583 ****
 22 C8      2.068388    3.776811   -4.016398 CK    1 CAP     0.1330 ****
 23 H8      2.712150    3.103014   -4.542682 H5    1 CAP     0.1697 ****
 24 N7      1.204275    4.550897   -4.557254 NB    1 CAP    -0.5627 ****
 25 C6     -0.402401    6.153321   -3.384917 CA    1 CAP     0.6591 ****
 26 N6     -1.010141    6.679483   -4.462658 N2    1 CAP    -0.8878 ****
 27 H61    -0.863299    6.268980   -5.355831 H     1 CAP     0.4046 ****
 28 H62    -1.820927    7.236423   -4.319130 H     1 CAP     0.4046 ****
 29 C5      0.614390    5.197601   -3.486992 CB    1 CAP     0.0379 ****
 30 C4      1.164160    4.775405   -2.299305 CB    1 CAP     0.4242 ****
 31 N3      0.818087    5.172773   -1.080538 NC    1 CAP    -0.7142 ****
 32 C2     -0.147642    6.063306   -1.117589 CQ    1 CAP     0.5704 ****
 33 H2     -0.487004    6.434296   -0.167575 H5    1 CAP     0.0554 ****
 34 N1     -0.767261    6.567034   -2.176915 NC    1 CAP    -0.7415 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1    13 1
    4     4     5 1
    5     4     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    16 1
   14    13    14 1
```

```
   15    13    15 1
   16    16    17 1
   17    16    18 1
   18    18    19 1
   19    18    20 1
   20    18    21 1
   21    21    22 1
   22    21    30 1
   23    22    23 1
   24    22    24 1
   25    24    29 1
   26    25    26 1
   27    25    29 1
   28    25    34 1
   29    26    27 1
   30    26    28 1
   31    29    30 1
   32    30    31 1
   33    31    32 1
   34    32    33 1
   35    32    34 1
@<TRIPOS>SUBSTRUCTURE
      1 CAP              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

Tripos4: MAPN

```
@<TRIPOS>MOLECULE
MAPN
   45    46     3     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 CH3      3.669350   -3.740306    0.749078 CT    1 ACE    -0.1986 ****
  2 HH31     2.948500   -4.543905    0.795444 HC    1 ACE     0.0639 ****
  3 HH32     3.800485   -3.308336    1.734877 HC    1 ACE     0.0639 ****
  4 HH33     4.620396   -4.154872    0.429549 HC    1 ACE     0.0639 ****
  5 C        3.208694   -2.725793   -0.275020 C     1 ACE     0.6513 ****
  6 O        2.527394   -3.029687   -1.215905 O     1 ACE    -0.5590 ****
  7 N1'      3.621054   -1.450373   -0.068677 N     2 APN    -0.5256 ****
  8 H1'1     4.246165   -1.271793    0.682602 H     2 APN     0.3154 ****
  9 C2'      3.457784   -0.420082   -1.077294 CT    2 APN    -0.1667 ****
 10 H2'1     3.628216   -0.847867   -2.056986 H1    2 APN     0.1407 ****
 11 H2'2     4.227242    0.322257   -0.905941 H1    2 APN     0.1407 ****
 12 C3'      2.080030    0.240647   -1.080497 CT    2 APN    -0.0601 ****
 13 H3'1     1.323813   -0.470938   -1.363061 H1    2 APN     0.0992 ****
 14 H3'2     2.071373    1.016457   -1.836440 H1    2 APN     0.0992 ****
 15 N4'      1.741210    0.867720    0.196197 N     2 APN    -0.0789 ****
 16 C5'      2.298716    2.197163    0.430342 CT    2 APN    -0.0497 ****
 17 H5'1     2.391804    2.356515    1.493133 H1    2 APN     0.0561 ****
 18 H5'2     3.275899    2.250678   -0.024827 H1    2 APN     0.0561 ****
 19 C'       1.449604    3.289214   -0.217087 C     2 APN     0.5520 ****
 20 O1'      1.767998    3.770682   -1.273205 O     2 APN    -0.5667 ****
 21 C7'      0.857680    0.402551    1.099946 C     2 APN     0.4519 ****
 22 O7'      0.505502    1.058337    2.054036 O     2 APN    -0.5582 ****
 23 C8'      0.288243   -1.011489    0.977878 CT    2 APN    -0.1384 ****
```

```
 24 H8'1    0.638353   -1.548180    1.850536 H1    2 APN     0.0996 ****
 25 H8'2    0.593037   -1.557702    0.104279 H1    2 APN     0.0996 ****
 26 N9     -1.153273   -0.974097    1.005139 N*    2 APN    -0.0306 ****
 27 C8     -1.981619   -0.898844    2.098267 CK    2 APN     0.1329 ****
 28 H8     -1.577259   -0.897330    3.089174 H5    2 APN     0.1676 ****
 29 N7     -3.225847   -0.831369    1.805401 NB    2 APN    -0.5631 ****
 30 C6     -4.278736   -0.782229   -0.521089 CA    2 APN     0.6557 ****
 31 N6     -5.576223   -0.734082   -0.171904 N2    2 APN    -0.8908 ****
 32 H61    -5.821383   -0.567634    0.776793 H     2 APN     0.4050 ****
 33 H62    -6.246165   -0.541126   -0.880495 H     2 APN     0.4050 ****
 34 C5     -3.247873   -0.843054    0.423001 CB    2 APN     0.0465 ****
 35 C4     -1.971958   -0.920512   -0.084086 CB    2 APN     0.3838 ****
 36 N3     -1.628350   -0.929692   -1.366627 NC    2 APN    -0.6473 ****
 37 C2     -2.681040   -0.862225   -2.150429 CQ    2 APN     0.5169 ****
 38 H2     -2.485791   -0.862420   -3.207408 H5    2 APN     0.0639 ****
 39 N1     -3.961295   -0.789586   -1.810556 NC    2 APN    -0.7226 ****
 40 N       0.351653    3.646064    0.468753 N     3 NME    -0.4149 ****
 41 H       0.105581    3.102507    1.266835 H     3 NME     0.3019 ****
 42 CH3    -0.608954    4.579103   -0.078842 CT    3 NME    -0.1405 ****
 43 HH31   -0.104131    5.471506   -0.420612 H1    3 NME     0.0930 ****
 44 HH32   -1.311496    4.847295    0.698620 H1    3 NME     0.0930 ****
 45 HH33   -1.153819    4.153269   -0.914556 H1    3 NME     0.0930 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    14 1
   14    12    15 1
   15    15    16 1
   16    15    21 1
   17    16    17 1
   18    16    18 1
   19    16    19 1
   20    19    20 1
   21    19    40 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
   25    23    25 1
   26    23    26 1
   27    26    27 1
   28    26    35 1
   29    27    28 1
   30    27    29 1
   31    29    34 1
   32    30    31 1
   33    30    34 1
   34    30    39 1
   35    31    32 1
   36    31    33 1
   37    34    35 1
   38    35    36 1
   39    36    37 1
```

```
   40    37    38 1
   41    37    39 1
   42    40    41 1
   43    40    42 1
   44    42    43 1
   45    42    44 1
   46    42    45 1
@<TRIPOS>SUBSTRUCTURE
      1  ACE             1 ****               0 ****  ****
      2  APN             7 ****               0 ****  ****
      3  NME            40 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 C 0 0 0 0
2 N1' C' 0 0 0 0
3 N 0 0 0 0 0
```

Tripos5: CPN

```
@<TRIPOS>MOLECULE
CPN
   31    31     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'     3.312760   -1.543352    0.212583 N     1 CPN    -0.4030 ****
  2 H1'1    3.702917   -1.507731    1.125329 H     1 CPN     0.2842 ****
  3 C2'     3.409872   -0.355316   -0.616244 CT    1 CPN    -0.1241 ****
  4 H2'1    3.831200   -0.625077   -1.576213 H1    1 CPN     0.1147 ****
  5 H2'2    4.106279    0.317173   -0.130994 H1    1 CPN     0.1147 ****
  6 C3'     2.078572    0.351137   -0.861740 CT    1 CPN    -0.0522 ****
  7 H3'1    1.419465   -0.270223   -1.440286 H1    1 CPN     0.1033 ****
  8 H3'2    2.265962    1.235931   -1.457963 H1    1 CPN     0.1033 ****
  9 N4'     1.418014    0.778575    0.370475 N     1 CPN    -0.0941 ****
 10 C5'     1.868150    2.048826    0.930762 CT    1 CPN    -0.1112 ****
 11 H5'1    1.604383    2.082404    1.976237 H1    1 CPN     0.0678 ****
 12 H5'2    2.941525    2.112645    0.832466 H1    1 CPN     0.0678 ****
 13 C'      1.286287    3.248636    0.184378 C     1 CPN     0.6209 ****
 14 O1'     1.962137    3.876561   -0.588907 O     1 CPN    -0.5807 ****
 15 C7'     0.289974    0.243420    0.877641 C     1 CPN     0.4391 ****
 16 O7'    -0.352871    0.790958    1.744232 O     1 CPN    -0.5502 ****
 17 C8'    -0.168725   -1.146666    0.438193 CT    1 CPN    -0.1769 ****
 18 H8'1    0.067701   -1.813354    1.258285 H1    1 CPN     0.1116 ****
 19 H8'2    0.309278   -1.520521   -0.446642 H1    1 CPN     0.1116 ****
 20 N1     -1.603343   -1.156389    0.242311 N*    1 CPN    -0.0586 ****
 21 C2     -2.082244   -0.489847   -0.899052 C     1 CPN     0.8504 ****
 22 O2     -1.272044    0.028044   -1.621377 O     1 CPN    -0.6164 ****
 23 N3     -3.421586   -0.470877   -1.105266 NC    1 CPN    -0.8062 ****
 24 C4     -4.227530   -0.981286   -0.225184 CA    1 CPN     0.8936 ****
 25 N4     -5.545324   -0.953373   -0.495430 N2    1 CPN    -0.9941 ****
 26 H41    -6.217094   -1.159663    0.205670 H     1 CPN     0.4267 ****
 27 H42    -5.840068   -0.434706   -1.291886 H     1 CPN     0.4267 ****
 28 C5     -3.781194   -1.582483    1.002879 CM    1 CPN    -0.5360 ****
 29 H5     -4.464124   -1.972947    1.730361 HA    1 CPN     0.2051 ****
 30 C6     -2.453831   -1.628734    1.178645 CM    1 CPN    -0.0567 ****
 31 H6     -2.006105   -2.050482    2.057232 H4    1 CPN     0.2189 ****
```

```
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    15 1
   11    10    11 1
   12    10    12 1
   13    10    13 1
   14    13    14 1
   15    15    16 1
   16    15    17 1
   17    17    18 1
   18    17    19 1
   19    17    20 1
   20    20    21 1
   21    20    30 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
   25    24    25 1
   26    24    28 1
   27    25    26 1
   28    25    27 1
   29    28    29 1
   30    28    30 1
   31    30    31 1
@<TRIPOS>SUBSTRUCTURE
      1  CPN              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

Tripos6: NCP

```
@<TRIPOS>MOLECULE
NCP
   33    33     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'     0.000000    0.000000    0.000000 N3    1 NCP     0.0049 ****
  2 H1'1   -0.372154   -0.940539    0.000008 H     1 NCP     0.2531 ****
  3 H1'2   -0.372155    0.470271    0.814530 H     1 NCP     0.2531 ****
  4 H1'3   -0.372159    0.470265   -0.814535 H     1 NCP     0.2531 ****
  5 C2'     1.451831    0.000000    0.000000 CT    1 NCP     0.0723 ****
  6 H2'1    1.807299    0.707039   -0.738626 HP    1 NCP     0.0819 ****
  7 H2'2    1.771691   -0.986601   -0.311450 HP    1 NCP     0.0819 ****
  8 C3'     2.081023    0.358531    1.344350 CT    1 NCP    -0.0606 ****
  9 H3'1    1.858758    1.376326    1.608771 H1    1 NCP     0.0894 ****
 10 H3'2    3.157960    0.296916    1.245012 H1    1 NCP     0.0894 ****
```

```
 11 N4'     1.683160   -0.542622    2.424638 N     1 NCP    -0.0615 ****
 12 C5'     2.432859   -1.790720    2.526052 CT    1 NCP    -0.1251 ****
 13 H5'1    1.845848   -2.506708    3.079674 H1    1 NCP     0.0674 ****
 14 H5'2    2.612995   -2.171254    1.531762 H1    1 NCP     0.0674 ****
 15 C'      3.801844   -1.591844    3.175304 C     1 NCP     0.6140 ****
 16 O1'     4.802339   -1.580938    2.506033 O     1 NCP    -0.5754 ****
 17 C7'     0.880256   -0.217213    3.456886 C     1 NCP     0.4500 ****
 18 O7'     0.790585   -0.904948    4.448353 O     1 NCP    -0.5598 ****
 19 C8'    -0.037062    1.001789    3.365122 CT    1 NCP    -0.1631 ****
 20 H8'1   -1.034977    0.614161    3.201317 H1    1 NCP     0.1106 ****
 21 H8'2    0.194123    1.681153    2.567497 H1    1 NCP     0.1106 ****
 22 N1     -0.029154    1.723758    4.620227 N*    1 NCP    -0.0668 ****
 23 C2      1.135831    2.454055    4.913163 C     1 NCP     0.8503 ****
 24 O2      2.026180    2.428175    4.105080 O     1 NCP    -0.6188 ****
 25 N3      1.179491    3.131882    6.085923 NC    1 NCP    -0.8037 ****
 26 C4      0.205489    3.031106    6.937845 CA    1 NCP     0.8897 ****
 27 N4      0.294463    3.745115    8.074801 N2    1 NCP    -0.9926 ****
 28 H41    -0.319525    3.575671    8.836168 H     1 NCP     0.4265 ****
 29 H42     1.153857    4.212176    8.257903 H     1 NCP     0.4265 ****
 30 C5     -0.957698    2.214143    6.717936 CM    1 NCP    -0.5319 ****
 31 H5     -1.738204    2.121610    7.446177 HA    1 NCP     0.2039 ****
 32 C6     -1.007101    1.578182    5.539786 CM    1 NCP    -0.0541 ****
 33 H6     -1.823841    0.938782    5.267118 H4    1 NCP     0.2174 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5     8 1
    8     8     9 1
    9     8    10 1
   10     8    11 1
   11    11    12 1
   12    11    17 1
   13    12    13 1
   14    12    14 1
   15    12    15 1
   16    15    16 1
   17    17    18 1
   18    17    19 1
   19    19    20 1
   20    19    21 1
   21    19    22 1
   22    22    23 1
   23    22    32 1
   24    23    24 1
   25    23    25 1
   26    25    26 1
   27    26    27 1
   28    26    30 1
   29    27    28 1
   30    27    29 1
   31    30    31 1
   32    30    32 1
   33    32    33 1
@<TRIPOS>SUBSTRUCTURE
      1 NCP              1 ****              0 ****  ****
@<TRIPOS>HEADTAIL
0 0
C' 1
```

```
@<TRIPOS>RESIDUECONNECT
1 0 C' 0 0 0 0
```

Tripos7:CCP

```
@<TRIPOS>MOLECULE
CCP
   32    32     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 C'      0.000000    0.000000    0.000000 C     1 CCP     0.7881 ****
  2 O1'    -0.511689   -1.124427    0.000000 O2    1 CCP    -0.8206 ****
  3 OXT    -0.540795    1.108412    0.000000 O2    1 CCP    -0.8206 ****
  4 N1'     4.547802    2.222175    1.204150 N     1 CCP    -0.4302 ****
  5 H1'1    5.114203    1.413832    1.092941 H     1 CCP     0.2931 ****
  6 C2'     3.247174    2.061524    1.828949 CT    1 CCP    -0.0573 ****
  7 H2'1    3.150528    2.775088    2.637228 H1    1 CCP     0.0960 ****
  8 H2'2    3.221350    1.069960    2.263488 H1    1 CCP     0.0960 ****
  9 C3'     2.065683    2.252761    0.880682 CT    1 CCP    -0.0479 ****
 10 H3'1    2.019998    3.267760    0.530088 H1    1 CCP     0.0959 ****
 11 H3'2    1.151130    2.079500    1.434785 H1    1 CCP     0.0959 ****
 12 N4'     2.080412    1.327413   -0.251091 N     1 CCP    -0.0212 ****
 13 C5'     1.528135    0.000000    0.000000 CT    1 CCP    -0.4798 ****
 14 H5'1    1.911977   -0.683800   -0.740811 H1    1 CCP     0.1621 ****
 15 H5'2    1.838723   -0.327232    0.980981 H1    1 CCP     0.1621 ****
 16 C7'     2.318781    1.666236   -1.533457 C     1 CCP     0.4669 ****
 17 O7'     2.067377    0.924611   -2.455853 O     1 CCP    -0.5706 ****
 18 C8'     3.020906    2.982673   -1.864926 CT    1 CCP    -0.1739 ****
 19 H8'1    4.034933    2.720339   -2.140338 H1    1 CCP     0.1088 ****
 20 H8'2    3.064266    3.686705   -1.056442 H1    1 CCP     0.1088 ****
 21 N1      2.386611    3.609236   -3.005839 N*    1 CCP    -0.0612 ****
 22 C2      1.123455    4.183709   -2.780559 C     1 CCP     0.8437 ****
 23 O2      0.672528    4.116382   -1.667670 O     1 CCP    -0.6180 ****
 24 N3      0.497863    4.768347   -3.831061 NC    1 CCP    -0.8009 ****
 25 C4      1.021589    4.716284   -5.017487 CA    1 CCP     0.8885 ****
 26 N4      0.365905    5.334296   -6.016754 N2    1 CCP    -0.9912 ****
 27 H41     0.614522    5.180692   -6.965420 H     1 CCP     0.4259 ****
 28 H42    -0.542564    5.689348   -5.819643 H     1 CCP     0.4259 ****
 29 C5      2.263387    4.050556   -5.306514 CM    1 CCP    -0.5351 ****
 30 H5      2.665244    3.993955   -6.298176 HA    1 CCP     0.2044 ****
 31 C6      2.890965    3.508437   -4.254289 CM    1 CCP    -0.0519 ****
 32 H6      3.821705    2.984115   -4.349449 H4    1 CCP     0.2183 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1    13 1
    4     4     5 1
    5     4     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    16 1
   14    13    14 1
```

```
   15    13    15 1
   16    16    17 1
   17    16    18 1
   18    18    19 1
   19    18    20 1
   20    18    21 1
   21    21    22 1
   22    21    31 1
   23    22    23 1
   24    22    24 1
   25    24    25 1
   26    25    26 1
   27    25    29 1
   28    26    27 1
   29    26    28 1
   30    29    30 1
   31    29    31 1
   32    31    32 1
@<TRIPOS>SUBSTRUCTURE
      1 CCP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

Tripos8: MCPN

```
@<TRIPOS>MOLECULE
MCPN
   43    43     3     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 CH3     3.144181   -3.930966    0.636165 CT    1 ACE    -0.1642 ****
  2 HH31    2.416220   -4.688501    0.382760 HC    1 ACE     0.0520 ****
  3 HH32    3.042308   -3.659752    1.681072 HC    1 ACE     0.0520 ****
  4 HH33    4.134273   -4.350512    0.488026 HC    1 ACE     0.0520 ****
  5 C       2.969595   -2.750664   -0.296010 C     1 ACE     0.6303 ****
  6 O       2.562419   -2.880862   -1.418429 O     1 ACE    -0.5637 ****
  7 N1'     3.312760   -1.543352    0.212583 N     2 CPN    -0.4588 ****
  8 H1'1    3.702917   -1.507731    1.125329 H     2 CPN     0.2940 ****
  9 C2'     3.409872   -0.355316   -0.616244 CT    2 CPN    -0.2064 ****
 10 H2'1    3.831200   -0.625077   -1.576213 H1    2 CPN     0.1369 ****
 11 H2'2    4.106279    0.317173   -0.130994 H1    2 CPN     0.1369 ****
 12 C3'     2.078572    0.351137   -0.861740 CT    2 CPN    -0.0645 ****
 13 H3'1    1.419465   -0.270223   -1.440286 H1    2 CPN     0.1270 ****
 14 H3'2    2.265962    1.235931   -1.457963 H1    2 CPN     0.1270 ****
 15 N4'     1.418014    0.778575    0.370475 N     2 CPN    -0.0696 ****
 16 C5'     1.868150    2.048826    0.930762 CT    2 CPN    -0.1121 ****
 17 H5'1    1.604383    2.082404    1.976237 H1    2 CPN     0.0694 ****
 18 H5'2    2.941525    2.112645    0.832466 H1    2 CPN     0.0694 ****
 19 C'      1.286287    3.248636    0.184378 C     2 CPN     0.5820 ****
 20 O1'     1.962137    3.876561   -0.588907 O     2 CPN    -0.5721 ****
 21 C7'     0.289974    0.243420    0.877641 C     2 CPN     0.4343 ****
 22 O7'    -0.352871    0.790958    1.744232 O     2 CPN    -0.5623 ****
 23 C8'    -0.168725   -1.146666    0.438193 CT    2 CPN    -0.0643 ****
 24 H8'1    0.067701   -1.813354    1.258285 H1    2 CPN     0.0679 ****
```

```
 25 H8'2    0.309278   -1.520521   -0.446642 H1    2 CPN     0.0679 ****
 26 N1     -1.603343   -1.156389    0.242311 N*    2 CPN    -0.0882 ****
 27 C2     -2.082244   -0.489847   -0.899052 C     2 CPN     0.8972 ****
 28 O2     -1.272044    0.028044   -1.621377 O     2 CPN    -0.6563 ****
 29 N3     -3.421586   -0.470877   -1.105266 NC    2 CPN    -0.7991 ****
 30 C4     -4.227530   -0.981286   -0.225184 CA    2 CPN     0.8624 ****
 31 N4     -5.545324   -0.953373   -0.495430 N2    2 CPN    -0.9823 ****
 32 H41    -6.217094   -1.159663    0.205670 H     2 CPN     0.4253 ****
 33 H42    -5.840068   -0.434706   -1.291886 H     2 CPN     0.4253 ****
 34 C5     -3.781194   -1.582483    1.002879 CM    2 CPN    -0.5186 ****
 35 H5     -4.464124   -1.972947    1.730361 HA    2 CPN     0.2026 ****
 36 C6     -2.453831   -1.628734    1.178645 CM    2 CPN    -0.0638 ****
 37 H6     -2.006105   -2.050482    2.057232 H4    2 CPN     0.2201 ****
 38 N       0.001889    3.528780    0.459686 N     3 NME    -0.4595 ****
 39 H      -0.505206    2.868136    1.006995 H     3 NME     0.3202 ****
 40 CH3    -0.729953    4.522394   -0.295657 CT    3 NME    -0.0875 ****
 41 HH31   -0.147181    5.429231   -0.367286 H1    3 NME     0.0804 ****
 42 HH32   -1.655365    4.738129    0.221813 H1    3 NME     0.0804 ****
 43 HH33   -0.959308    4.177959   -1.298297 H1    3 NME     0.0804 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    14 1
   14    12    15 1
   15    15    16 1
   16    15    21 1
   17    16    17 1
   18    16    18 1
   19    16    19 1
   20    19    20 1
   21    19    38 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
   25    23    25 1
   26    23    26 1
   27    26    27 1
   28    26    36 1
   29    27    28 1
   30    27    29 1
   31    29    30 1
   32    30    31 1
   33    30    34 1
   34    31    32 1
   35    31    33 1
   36    34    35 1
   37    34    36 1
   38    36    37 1
   39    38    39 1
   40    38    40 1
   41    40    41 1
   42    40    42 1
```

```
   43    40    43 1
@<TRIPOS>SUBSTRUCTURE
      1  ACE              1 ****               0 ****  ****
      2  CPN              7 ****               0 ****  ****
      3  NME             38 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 C 0 0 0 0
2 N1' C' 0 0 0 0
3 N 0 0 0 0 0
```

Tripos9: GPN

```
@<TRIPOS>MOLECULE
GPN
   34    35     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'     4.122369   -0.011061   -0.354039 N     1 GPN    -0.4030 ****
  2 H1'1    4.685538    0.387828    0.360264 H     1 GPN     0.2842 ****
  3 C2'     3.375242    0.883187   -1.218440 CT    1 GPN    -0.1241 ****
  4 H2'1    3.602902    0.654488   -2.252239 H1    1 GPN     0.1147 ****
  5 H2'2    3.728294    1.888000   -1.020195 H1    1 GPN     0.1147 ****
  6 C3'     1.861091    0.807184   -1.034211 CT    1 GPN    -0.0522 ****
  7 H3'1    1.508956   -0.169250   -1.308675 H1    1 GPN     0.1033 ****
  8 H3'2    1.385586    1.503556   -1.712045 H1    1 GPN     0.1033 ****
  9 N4'     1.442012    1.143452    0.322917 N     1 GPN    -0.0941 ****
 10 C5'     1.274774    2.560901    0.611615 CT    1 GPN    -0.1112 ****
 11 H5'1    1.507660    2.748575    1.647866 H1    1 GPN     0.0678 ****
 12 H5'2    1.958337    3.123362   -0.007239 H1    1 GPN     0.0678 ****
 13 C'     -0.136885    3.034278    0.262312 C     1 GPN     0.6209 ****
 14 O1'    -0.436875    3.291728   -0.881342 O     1 GPN    -0.5807 ****
 15 C7'     0.908973    0.291353    1.227567 C     1 GPN     0.4391 ****
 16 O7'     0.434614    0.681319    2.265912 O     1 GPN    -0.5502 ****
 17 C8'     0.930894   -1.218898    0.973741 CT    1 GPN    -0.2359 ****
 18 H8'1    1.388371   -1.656028    1.850521 H1    1 GPN     0.1418 ****
 19 H8'2    1.499243   -1.518921    0.110264 H1    1 GPN     0.1418 ****
 20 N9     -0.419969   -1.732213    0.885575 N*    1 GPN     0.0268 ****
 21 C8     -0.938170   -2.847500    1.505896 CK    1 GPN     0.0862 ****
 22 H8     -0.350748   -3.413771    2.199667 H5    1 GPN     0.1623 ****
 23 N7     -2.143275   -3.103753    1.168112 NB    1 GPN    -0.5377 ****
 24 C6     -3.680934   -1.852372   -0.458440 C     1 GPN     0.4075 ****
 25 O6     -4.720758   -2.438010   -0.491249 O     1 GPN    -0.5295 ****
 26 C5     -2.470564   -2.111875    0.269409 CB    1 GPN     0.2319 ****
 27 C4     -1.413271   -1.264990    0.091765 CB    1 GPN     0.0882 ****
 28 N3     -1.328232   -0.155400   -0.676135 NC    1 GPN    -0.3651 ****
 29 C2     -2.418087    0.101743   -1.321498 CA    1 GPN     0.4620 ****
 30 N2     -2.476606    1.182902   -2.135232 N2    1 GPN    -0.8775 ****
 31 H21    -1.767274    1.870625   -1.969915 H     1 GPN     0.4042 ****
 32 H22    -3.371292    1.551625   -2.367077 H     1 GPN     0.4042 ****
 33 N1     -3.526182   -0.682293   -1.245267 NA    1 GPN    -0.3251 ****
 34 H1     -4.323926   -0.465858   -1.804032 H     1 GPN     0.3136 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
```

```
    3     3     4 1
    4     3     5 1
    5     3     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    15 1
   11    10    11 1
   12    10    12 1
   13    10    13 1
   14    13    14 1
   15    15    16 1
   16    15    17 1
   17    17    18 1
   18    17    19 1
   19    17    20 1
   20    20    21 1
   21    20    27 1
   22    21    22 1
   23    21    23 1
   24    23    26 1
   25    24    25 1
   26    24    26 1
   27    24    33 1
   28    26    27 1
   29    27    28 1
   30    28    29 1
   31    29    30 1
   32    29    33 1
   33    30    31 1
   34    30    32 1
   35    33    34 1
@<TRIPOS>SUBSTRUCTURE
      1  GPN             1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

Tripos10: NGP

```
@<TRIPOS>MOLECULE
NGP
   36    37     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'     0.000000    0.000000    0.000000 N3    1 NGP     0.0049 ****
  2 H1'1   -0.372154   -0.940539    0.000008 H     1 NGP     0.2531 ****
  3 H1'2   -0.372155    0.470271    0.814530 H     1 NGP     0.2531 ****
  4 H1'3   -0.372159    0.470265   -0.814535 H     1 NGP     0.2531 ****
  5 C2'     1.450885    0.000000    0.000000 CT    1 NGP     0.0723 ****
  6 H2'1    1.808606    0.667951   -0.773789 HP    1 NGP     0.0819 ****
  7 H2'2    1.770286   -1.002133   -0.259438 HP    1 NGP     0.0819 ****
  8 C3'     2.073987    0.428275    1.326911 CT    1 NGP    -0.0606 ****
  9 H3'1    1.817015    1.449196    1.537716 H1    1 NGP     0.0894 ****
```

```
 10 H3'2    3.151890    0.397925    1.239142 H1    1 NGP     0.0894 ****
 11 N4'     1.688505   -0.439558    2.435410 N     1 NGP    -0.0615 ****
 12 C5'     2.476264   -1.652240    2.606625 CT    1 NGP    -0.1251 ****
 13 H5'1    1.854640   -2.431969    3.017574 H1    1 NGP     0.0674 ****
 14 H5'2    2.839635   -1.968519    1.639914 H1    1 NGP     0.0674 ****
 15 C'      3.703062   -1.396381    3.483199 C     1 NGP     0.6140 ****
 16 O1'     4.697580   -0.885744    3.020161 O     1 NGP    -0.5754 ****
 17 C7'     0.898835   -0.098127    3.478694 C     1 NGP     0.4500 ****
 18 O7'     0.764838   -0.818272    4.437175 O     1 NGP    -0.5598 ****
 19 C8'     0.107932    1.213202    3.453096 CT    1 NGP    -0.2061 ****
 20 H8'1   -0.919431    0.938764    3.648851 H1    1 NGP     0.1347 ****
 21 H8'2    0.144783    1.739285    2.514744 H1    1 NGP     0.1347 ****
 22 N9      0.539700    2.080286    4.529125 N*    1 NGP     0.0149 ****
 23 C8     -0.250429    2.776796    5.416686 CK    1 NGP     0.0908 ****
 24 H8     -1.315270    2.662171    5.412165 H5    1 NGP     0.1618 ****
 25 N7      0.413437    3.532922    6.203830 NB    1 NGP    -0.5404 ****
 26 C6      2.945591    3.910922    6.349885 C     1 NGP     0.4042 ****
 27 O6      3.139635    4.708326    7.216936 O     1 NGP    -0.5290 ****
 28 C5      1.728738    3.345237    5.838641 CB    1 NGP     0.2410 ****
 29 C4      1.812100    2.452105    4.808104 CB    1 NGP     0.0795 ****
 30 N3      2.909697    1.973985    4.179794 NC    1 NGP    -0.3622 ****
 31 C2      4.013894    2.459569    4.643861 CA    1 NGP     0.4564 ****
 32 N2      5.195199    2.072913    4.105801 N2    1 NGP    -0.8711 ****
 33 H21     5.155316    1.226248    3.571936 H     1 NGP     0.4024 ****
 34 H22     6.021302    2.187023    4.648668 H     1 NGP     0.4024 ****
 35 N1      4.055848    3.374183    5.648907 NA    1 NGP    -0.3229 ****
 36 H1      4.932940    3.754317    5.934680 H     1 NGP     0.3134 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5     8 1
    8     8     9 1
    9     8    10 1
   10     8    11 1
   11    11    12 1
   12    11    17 1
   13    12    13 1
   14    12    14 1
   15    12    15 1
   16    15    16 1
   17    17    18 1
   18    17    19 1
   19    19    20 1
   20    19    21 1
   21    19    22 1
   22    22    23 1
   23    22    29 1
   24    23    24 1
   25    23    25 1
   26    25    28 1
   27    26    27 1
   28    26    28 1
   29    26    35 1
   30    28    29 1
   31    29    30 1
   32    30    31 1
   33    31    32 1
   34    31    35 1
```

```
   35    32    33 1
   36    32    34 1
   37    35    36 1
@<TRIPOS>SUBSTRUCTURE
      1 NGP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
C' 1
@<TRIPOS>RESIDUECONNECT
1 0 C' 0 0 0 0
```

Tripos11: CGP

```
@<TRIPOS>MOLECULE
CGP
   35    36     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 C'      0.000000    0.000000    0.000000 C     1 CGP     0.7881 ****
  2 O1'    -0.511689   -1.124427    0.000000 O2    1 CGP    -0.8206 ****
  3 OXT    -0.540795    1.108412    0.000000 O2    1 CGP    -0.8206 ****
  4 N1'     4.733360    2.037702    1.112789 N     1 CGP    -0.4302 ****
  5 H1'1    5.292873    1.307794    0.737716 H     1 CGP     0.2931 ****
  6 C2'     3.569496    1.674956    1.899493 CT    1 CGP    -0.0573 ****
  7 H2'1    3.614306    2.178052    2.857491 H1    1 CGP     0.0960 ****
  8 H2'2    3.629641    0.609550    2.086275 H1    1 CGP     0.0960 ****
  9 C3'     2.237460    2.023037    1.238516 CT    1 CGP    -0.0479 ****
 10 H3'1    2.151970    3.087675    1.129016 H1    1 CGP     0.0959 ****
 11 H3'2    1.428178    1.718262    1.888649 H1    1 CGP     0.0959 ****
 12 N4'     2.056513    1.356583   -0.047407 N     1 CGP    -0.0212 ****
 13 C5'     1.529340    0.000000    0.000000 CT    1 CGP    -0.4798 ****
 14 H5'1    1.922895   -0.569746   -0.826908 H1    1 CGP     0.1621 ****
 15 H5'2    1.844858   -0.462054    0.923881 H1    1 CGP     0.1621 ****
 16 C7'     2.034865    1.954718   -1.259991 C     1 CGP     0.4669 ****
 17 O7'     1.713460    1.357833   -2.257827 O     1 CGP    -0.5706 ****
 18 C8'     2.464593    3.417266   -1.408498 CT    1 CGP    -0.2258 ****
 19 H8'1    3.222431    3.422366   -2.179709 H1    1 CGP     0.1344 ****
 20 H8'2    2.884856    3.852372   -0.518167 H1    1 CGP     0.1344 ****
 21 N9      1.356427    4.219949   -1.881533 N*    1 CGP     0.0345 ****
 22 C8      1.364998    5.150565   -2.896952 CK    1 CGP     0.0831 ****
 23 H8      2.240954    5.305776   -3.493320 H5    1 CGP     0.1624 ****
 24 N7      0.254792    5.768411   -3.030183 NB    1 CGP    -0.5375 ****
 25 C6     -1.923394    5.510352   -1.701674 C     1 CGP     0.4092 ****
 26 O6     -2.709425    6.290962   -2.146741 O     1 CGP    -0.5305 ****
 27 C5     -0.559594    5.233458   -2.055910 CB    1 CGP     0.2325 ****
 28 C4      0.113631    4.279918   -1.345862 CB    1 CGP     0.0799 ****
 29 N3     -0.326715    3.500988   -0.332227 NC    1 CGP    -0.3583 ****
 30 C2     -1.559702    3.725474   -0.016398 CA    1 CGP     0.4443 ****
 31 N2     -2.134228    3.020094    0.987022 N2    1 CGP    -0.8613 ****
 32 H21    -1.654590    2.176015    1.233849 H     1 CGP     0.3992 ****
 33 H22    -3.127153    2.960676    1.014221 H     1 CGP     0.3992 ****
 34 N1     -2.322437    4.666699   -0.633431 NA    1 CGP    -0.3215 ****
 35 H1     -3.253411    4.833994   -0.316008 H     1 CGP     0.3139 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1    13 1
```

```
    4      4      5 1
    5      4      6 1
    6      6      7 1
    7      6      8 1
    8      6      9 1
    9      9     10 1
   10      9     11 1
   11      9     12 1
   12     12     13 1
   13     12     16 1
   14     13     14 1
   15     13     15 1
   16     16     17 1
   17     16     18 1
   18     18     19 1
   19     18     20 1
   20     18     21 1
   21     21     22 1
   22     21     28 1
   23     22     23 1
   24     22     24 1
   25     24     27 1
   26     25     26 1
   27     25     27 1
   28     25     34 1
   29     27     28 1
   30     28     29 1
   31     29     30 1
   32     30     31 1
   33     30     34 1
   34     31     32 1
   35     31     33 1
   36     34     35 1
@<TRIPOS>SUBSTRUCTURE
      1 CGP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

Tripos12: MGPN

```
@<TRIPOS>MOLECULE
MGPN
   46    47     3     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 CH3      5.250486   -2.082960    0.231435 CT    1 ACE    -0.2270 ****
  2 HH31     4.956259   -3.121745    0.274719 HC    1 ACE     0.0719 ****
  3 HH32     5.319264   -1.679633    1.235446 HC    1 ACE     0.0719 ****
  4 HH33     6.233092   -2.026775   -0.226492 HC    1 ACE     0.0719 ****
  5 C        4.263906   -1.328219   -0.632833 C     1 ACE     0.6551 ****
  6 O        3.650754   -1.861271   -1.518940 O     1 ACE    -0.5668 ****
  7 N1'      4.122369   -0.011061   -0.354039 N     2 GPN    -0.5393 ****
  8 H1'1     4.685538    0.387828    0.360264 H     2 GPN     0.3199 ****
  9 C2'      3.375242    0.883187   -1.218440 CT    2 GPN    -0.1343 ****
```

```
 10 H2'1    3.602902    0.654488   -2.252239 H1    2 GPN     0.1313 ****
 11 H2'2    3.728294    1.888000   -1.020195 H1    2 GPN     0.1313 ****
 12 C3'     1.861091    0.807184   -1.034211 CT    2 GPN    -0.0753 ****
 13 H3'1    1.508956   -0.169250   -1.308675 H1    2 GPN     0.0928 ****
 14 H3'2    1.385586    1.503556   -1.712045 H1    2 GPN     0.0928 ****
 15 N4'     1.442012    1.143452    0.322917 N     2 GPN    -0.0660 ****
 16 C5'     1.274774    2.560901    0.611615 CT    2 GPN    -0.0463 ****
 17 H5'1    1.507660    2.748575    1.647866 H1    2 GPN     0.0638 ****
 18 H5'2    1.958337    3.123362   -0.007239 H1    2 GPN     0.0638 ****
 19 C'     -0.136885    3.034278    0.262312 C     2 GPN     0.5106 ****
 20 O1'    -0.436875    3.291728   -0.881342 O     2 GPN    -0.5496 ****
 21 C7'     0.908973    0.291353    1.227567 C     2 GPN     0.3502 ****
 22 O7'     0.434614    0.681319    2.265912 O     2 GPN    -0.5188 ****
 23 C8'     0.930894   -1.218898    0.973741 CT    2 GPN    -0.1247 ****
 24 H8'1    1.388371   -1.656028    1.850521 H1    2 GPN     0.1220 ****
 25 H8'2    1.499243   -1.518921    0.110264 H1    2 GPN     0.1220 ****
 26 N9     -0.419969   -1.732213    0.885575 N*    2 GPN    -0.0216 ****
 27 C8     -0.938170   -2.847500    1.505896 CK    2 GPN     0.0950 ****
 28 H8     -0.350748   -3.413771    2.199667 H5    2 GPN     0.1612 ****
 29 N7     -2.143275   -3.103753    1.168112 NB    2 GPN    -0.5400 ****
 30 C6     -3.680934   -1.852372   -0.458440 C     2 GPN     0.3790 ****
 31 O6     -4.720758   -2.438010   -0.491249 O     2 GPN    -0.5262 ****
 32 C5     -2.470564   -2.111875    0.269409 CB    2 GPN     0.2702 ****
 33 C4     -1.413271   -1.264990    0.091765 CB    2 GPN     0.0213 ****
 34 N3     -1.328232   -0.155400   -0.676135 NC    2 GPN    -0.2338 ****
 35 C2     -2.418087    0.101743   -1.321498 CA    2 GPN     0.3562 ****
 36 N2     -2.476606    1.182902   -2.135232 N2    2 GPN    -0.8521 ****
 37 H21    -1.767274    1.870625   -1.969915 H     2 GPN     0.4017 ****
 38 H22    -3.371292    1.551625   -2.367077 H     2 GPN     0.4017 ****
 39 N1     -3.526182   -0.682293   -1.245267 NA    2 GPN    -0.2798 ****
 40 H1     -4.323926   -0.465858   -1.804032 H     2 GPN     0.3090 ****
 41 N      -0.974453    3.159616    1.301021 N     3 NME    -0.4710 ****
 42 H      -0.720015    2.669624    2.131110 H     3 NME     0.3430 ****
 43 CH3    -2.367489    3.503665    1.118477 CT    3 NME    -0.0758 ****
 44 HH31   -2.447842    4.310199    0.405014 H1    3 NME     0.0796 ****
 45 HH32   -2.772615    3.831079    2.066429 H1    3 NME     0.0796 ****
 46 HH33   -2.951120    2.661155    0.762137 H1    3 NME     0.0796 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    14 1
   14    12    15 1
   15    15    16 1
   16    15    21 1
   17    16    17 1
   18    16    18 1
   19    16    19 1
   20    19    20 1
   21    19    41 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
```

```
   25    23    25 1
   26    23    26 1
   27    26    27 1
   28    26    33 1
   29    27    28 1
   30    27    29 1
   31    29    32 1
   32    30    31 1
   33    30    32 1
   34    30    39 1
   35    32    33 1
   36    33    34 1
   37    34    35 1
   38    35    36 1
   39    35    39 1
   40    36    37 1
   41    36    38 1
   42    39    40 1
   43    41    42 1
   44    41    43 1
   45    43    44 1
   46    43    45 1
   47    43    46 1
@<TRIPOS>SUBSTRUCTURE
      1  ACE              1 ****               0 ****  ****
      2  GPN              7 ****               0 ****  ****
      3  NME             41 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 C 0 0 0 0
2 N1' C' 0 0 0 0
3 N 0 0 0 0 0
```

Tripos13: IPN

```
@<TRIPOS>MOLECULE
IPN
   32    33     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'     3.559657   -1.542274   -0.052240 N     1 IPN    -0.4030 ****
  2 H1'1    4.203196   -1.390317    0.689405 H     1 IPN     0.2842 ****
  3 C2'     3.435242   -0.512090   -1.066643 CT    1 IPN    -0.1241 ****
  4 H2'1    3.600955   -0.948354   -2.043642 H1    1 IPN     0.1147 ****
  5 H2'2    4.224092    0.208605   -0.891637 H1    1 IPN     0.1147 ****
  6 C3'     2.077227    0.187561   -1.082969 CT    1 IPN    -0.0522 ****
  7 H3'1    1.302946   -0.511790   -1.346386 H1    1 IPN     0.1033 ****
  8 H3'2    2.088007    0.945093   -1.857007 H1    1 IPN     0.1033 ****
  9 N4'     1.758896    0.854591    0.178000 N     1 IPN    -0.0941 ****
 10 C5'     2.344620    2.178591    0.373362 CT    1 IPN    -0.1112 ****
 11 H5'1    2.460129    2.359702    1.430258 H1    1 IPN     0.0678 ****
 12 H5'2    3.314042    2.204119   -0.100560 H1    1 IPN     0.0678 ****
 13 C'      1.499276    3.265679   -0.286995 C     1 IPN     0.6209 ****
 14 O1'     1.779240    3.680739   -1.381926 O     1 IPN    -0.5807 ****
 15 C7'     0.874076    0.433347    1.102883 C     1 IPN     0.4391 ****
```

```
 16 O7'     0.553995    1.118469    2.046300 O     1 IPN    -0.5502 ****
 17 C8'     0.262332   -0.965693    1.015962 CT    1 IPN    -0.0806 ****
 18 H8'1    0.605179   -1.494691    1.895814 H1    1 IPN     0.0966 ****
 19 H8'2    0.541523   -1.538143    0.151025 H1    1 IPN     0.0966 ****
 20 N9     -1.179020   -0.886381    1.060590 N*    1 IPN    -0.0288 ****
 21 C8     -1.985563   -0.743800    2.160091 CK    1 IPN     0.0720 ****
 22 H8     -1.569220   -0.695999    3.144244 H5    1 IPN     0.1798 ****
 23 N7     -3.232071   -0.672045    1.871360 NB    1 IPN    -0.5189 ****
 24 C6     -4.397681   -0.731078   -0.404605 C     1 IPN     0.5198 ****
 25 O6     -5.568101   -0.634105   -0.193600 O     1 IPN    -0.5429 ****
 26 C5     -3.276262   -0.755586    0.501537 CB    1 IPN     0.1278 ****
 27 C4     -2.015744   -0.881579   -0.008061 CB    1 IPN     0.3258 ****
 28 N3     -1.636599   -0.981220   -1.311904 NC    1 IPN    -0.5586 ****
 29 C2     -2.632626   -0.957938   -2.110315 CQ    1 IPN     0.2453 ****
 30 H2     -2.463275   -1.030037   -3.168844 H5    1 IPN     0.1476 ****
 31 N1     -3.926211   -0.845529   -1.730193 NA    1 IPN    -0.4020 ****
 32 H1     -4.639593   -0.837005   -2.428008 H     1 IPN     0.3202 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    15 1
   11    10    11 1
   12    10    12 1
   13    10    13 1
   14    13    14 1
   15    15    16 1
   16    15    17 1
   17    17    18 1
   18    17    19 1
   19    17    20 1
   20    20    21 1
   21    20    27 1
   22    21    22 1
   23    21    23 1
   24    23    26 1
   25    24    25 1
   26    24    26 1
   27    24    31 1
   28    26    27 1
   29    27    28 1
   30    28    29 1
   31    29    30 1
   32    29    31 1
   33    31    32 1
@<TRIPOS>SUBSTRUCTURE
      1  IPN             1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

Tripos14: NIP

```
@<TRIPOS>MOLECULE
NIP
   34    35     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'     0.000000    0.000000    0.000000 N3    1 NIP     0.0049 ****
  2 H1'1   -0.372154   -0.940539    0.000008 H     1 NIP     0.2531 ****
  3 H1'2   -0.372155    0.470271    0.814530 H     1 NIP     0.2531 ****
  4 H1'3   -0.372159    0.470265   -0.814535 H     1 NIP     0.2531 ****
  5 C2'     1.451128    0.000000    0.000000 CT    1 NIP     0.0723 ****
  6 H2'1    1.810174    0.717285   -0.727257 HP    1 NIP     0.0819 ****
  7 H2'2    1.772793   -0.982310   -0.322352 HP    1 NIP     0.0819 ****
  8 C3'     2.075668    0.341031    1.351899 CT    1 NIP    -0.0606 ****
  9 H3'1    1.829710    1.351129    1.629745 H1    1 NIP     0.0894 ****
 10 H3'2    3.153618    0.299269    1.255016 H1    1 NIP     0.0894 ****
 11 N4'     1.695025   -0.590007    2.412363 N     1 NIP    -0.0615 ****
 12 C5'     2.448174   -1.840266    2.474299 CT    1 NIP    -0.1251 ****
 13 H5'1    1.828027   -2.602908    2.918110 H1    1 NIP     0.0674 ****
 14 H5'2    2.714474   -2.136679    1.471176 H1    1 NIP     0.0674 ****
 15 C'      3.754015   -1.667339    3.247160 C     1 NIP     0.6140 ****
 16 O1'     4.790076   -1.471621    2.665955 O     1 NIP    -0.5754 ****
 17 C7'     0.825303   -0.348270    3.412817 C     1 NIP     0.4500 ****
 18 O7'     0.679636   -1.113192    4.337779 O     1 NIP    -0.5598 ****
 19 C8'    -0.054693    0.902491    3.395346 CT    1 NIP    -0.0662 ****
 20 H8'1   -1.074689    0.546787    3.327489 H1    1 NIP     0.0979 ****
 21 H8'2    0.119606    1.577580    2.578255 H1    1 NIP     0.0979 ****
 22 N9      0.093992    1.637487    4.629627 N*    1 NIP    -0.0507 ****
 23 C8     -0.504236    1.392065    5.838607 CK    1 NIP     0.0847 ****
 24 H8     -1.193964    0.582940    5.955980 H5    1 NIP     0.1775 ****
 25 N7     -0.144589    2.207433    6.759544 NB    1 NIP    -0.5241 ****
 26 C6      1.504439    4.160721    6.672818 C     1 NIP     0.5174 ****
 27 O6      1.526128    4.640554    7.765108 O     1 NIP    -0.5423 ****
 28 C5      0.757460    3.043819    6.149533 CB    1 NIP     0.1370 ****
 29 C4      0.916174    2.696999    4.838428 CB    1 NIP     0.3238 ****
 30 N3      1.724375    3.269196    3.904037 NC    1 NIP    -0.5633 ****
 31 C2      2.384424    4.256723    4.372222 CQ    1 NIP     0.2549 ****
 32 H2      3.058679    4.793525    3.730672 H5    1 NIP     0.1456 ****
 33 N1      2.309411    4.702297    5.647497 NA    1 NIP    -0.4068 ****
 34 H1      2.864429    5.484515    5.923219 H     1 NIP     0.3212 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5     8 1
    8     8     9 1
    9     8    10 1
   10     8    11 1
   11    11    12 1
   12    11    17 1
   13    12    13 1
   14    12    14 1
   15    12    15 1
   16    15    16 1
   17    17    18 1
```

```
   18    17    19 1
   19    19    20 1
   20    19    21 1
   21    19    22 1
   22    22    23 1
   23    22    29 1
   24    23    24 1
   25    23    25 1
   26    25    28 1
   27    26    27 1
   28    26    28 1
   29    26    33 1
   30    28    29 1
   31    29    30 1
   32    30    31 1
   33    31    32 1
   34    31    33 1
   35    33    34 1
@<TRIPOS>SUBSTRUCTURE
      1 NIP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
C' 1
@<TRIPOS>RESIDUECONNECT
1 0 C' 0 0 0 0
```

Tripos15: CIP

```
@<TRIPOS>MOLECULE
CIP
   33    34     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 C'      0.000000    0.000000    0.000000 C     1 CIP     0.7881 ****
  2 O1'    -0.511689   -1.124427    0.000000 O2    1 CIP    -0.8206 ****
  3 OXT    -0.540795    1.108412    0.000000 O2    1 CIP    -0.8206 ****
  4 N1'     4.664268    2.092576    1.132439 N     1 CIP    -0.4302 ****
  5 H1'1    5.232991    1.304479    0.925731 H     1 CIP     0.2931 ****
  6 C2'     3.423499    1.866503    1.850182 CT    1 CIP    -0.0573 ****
  7 H2'1    3.403315    2.487593    2.736835 H1    1 CIP     0.0960 ****
  8 H2'2    3.422817    0.833296    2.173904 H1    1 CIP     0.0960 ****
  9 C3'     2.166747    2.163102    1.033736 CT    1 CIP    -0.0479 ****
 10 H3'1    2.122073    3.210049    0.788997 H1    1 CIP     0.0959 ****
 11 H3'2    1.298816    1.949844    1.645554 H1    1 CIP     0.0959 ****
 12 N4'     2.060980    1.348561   -0.175248 N     1 CIP    -0.0212 ****
 13 C5'     1.527232    0.000000    0.000000 CT    1 CIP    -0.4798 ****
 14 H5'1    1.919242   -0.637197   -0.776840 H1    1 CIP     0.1621 ****
 15 H5'2    1.840731   -0.376559    0.961746 H1    1 CIP     0.1621 ****
 16 C7'     2.270971    1.765056   -1.439446 C     1 CIP     0.4669 ****
 17 O7'     2.014053    1.072097   -2.396358 O     1 CIP    -0.5706 ****
 18 C8'     2.890620    3.135760   -1.715681 CT    1 CIP    -0.0900 ****
 19 H8'1    3.837370    2.940752   -2.202670 H1    1 CIP     0.0987 ****
 20 H8'2    3.078639    3.740180   -0.847875 H1    1 CIP     0.0987 ****
 21 N9      2.055654    3.890301   -2.620820 N*    1 CIP    -0.0296 ****
 22 C8      1.983143    3.792577   -3.986447 CK    1 CIP     0.0759 ****
 23 H8      2.605105    3.106501   -4.521840 H5    1 CIP     0.1819 ****
 24 N7      1.117264    4.580279   -4.508057 NB    1 CIP    -0.5250 ****
 25 C6     -0.469997    6.247354   -3.392885 C     1 CIP     0.5174 ****
```

```
 26 O6     -1.095631    6.763843   -4.267890 O     1 CIP    -0.5427 ****
 27 C5      0.559973    5.239285   -3.440122 CB    1 CIP     0.1399 ****
 28 C4      1.127025    4.816678   -2.272079 CB    1 CIP     0.3135 ****
 29 N3      0.844046    5.215824   -1.001531 NC    1 CIP    -0.5606 ****
 30 C2     -0.069064    6.107422   -0.964717 CQ    1 CIP     0.2509 ****
 31 H2     -0.381701    6.504926   -0.016766 H5    1 CIP     0.1466 ****
 32 N1     -0.700734    6.612927   -2.049128 NA    1 CIP    -0.4040 ****
 33 H1     -1.403395    7.310109   -1.922096 H     1 CIP     0.3205 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1    13 1
    4     4     5 1
    5     4     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    16 1
   14    13    14 1
   15    13    15 1
   16    16    17 1
   17    16    18 1
   18    18    19 1
   19    18    20 1
   20    18    21 1
   21    21    22 1
   22    21    28 1
   23    22    23 1
   24    22    24 1
   25    24    27 1
   26    25    26 1
   27    25    27 1
   28    25    32 1
   29    27    28 1
   30    28    29 1
   31    29    30 1
   32    30    31 1
   33    30    32 1
   34    32    33 1
@<TRIPOS>SUBSTRUCTURE
      1 CIP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

Tripos16: MIPN

```
@<TRIPOS>MOLECULE
MIPN
   44    45     3     0     1
SMALL
USER_CHARGES
```

```
@<TRIPOS>ATOM
  1 CH3     3.516313   -3.831103    0.770273 CT    1 ACE    -0.2252 ****
  2 HH31    2.760822   -4.601552    0.826535 HC    1 ACE     0.0719 ****
  3 HH32    3.677193   -3.403266    1.753367 HC    1 ACE     0.0719 ****
  4 HH33    4.444325   -4.288559    0.441684 HC    1 ACE     0.0719 ****
  5 C       3.091040   -2.799280   -0.251333 C     1 ACE     0.6711 ****
  6 O       2.387473   -3.074039   -1.185429 O     1 ACE    -0.5676 ****
  7 N1'     3.559657   -1.542274   -0.052240 N     2 IPN    -0.5493 ****
  8 H1'1    4.203196   -1.390317    0.689405 H     2 IPN     0.3236 ****
  9 C2'     3.435242   -0.512090   -1.066643 CT    2 IPN    -0.1739 ****
 10 H2'1    3.600955   -0.948354   -2.043642 H1    2 IPN     0.1477 ****
 11 H2'2    4.224092    0.208605   -0.891637 H1    2 IPN     0.1477 ****
 12 C3'     2.077227    0.187561   -1.082969 CT    2 IPN    -0.0463 ****
 13 H3'1    1.302946   -0.511790   -1.346386 H1    2 IPN     0.0834 ****
 14 H3'2    2.088007    0.945093   -1.857007 H1    2 IPN     0.0834 ****
 15 N4'     1.758896    0.854591    0.178000 N     2 IPN    -0.0811 ****
 16 C5'     2.344620    2.178591    0.373362 CT    2 IPN    -0.0374 ****
 17 H5'1    2.460129    2.359702    1.430258 H1    2 IPN     0.0523 ****
 18 H5'2    3.314042    2.204119   -0.100560 H1    2 IPN     0.0523 ****
 19 C'      1.499276    3.265679   -0.286995 C     2 IPN     0.5520 ****
 20 O1'     1.779240    3.680739   -1.381926 O     2 IPN    -0.5680 ****
 21 C7'     0.874076    0.433347    1.102883 C     2 IPN     0.4496 ****
 22 O7'     0.553995    1.118469    2.046300 O     2 IPN    -0.5590 ****
 23 C8'     0.262332   -0.965693    1.015962 CT    2 IPN    -0.0753 ****
 24 H8'1    0.605179   -1.494691    1.895814 H1    2 IPN     0.0810 ****
 25 H8'2    0.541523   -1.538143    0.151025 H1    2 IPN     0.0810 ****
 26 N9     -1.179020   -0.886381    1.060590 N*    2 IPN    -0.0092 ****
 27 C8     -1.985563   -0.743800    2.160091 CK    2 IPN     0.0814 ****
 28 H8     -1.569220   -0.695999    3.144244 H5    2 IPN     0.1778 ****
 29 N7     -3.232071   -0.672045    1.871360 NB    2 IPN    -0.5319 ****
 30 C6     -4.397681   -0.731078   -0.404605 C     2 IPN     0.4676 ****
 31 O6     -5.568101   -0.634105   -0.193600 O     2 IPN    -0.5361 ****
 32 C5     -3.276262   -0.755586    0.501537 CB    2 IPN     0.1935 ****
 33 C4     -2.015744   -0.881579   -0.008061 CB    2 IPN     0.2153 ****
 34 N3     -1.636599   -0.981220   -1.311904 NC    2 IPN    -0.4249 ****
 35 C2     -2.632626   -0.957938   -2.110315 CQ    2 IPN     0.1272 ****
 36 H2     -2.463275   -1.030037   -3.168844 H5    2 IPN     0.1677 ****
 37 N1     -3.926211   -0.845529   -1.730193 NA    2 IPN    -0.3258 ****
 38 H1     -4.639593   -0.837005   -2.428008 H     2 IPN     0.3083 ****
 39 N       0.448069    3.692648    0.430458 N     3 NME    -0.4086 ****
 40 H       0.218097    3.191295    1.260374 H     3 NME     0.3051 ****
 41 CH3    -0.506205    4.633386   -0.115671 CT    3 NME    -0.1585 ****
 42 HH31    0.011491    5.490432   -0.522330 H1    3 NME     0.0978 ****
 43 HH32   -1.161562    4.961288    0.679602 H1    3 NME     0.0978 ****
 44 HH33   -1.104509    4.189635   -0.904257 H1    3 NME     0.0978 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    14 1
   14    12    15 1
   15    15    16 1
   16    15    21 1
```

```
   17    16    17 1
   18    16    18 1
   19    16    19 1
   20    19    20 1
   21    19    39 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
   25    23    25 1
   26    23    26 1
   27    26    27 1
   28    26    33 1
   29    27    28 1
   30    27    29 1
   31    29    32 1
   32    30    31 1
   33    30    32 1
   34    30    37 1
   35    32    33 1
   36    33    34 1
   37    34    35 1
   38    35    36 1
   39    35    37 1
   40    37    38 1
   41    39    40 1
   42    39    41 1
   43    41    42 1
   44    41    43 1
   45    41    44 1
@<TRIPOS>SUBSTRUCTURE
      1  ACE              1 ****               0 ****  ****
      2  IPN              7 ****               0 ****  ****
      3  NME             39 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 C 0 0 0 0
2 N1' C' 0 0 0 0
3 N 0 0 0 0 0
```

Tripos17: TPN

```
@<TRIPOS>MOLECULE
TPN
   33    33     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'    -3.164169   -2.076590   -0.366700 N     1 TPN    -0.4030 ****
  2 H1'1   -3.506187   -2.157192   -1.296066 H     1 TPN     0.2842 ****
  3 C2'    -3.574308   -0.918849    0.407734 CT    1 TPN    -0.1241 ****
  4 H2'1   -4.043432   -1.248639    1.326301 H1    1 TPN     0.1147 ****
  5 H2'2   -4.323694   -0.395224   -0.172374 H1    1 TPN     0.1147 ****
  6 C3'    -2.434601    0.029400    0.773185 CT    1 TPN    -0.0522 ****
  7 H3'1   -1.736380   -0.458108    1.429867 H1    1 TPN     0.1033 ****
  8 H3'2   -2.847904    0.863153    1.327406 H1    1 TPN     0.1033 ****
  9 N4'    -1.739615    0.576891   -0.391242 N     1 TPN    -0.0941 ****
 10 C5'    -2.365484    1.731600   -1.029253 CT    1 TPN    -0.1112 ****
```

```
 11 H5'1   -2.030029    1.789884   -2.052832 H1    1 TPN     0.0678 ****
 12 H5'2   -3.436772    1.599072   -1.013366 H1    1 TPN     0.0678 ****
 13 C'     -2.071035    3.030494   -0.280901 C     1 TPN     0.6209 ****
 14 O1'    -2.885875    3.506896    0.465867 O     1 TPN    -0.5807 ****
 15 C7'    -0.493775    0.252814   -0.788018 C     1 TPN     0.4391 ****
 16 O7'     0.116709    0.906616   -1.602495 O     1 TPN    -0.5502 ****
 17 C8'     0.167399   -1.027414   -0.277164 CT    1 TPN    -0.0664 ****
 18 H8'1    0.112897   -1.739168   -1.090636 H1    1 TPN     0.0917 ****
 19 H8'2   -0.302938   -1.469186    0.579881 H1    1 TPN     0.0917 ****
 20 N1      1.565822   -0.792717    0.023138 N*    1 TPN    -0.0415 ****
 21 C2      1.842742   -0.187207    1.221262 C     1 TPN     0.6366 ****
 22 O2      0.985969    0.112374    2.004874 O     1 TPN    -0.5738 ****
 23 N3      3.169727    0.044652    1.449507 NA    1 TPN    -0.5240 ****
 24 H3      3.389568    0.465526    2.327773 H     1 TPN     0.3452 ****
 25 C4      4.231392   -0.185367    0.586981 C     1 TPN     0.6300 ****
 26 O4      5.349034    0.081073    0.916386 O     1 TPN    -0.5603 ****
 27 C5      3.845209   -0.751534   -0.709285 CM    1 TPN     0.0069 ****
 28 C6      2.555494   -1.002612   -0.917275 CM    1 TPN    -0.2988 ****
 29 H6      2.203698   -1.400670   -1.848115 H4    1 TPN     0.2494 ****
 30 C7      4.928309   -0.997851   -1.721763 CT    1 TPN    -0.2318 ****
 31 H71     5.445785   -0.076709   -1.964474 HC    1 TPN     0.0816 ****
 32 H72     5.668103   -1.689340   -1.334509 HC    1 TPN     0.0816 ****
 33 H73     4.515420   -1.410559   -2.635231 HC    1 TPN     0.0816 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    15 1
   11    10    11 1
   12    10    12 1
   13    10    13 1
   14    13    14 1
   15    15    16 1
   16    15    17 1
   17    17    18 1
   18    17    19 1
   19    17    20 1
   20    20    21 1
   21    20    28 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
   25    23    25 1
   26    25    26 1
   27    25    27 1
   28    27    28 1
   29    27    30 1
   30    28    29 1
   31    30    31 1
   32    30    32 1
   33    30    33 1
@<TRIPOS>SUBSTRUCTURE
      1  TPN              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
```

```
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

Tripos18: NTP

```
@<TRIPOS>MOLECULE
NTP
   35    35     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'      0.000000    0.000000    0.000000 N3    1 NTP     0.0049 ****
  2 H1'1    -0.372154   -0.940539    0.000008 H     1 NTP     0.2531 ****
  3 H1'2    -0.372155    0.470271    0.814530 H     1 NTP     0.2531 ****
  4 H1'3    -0.372159    0.470265   -0.814535 H     1 NTP     0.2531 ****
  5 C2'      1.452008    0.000000    0.000000 CT    1 NTP     0.0723 ****
  6 H2'1     1.811485    0.709863   -0.734490 HP    1 NTP     0.0819 ****
  7 H2'2     1.771785   -0.985591   -0.314071 HP    1 NTP     0.0819 ****
  8 C3'      2.081072    0.353118    1.345826 CT    1 NTP    -0.0606 ****
  9 H3'1     1.845385    1.366929    1.616100 H1    1 NTP     0.0894 ****
 10 H3'2     3.158194    0.304327    1.243103 H1    1 NTP     0.0894 ****
 11 N4'      1.700248   -0.561133    2.421819 N     1 NTP    -0.0615 ****
 12 C5'      2.457440   -1.806866    2.505076 CT    1 NTP    -0.1251 ****
 13 H5'1     1.863229   -2.543139    3.023203 H1    1 NTP     0.0674 ****
 14 H5'2     2.662844   -2.156139    1.504430 H1    1 NTP     0.0674 ****
 15 C'       3.809063   -1.612312    3.190000 C     1 NTP     0.6140 ****
 16 O1'      4.817371   -1.511348    2.540609 O     1 NTP    -0.5754 ****
 17 C7'      0.878322   -0.269011    3.448307 C     1 NTP     0.4500 ****
 18 O7'      0.792780   -0.971906    4.429195 O     1 NTP    -0.5598 ****
 19 C8'     -0.056744    0.937589    3.365388 CT    1 NTP    -0.0520 ****
 20 H8'1    -1.042726    0.540421    3.161924 H1    1 NTP     0.0911 ****
 21 H8'2     0.180976    1.642556    2.592461 H1    1 NTP     0.0911 ****
 22 N1      -0.104447    1.632386    4.636542 N*    1 NTP    -0.0517 ****
 23 C2       0.939154    2.476729    4.913690 C     1 NTP     0.6341 ****
 24 O2       1.837971    2.664708    4.142556 O     1 NTP    -0.5750 ****
 25 N3       0.870934    3.085554    6.134928 NA    1 NTP    -0.5177 ****
 26 H3       1.612843    3.720536    6.342740 H     1 NTP     0.3442 ****
 27 C4      -0.072382    2.883646    7.131613 C     1 NTP     0.6257 ****
 28 O4       0.000058    3.479989    8.164933 O     1 NTP    -0.5597 ****
 29 C5      -1.106094    1.899713    6.795010 CM    1 NTP     0.0099 ****
 30 C6      -1.052923    1.330010    5.594062 CM    1 NTP    -0.2934 ****
 31 H6      -1.767408    0.590947    5.290795 H4    1 NTP     0.2468 ****
 32 C7      -2.148437    1.589409    7.832373 CT    1 NTP    -0.2418 ****
 33 H71     -1.689595    1.210718    8.738601 HC    1 NTP     0.0843 ****
 34 H72     -2.702209    2.481776    8.101512 HC    1 NTP     0.0843 ****
 35 H73     -2.848081    0.846914    7.465728 HC    1 NTP     0.0843 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5     8 1
    8     8     9 1
    9     8    10 1
   10     8    11 1
   11    11    12 1
```

```
   12    11    17 1
   13    12    13 1
   14    12    14 1
   15    12    15 1
   16    15    16 1
   17    17    18 1
   18    17    19 1
   19    19    20 1
   20    19    21 1
   21    19    22 1
   22    22    23 1
   23    22    30 1
   24    23    24 1
   25    23    25 1
   26    25    26 1
   27    25    27 1
   28    27    28 1
   29    27    29 1
   30    29    30 1
   31    29    32 1
   32    30    31 1
   33    32    33 1
   34    32    34 1
   35    32    35 1
@<TRIPOS>SUBSTRUCTURE
      1 NTP              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
C' 1
@<TRIPOS>RESIDUECONNECT
1 0 C' 0 0 0 0
```

Tripos19: CTP

```
@<TRIPOS>MOLECULE
CTP
   34    34     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 C'      0.000000    0.000000    0.000000 C     1 CTP     0.7881 ****
  2 O1'    -0.511689   -1.124427    0.000000 O2    1 CTP    -0.8206 ****
  3 OXT    -0.540795    1.108412    0.000000 O2    1 CTP    -0.8206 ****
  4 N1'     4.594918    2.187026    1.178241 N     1 CTP    -0.4302 ****
  5 H1'1    5.184625    1.392533    1.087575 H     1 CTP     0.2931 ****
  6 C2'     3.310261    2.005883    1.830291 CT    1 CTP    -0.0573 ****
  7 H2'1    3.231111    2.690954    2.665166 H1    1 CTP     0.0960 ****
  8 H2'2    3.293666    0.999863    2.230210 H1    1 CTP     0.0960 ****
  9 C3'     2.105344    2.229808    0.919414 CT    1 CTP    -0.0479 ****
 10 H3'1    2.063583    3.254873    0.597069 H1    1 CTP     0.0959 ****
 11 H3'2    1.204634    2.050699    1.493665 H1    1 CTP     0.0959 ****
 12 N4'     2.076301    1.332626   -0.235083 N     1 CTP    -0.0212 ****
 13 C5'     1.527696    0.000000    0.000000 CT    1 CTP    -0.4798 ****
 14 H5'1    1.914893   -0.674301   -0.747689 H1    1 CTP     0.1621 ****
 15 H5'2    1.839074   -0.336457    0.977402 H1    1 CTP     0.1621 ****
 16 C7'     2.306081    1.688413   -1.513829 C     1 CTP     0.4669 ****
 17 O7'     2.031510    0.967337   -2.445592 O     1 CTP    -0.5706 ****
 18 C8'     3.016889    3.004414   -1.830052 CT    1 CTP    -0.0673 ****
```

```
 19 H8'1    4.031033    2.740822   -2.100850 H1    1 CTP     0.0902 ****
 20 H8'2    3.063321    3.701184   -1.015644 H1    1 CTP     0.0902 ****
 21 N1      2.400705    3.654210   -2.969760 N*    1 CTP    -0.0463 ****
 22 C2      1.245600    4.351399   -2.727920 C     1 CTP     0.6285 ****
 23 O2      0.772164    4.453010   -1.630929 O     1 CTP    -0.5743 ****
 24 N3      0.680895    4.920280   -3.834368 NA    1 CTP    -0.5149 ****
 25 H3     -0.149541    5.449972   -3.671221 H     1 CTP     0.3440 ****
 26 C4      1.094352    4.802409   -5.153138 C     1 CTP     0.6249 ****
 27 O4      0.491041    5.348029   -6.028886 O     1 CTP    -0.5597 ****
 28 C5      2.285142    3.967683   -5.340862 CM    1 CTP     0.0064 ****
 29 C6      2.849082    3.438753   -4.258333 CM    1 CTP    -0.2916 ****
 30 H6      3.711306    2.805822   -4.326463 H4    1 CTP     0.2495 ****
 31 C7      2.781780    3.753244   -6.743094 CT    1 CTP    -0.2329 ****
 32 H71     2.017752    3.288368   -7.355748 HC    1 CTP     0.0818 ****
 33 H72     3.037418    4.697613   -7.210250 HC    1 CTP     0.0818 ****
 34 H73     3.659725    3.117307   -6.748038 HC    1 CTP     0.0818 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1    13 1
    4     4     5 1
    5     4     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    16 1
   14    13    14 1
   15    13    15 1
   16    16    17 1
   17    16    18 1
   18    18    19 1
   19    18    20 1
   20    18    21 1
   21    21    22 1
   22    21    29 1
   23    22    23 1
   24    22    24 1
   25    24    25 1
   26    24    26 1
   27    26    27 1
   28    26    28 1
   29    28    29 1
   30    28    31 1
   31    29    30 1
   32    31    32 1
   33    31    33 1
   34    31    34 1
@<TRIPOS>SUBSTRUCTURE
      1 CTP              1 ****              0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

Tripos20: MTPN

```
@<TRIPOS>MOLECULE
MTPN
   45    45     3     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 CH3     -2.511506   -4.400783   -0.663336 CT    1 ACE    -0.1543 ****
  2 HH31    -1.683830   -5.002251   -0.315556 HC    1 ACE     0.0513 ****
  3 HH32    -2.368080   -4.148382   -1.708026 HC    1 ACE     0.0513 ****
  4 HH33    -3.418621   -4.990733   -0.577015 HC    1 ACE     0.0513 ****
  5 C       -2.637354   -3.178981    0.220977 C     1 ACE     0.6257 ****
  6 O       -2.298076   -3.186372    1.373090 O     1 ACE    -0.5596 ****
  7 N1'     -3.164169   -2.076590   -0.366700 N     2 TPN    -0.4761 ****
  8 H1'1    -3.506187   -2.157192   -1.296066 H     2 TPN     0.3011 ****
  9 C2'     -3.574308   -0.918849    0.407734 CT    2 TPN    -0.1798 ****
 10 H2'1    -4.043432   -1.248639    1.326301 H1    2 TPN     0.1351 ****
 11 H2'2    -4.323694   -0.395224   -0.172374 H1    2 TPN     0.1351 ****
 12 C3'     -2.434601    0.029400    0.773185 CT    2 TPN    -0.0676 ****
 13 H3'1    -1.736380   -0.458108    1.429867 H1    2 TPN     0.1221 ****
 14 H3'2    -2.847904    0.863153    1.327406 H1    2 TPN     0.1221 ****
 15 N4'     -1.739615    0.576891   -0.391242 N     2 TPN    -0.1000 ****
 16 C5'     -2.365484    1.731600   -1.029253 CT    2 TPN    -0.0992 ****
 17 H5'1    -2.030029    1.789884   -2.052832 H1    2 TPN     0.0703 ****
 18 H5'2    -3.436772    1.599072   -1.013366 H1    2 TPN     0.0703 ****
 19 C'      -2.071035    3.030494   -0.280901 C     2 TPN     0.5707 ****
 20 O1'     -2.885875    3.506896    0.465867 O     2 TPN    -0.5674 ****
 21 C7'     -0.493775    0.252814   -0.788018 C     2 TPN     0.4407 ****
 22 O7'      0.116709    0.906616   -1.602495 O     2 TPN    -0.5491 ****
 23 C8'      0.167399   -1.027414   -0.277164 CT    2 TPN    -0.0520 ****
 24 H8'1     0.112897   -1.739168   -1.090636 H1    2 TPN     0.0755 ****
 25 H8'2    -0.302938   -1.469186    0.579881 H1    2 TPN     0.0755 ****
 26 N1       1.565822   -0.792717    0.023138 N*    2 TPN    -0.0412 ****
 27 C2       1.842742   -0.187207    1.221262 C     2 TPN     0.6620 ****
 28 O2       0.985969    0.112374    2.004874 O     2 TPN    -0.5893 ****
 29 N3       3.169727    0.044652    1.449507 NA    2 TPN    -0.5381 ****
 30 H3       3.389568    0.465526    2.327773 H     2 TPN     0.3504 ****
 31 C4       4.231392   -0.185367    0.586981 C     2 TPN     0.6296 ****
 32 O4       5.349034    0.081073    0.916386 O     2 TPN    -0.5598 ****
 33 C5       3.845209   -0.751534   -0.709285 CM    2 TPN     0.0092 ****
 34 C6       2.555494   -1.002612   -0.917275 CM    2 TPN    -0.3059 ****
 35 H6       2.203698   -1.400670   -1.848115 H4    2 TPN     0.2503 ****
 36 C7       4.928309   -0.997851   -1.721763 CT    2 TPN    -0.2157 ****
 37 H71      5.445785   -0.076709   -1.964474 HC    2 TPN     0.0772 ****
 38 H72      5.668103   -1.689340   -1.334509 HC    2 TPN     0.0772 ****
 39 H73      4.515420   -1.410559   -2.635231 HC    2 TPN     0.0772 ****
 40 N       -0.865254    3.571266   -0.519214 N     3 NME    -0.4370 ****
 41 H       -0.213735    3.033768   -1.047775 H     3 NME     0.3127 ****
 42 CH3     -0.399200    4.726533    0.216784 CT    3 NME    -0.1122 ****
 43 HH31    -1.150884    5.502623    0.198654 H1    3 NME     0.0868 ****
 44 HH32     0.500950    5.096914   -0.255148 H1    3 NME     0.0868 ****
 45 HH33    -0.181298    4.484551    1.251482 H1    3 NME     0.0868 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
```

```
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    14 1
   14    12    15 1
   15    15    16 1
   16    15    21 1
   17    16    17 1
   18    16    18 1
   19    16    19 1
   20    19    20 1
   21    19    40 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
   25    23    25 1
   26    23    26 1
   27    26    27 1
   28    26    34 1
   29    27    28 1
   30    27    29 1
   31    29    30 1
   32    29    31 1
   33    31    32 1
   34    31    33 1
   35    33    34 1
   36    33    36 1
   37    34    35 1
   38    36    37 1
   39    36    38 1
   40    36    39 1
   41    40    41 1
   42    40    42 1
   43    42    43 1
   44    42    44 1
   45    42    45 1
@<TRIPOS>SUBSTRUCTURE
      1  ACE              1 ****               0 ****  ****
      2  TPN              7 ****               0 ****  ****
      3  NME             40 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 C 0 0 0 0
2 N1' C' 0 0 0 0
3 N 0 0 0 0 0
```

Tripos21: UPN

```
@<TRIPOS>MOLECULE
UPN
   30    30     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
```

```
  1 N1'    -3.254846   -1.627229   -0.278295 N     1 UPN    -0.4030 ****
  2 H1'1   -3.684063   -1.618138   -1.174397 H     1 UPN     0.2842 ****
  3 C2'    -3.417411   -0.453557    0.561161 CT    1 UPN    -0.1241 ****
  4 H2'1   -3.861967   -0.744434    1.504820 H1    1 UPN     0.1147 ****
  5 H2'2   -4.116231    0.203849    0.059476 H1    1 UPN     0.1147 ****
  6 C3'    -2.118053    0.290355    0.861095 CT    1 UPN    -0.0522 ****
  7 H3'1   -1.458527   -0.326589    1.444893 H1    1 UPN     0.1033 ****
  8 H3'2   -2.351584    1.154613    1.470650 H1    1 UPN     0.1033 ****
  9 N4'    -1.434427    0.771128   -0.339044 N     1 UPN    -0.0941 ****
 10 C5'    -1.916897    2.033395   -0.892647 CT    1 UPN    -0.1112 ****
 11 H5'1   -1.656731    2.078589   -1.938533 H1    1 UPN     0.0678 ****
 12 H5'2   -2.991117    2.068858   -0.791737 H1    1 UPN     0.0678 ****
 13 C'     -1.362836    3.239471   -0.136210 C     1 UPN     0.6209 ****
 14 O1'    -2.029484    3.803357    0.692087 O     1 UPN    -0.5807 ****
 15 C7'    -0.290973    0.272717   -0.845814 C     1 UPN     0.4391 ****
 16 O7'     0.350918    0.853260   -1.691252 O     1 UPN    -0.5502 ****
 17 C8'     0.198827   -1.112754   -0.424350 CT    1 UPN    -0.0516 ****
 18 H8'1   -0.027035   -1.775518   -1.249653 H1    1 UPN     0.0866 ****
 19 H8'2   -0.265877   -1.511529    0.456298 H1    1 UPN     0.0866 ****
 20 N1      1.636338   -1.105349   -0.233737 N*    1 UPN    -0.0156 ****
 21 C2      2.097211   -0.613701    0.965992 C     1 UPN     0.5893 ****
 22 O2      1.354926   -0.236002    1.827302 O     1 UPN    -0.5655 ****
 23 N3      3.456597   -0.592628    1.095482 NA    1 UPN    -0.4340 ****
 24 H3      3.799540   -0.257334    1.971222 H     1 UPN     0.3197 ****
 25 C4      4.409555   -0.925403    0.138007 C     1 UPN     0.6418 ****
 26 O4      5.574047   -0.843589    0.387444 O     1 UPN    -0.5643 ****
 27 C5      3.828236   -1.349290   -1.130494 CM    1 UPN    -0.3150 ****
 28 H5      4.487720   -1.598074   -1.935910 HA    1 UPN     0.1763 ****
 29 C6      2.505560   -1.401481   -1.258114 CM    1 UPN    -0.1842 ****
 30 H6      2.035852   -1.690560   -2.176543 H4    1 UPN     0.2296 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    15 1
   11    10    11 1
   12    10    12 1
   13    10    13 1
   14    13    14 1
   15    15    16 1
   16    15    17 1
   17    17    18 1
   18    17    19 1
   19    17    20 1
   20    20    21 1
   21    20    29 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
   25    23    25 1
   26    25    26 1
   27    25    27 1
   28    27    28 1
   29    27    29 1
   30    29    30 1
@<TRIPOS>SUBSTRUCTURE
```

```
      1  UPN              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

Tripos22: NUP

```
@<TRIPOS>MOLECULE
NUP
   32    32     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 N1'     0.000000    0.000000    0.000000 N3    1 NUP     0.0049 ****
  2 H1'1   -0.372154   -0.940539    0.000008 H     1 NUP     0.2531 ****
  3 H1'2   -0.372155    0.470271    0.814530 H     1 NUP     0.2531 ****
  4 H1'3   -0.372159    0.470265   -0.814535 H     1 NUP     0.2531 ****
  5 C2'     1.452109    0.000000    0.000000 CT    1 NUP     0.0723 ****
  6 H2'1    1.812299    0.711263   -0.732871 HP    1 NUP     0.0819 ****
  7 H2'2    1.771672   -0.985033   -0.315927 HP    1 NUP     0.0819 ****
  8 C3'     2.081304    0.350353    1.346500 CT    1 NUP    -0.0606 ****
  9 H3'1    1.846313    1.363945    1.618205 H1    1 NUP     0.0894 ****
 10 H3'2    3.158368    0.300601    1.244060 H1    1 NUP     0.0894 ****
 11 N4'     1.699564   -0.565259    2.421094 N     1 NUP    -0.0615 ****
 12 C5'     2.453774   -1.813144    2.501840 CT    1 NUP    -0.1251 ****
 13 H5'1    1.856556   -2.549965    3.015681 H1    1 NUP     0.0674 ****
 14 H5'2    2.661033   -2.159104    1.500469 H1    1 NUP     0.0674 ****
 15 C'      3.803854   -1.622760    3.190926 C     1 NUP     0.6140 ****
 16 O1'     4.813082   -1.515316    2.544106 O     1 NUP    -0.5754 ****
 17 C7'     0.875750   -0.275429    3.445811 C     1 NUP     0.4500 ****
 18 O7'     0.784373   -0.981482    4.424055 O     1 NUP    -0.5598 ****
 19 C8'    -0.055249    0.934575    3.367104 CT    1 NUP    -0.0400 ****
 20 H8'1   -1.042748    0.541471    3.162874 H1    1 NUP     0.0869 ****
 21 H8'2    0.183561    1.643560    2.598454 H1    1 NUP     0.0869 ****
 22 N1     -0.100003    1.622919    4.642646 N*    1 NUP    -0.0241 ****
 23 C2      0.939335    2.481954    4.917068 C     1 NUP     0.5863 ****
 24 O2      1.825630    2.680119    4.135588 O     1 NUP    -0.5667 ****
 25 N3      0.879040    3.086369    6.140256 NA    1 NUP    -0.4267 ****
 26 H3      1.617910    3.727759    6.339359 H     1 NUP     0.3185 ****
 27 C4     -0.050122    2.870611    7.153040 C     1 NUP     0.6353 ****
 28 O4      0.029837    3.463892    8.185793 O     1 NUP    -0.5631 ****
 29 C5     -1.060964    1.878469    6.805825 CM    1 NUP    -0.3089 ****
 30 H5     -1.801481    1.627518    7.536691 HA    1 NUP     0.1749 ****
 31 C6     -1.028849    1.305058    5.606387 CM    1 NUP    -0.1830 ****
 32 H6     -1.740853    0.559276    5.315589 H4    1 NUP     0.2282 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5     8 1
    8     8     9 1
    9     8    10 1
   10     8    11 1
```

```
   11    11    12 1
   12    11    17 1
   13    12    13 1
   14    12    14 1
   15    12    15 1
   16    15    16 1
   17    17    18 1
   18    17    19 1
   19    19    20 1
   20    19    21 1
   21    19    22 1
   22    22    23 1
   23    22    31 1
   24    23    24 1
   25    23    25 1
   26    25    26 1
   27    25    27 1
   28    27    28 1
   29    27    29 1
   30    29    30 1
   31    29    31 1
   32    31    32 1
@<TRIPOS>SUBSTRUCTURE
      1 NUP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
C' 1
@<TRIPOS>RESIDUECONNECT
1 0 C' 0 0 0 0
```

Tripos23: CUP

```
@<TRIPOS>MOLECULE
CUP
   31    31     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
  1 C'      0.000000    0.000000    0.000000 C     1 CUP     0.7881 ****
  2 O1'    -0.511689   -1.124427    0.000000 O2    1 CUP    -0.8206 ****
  3 OXT    -0.540795    1.108412    0.000000 O2    1 CUP    -0.8206 ****
  4 N1'     4.598729    2.180074    1.176275 N     1 CUP    -0.4302 ****
  5 H1'1    5.190931    1.386592    1.092655 H     1 CUP     0.2931 ****
  6 C2'     3.315434    1.999899    1.831492 CT    1 CUP    -0.0573 ****
  7 H2'1    3.239051    2.683291    2.668076 H1    1 CUP     0.0960 ****
  8 H2'2    3.298284    0.993071    2.229263 H1    1 CUP     0.0960 ****
  9 C3'     2.108361    2.227674    0.924414 CT    1 CUP    -0.0479 ****
 10 H3'1    2.067159    3.253702    0.605067 H1    1 CUP     0.0959 ****
 11 H3'2    1.208917    2.047858    1.500351 H1    1 CUP     0.0959 ****
 12 N4'     2.075114    1.333719   -0.232548 N     1 CUP    -0.0212 ****
 13 C5'     1.527678    0.000000    0.000000 CT    1 CUP    -0.4798 ****
 14 H5'1    1.915516   -0.672557   -0.748902 H1    1 CUP     0.1621 ****
 15 H5'2    1.839314   -0.337875    0.976787 H1    1 CUP     0.1621 ****
 16 C7'     2.304821    1.691703   -1.509967 C     1 CUP     0.4669 ****
 17 O7'     2.032312    0.972538   -2.443969 O     1 CUP    -0.5706 ****
 18 C8'     3.012295    3.009724   -1.825188 CT    1 CUP    -0.0535 ****
 19 H8'1    4.026105    2.748689   -2.099861 H1    1 CUP     0.0856 ****
 20 H8'2    3.059604    3.707024   -1.011525 H1    1 CUP     0.0856 ****
```

```
 21 N1      2.390707    3.658532   -2.963390 N*    1 CUP    -0.0186 ****
 22 C2      1.241358    4.373070   -2.714590 C     1 CUP     0.5820 ****
 23 O2      0.785902    4.483801   -1.611958 O     1 CUP    -0.5661 ****
 24 N3      0.667579    4.942185   -3.815511 NA    1 CUP    -0.4276 ****
 25 H3     -0.155136    5.480476   -3.641340 H     1 CUP     0.3191 ****
 26 C4      1.058777    4.812173   -5.144303 C     1 CUP     0.6385 ****
 27 O4      0.448336    5.358730   -6.012399 O     1 CUP    -0.5640 ****
 28 C5      2.232367    3.964335   -5.319130 CM    1 CUP    -0.3144 ****
 29 H5      2.588400    3.784728   -6.312352 HA    1 CUP     0.1756 ****
 30 C6      2.816473    3.428732   -4.251220 CM    1 CUP    -0.1798 ****
 31 H6      3.669815    2.786407   -4.334471 H4    1 CUP     0.2297 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1    13 1
    4     4     5 1
    5     4     6 1
    6     6     7 1
    7     6     8 1
    8     6     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    16 1
   14    13    14 1
   15    13    15 1
   16    16    17 1
   17    16    18 1
   18    18    19 1
   19    18    20 1
   20    18    21 1
   21    21    22 1
   22    21    30 1
   23    22    23 1
   24    22    24 1
   25    24    25 1
   26    24    26 1
   27    26    27 1
   28    26    28 1
   29    28    29 1
   30    28    30 1
   31    30    31 1
@<TRIPOS>SUBSTRUCTURE
      1 CUP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

Tripos24: MUPN

```
@<TRIPOS>MOLECULE
MUPN
   42    42     3     0     1
SMALL
USER_CHARGES
```

```
@<TRIPOS>ATOM
  1 CH3     -3.003190   -4.010912   -0.692202 CT     1 ACE    -0.1596 ****
  2 HH31    -2.257828   -4.748586   -0.431447 HC     1 ACE     0.0528 ****
  3 HH32    -2.903897   -3.742706   -1.738084 HC     1 ACE     0.0528 ****
  4 HH33    -3.983526   -4.453537   -0.545924 HC     1 ACE     0.0528 ****
  5 C       -2.861669   -2.820591    0.231677 C      1 ACE     0.6302 ****
  6 O       -2.433479   -2.925728    1.348865 O      1 ACE    -0.5606 ****
  7 N1'     -3.254846   -1.627229   -0.278295 N      2 UPN    -0.4780 ****
  8 H1'1    -3.684063   -1.618138   -1.174397 H      2 UPN     0.3011 ****
  9 C2'     -3.417411   -0.453557    0.561161 CT     2 UPN    -0.1800 ****
 10 H2'1    -3.861967   -0.744434    1.504820 H1     2 UPN     0.1370 ****
 11 H2'2    -4.116231    0.203849    0.059476 H1     2 UPN     0.1370 ****
 12 C3'     -2.118053    0.290355    0.861095 CT     2 UPN    -0.0797 ****
 13 H3'1    -1.458527   -0.326589    1.444893 H1     2 UPN     0.1269 ****
 14 H3'2    -2.351584    1.154613    1.470650 H1     2 UPN     0.1269 ****
 15 N4'     -1.434427    0.771128   -0.339044 N      2 UPN    -0.1075 ****
 16 C5'     -1.916897    2.033395   -0.892647 CT     2 UPN    -0.0994 ****
 17 H5'1    -1.656731    2.078589   -1.938533 H1     2 UPN     0.0708 ****
 18 H5'2    -2.991117    2.068858   -0.791737 H1     2 UPN     0.0708 ****
 19 C'      -1.362836    3.239471   -0.136210 C      2 UPN     0.5717 ****
 20 O1'     -2.029484    3.803357    0.692087 O      2 UPN    -0.5673 ****
 21 C7'     -0.290973    0.272717   -0.845814 C      2 UPN     0.4544 ****
 22 O7'      0.350918    0.853260   -1.691252 O      2 UPN    -0.5498 ****
 23 C8'      0.198827   -1.112754   -0.424350 CT     2 UPN    -0.0603 ****
 24 H8'1    -0.027035   -1.775518   -1.249653 H1     2 UPN     0.0739 ****
 25 H8'2    -0.265877   -1.511529    0.456298 H1     2 UPN     0.0739 ****
 26 N1       1.636338   -1.105349   -0.233737 N*     2 UPN    -0.0052 ****
 27 C2       2.097211   -0.613701    0.965992 C      2 UPN     0.6187 ****
 28 O2       1.354926   -0.236002    1.827302 O      2 UPN    -0.5825 ****
 29 N3       3.456597   -0.592628    1.095482 NA     2 UPN    -0.4525 ****
 30 H3       3.799540   -0.257334    1.971222 H      2 UPN     0.3261 ****
 31 C4       4.409555   -0.925403    0.138007 C      2 UPN     0.6427 ****
 32 O4       5.574047   -0.843589    0.387444 O      2 UPN    -0.5638 ****
 33 C5       3.828236   -1.349290   -1.130494 CM     2 UPN    -0.3153 ****
 34 H5       4.487720   -1.598074   -1.935910 HA     2 UPN     0.1770 ****
 35 C6       2.505560   -1.401481   -1.258114 CM     2 UPN    -0.1892 ****
 36 H6       2.035852   -1.690560   -2.176543 H4     2 UPN     0.2305 ****
 37 N       -0.110798    3.598190   -0.462810 N      3 NME    -0.4360 ****
 38 H        0.403391    2.990697   -1.062166 H      3 NME     0.3119 ****
 39 CH3      0.587826    4.638500    0.260751 CT     3 NME    -0.1202 ****
 40 HH31    -0.031804    5.521352    0.326512 H1     3 NME     0.0890 ****
 41 HH32     1.494443    4.883562   -0.275972 H1     3 NME     0.0890 ****
 42 HH33     0.846659    4.325829    1.266638 H1     3 NME     0.0890 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    12 1
   12    12    13 1
   13    12    14 1
   14    12    15 1
   15    15    16 1
   16    15    21 1
   17    16    17 1
   18    16    18 1
```

```
   19    16    19 1
   20    19    20 1
   21    19    37 1
   22    21    22 1
   23    21    23 1
   24    23    24 1
   25    23    25 1
   26    23    26 1
   27    26    27 1
   28    26    35 1
   29    27    28 1
   30    27    29 1
   31    29    30 1
   32    29    31 1
   33    31    32 1
   34    31    33 1
   35    33    34 1
   36    33    35 1
   37    35    36 1
   38    37    38 1
   39    37    39 1
   40    39    40 1
   41    39    41 1
   42    39    42 1
@<TRIPOS>SUBSTRUCTURE
      1  ACE              1 ****               0 ****  ****
      2  UPN              7 ****               0 ****  ****
      3  NME             37 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 C 0 0 0 0
2 N1' C' 0 0 0 0
3 N 0 0 0 0 0
```

frcmod.pna

```
PNA frcmod file: frcmod.pna
MASS

BOND
CQ-NA        502.0    1.324        by analogy to CQ-NC, parm99.dat

ANGL
C -CT-N*    50.00     110.13       FF cst by analogy to CT-CT-N*; Eq.val. from
QM geo.opt.
CQ-NA-C     70.0      118.60       by analogy to CA-NC-CQ, parm99.dat
H5-CQ-NA    50.0      115.45       by analogy to H5-CQ-NC, parm99.dat
NC-CQ-NA    70.0      129.10       by analogy to NC-CQ-NC, parm99.dat
CQ-NA-H     50.0      118.00       by analogy to CA-NA-H, parm99.dat

DIHEDRAL
X -CQ-NA-X    4    6.00        180.0         2.0      by analogy to X -CA-NA-X,
parm99.dat

NONB
```

8.2. Gamma modified PNA

GAP

```
@<TRIPOS>MOLECULE
GAP
    37    38     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'     2.658286  -1.774021   0.309197 N     1   GAP -0.3871  0.0000 ****
   2 H1'1    2.517951  -2.759747   0.268759 H     1   GAP  0.2784  0.0000 ****
   3 C2'     2.512277  -1.069824  -0.955646 CT    1   GAP  0.0528  0.0000 ****
   4 H2'1    3.410703  -0.490010  -1.113776 H1    1   GAP  0.1363  0.0000 ****
   5 C3'     1.288324  -0.144835  -0.973096 CT    1   GAP -0.1234  0.0000 ****
   6 H3'1    0.383739  -0.730991  -0.927254 H1    1   GAP  0.0809  0.0000 ****
   7 H3'2    1.264585   0.390541  -1.914799 H1    1   GAP  0.0809  0.0000 ****
   8 N4'     1.282584   0.854551   0.086253 N     1   GAP -0.1034  0.0000 ****
   9 C5'     2.106475   2.040867  -0.129854 CT    1   GAP -0.1424  0.0000 ****
  10 H5'1    2.401438   2.436687   0.827763 H1    1   GAP  0.0802  0.0000 ****
  11 H5'2    2.992408   1.751034  -0.670394 H1    1   GAP  0.0802  0.0000 ****
  12 C'      1.360971   3.072713  -0.969332 C     1   GAP  0.6199  0.0000 ****
  13 O1'     1.478995   3.099056  -2.168198 O     1   GAP -0.5806  0.0000 ****
  14 C7'     0.496122   0.835159   1.175840 C     1   GAP  0.5312  0.0000 ****
  15 O7'     0.374438   1.789833   1.910183 O     1   GAP -0.5649  0.0000 ****
  16 C8'    -0.284550  -0.426263   1.541418 CT    1   GAP -0.0325  0.0000 ****
  17 H8'1   -0.142787  -0.557902   2.605388 H1    1   GAP  0.0716  0.0000 ****
  18 H8'2    0.046395  -1.316442   1.039739 H1    1   GAP  0.0716  0.0000 ****
  19 C9'     2.387422  -2.077500  -2.092261 CT    1   GAP  0.1284  0.0000 ****
  20 H9'1    2.350846  -1.540092  -3.035476 H1    1   GAP  0.0497  0.0000 ****
  21 H9'2    1.458038  -2.637070  -1.992424 H1    1   GAP  0.0497  0.0000 ****
  22 O10'    3.491037  -2.940017  -2.041878 OH    1   GAP -0.6690  0.0000 ****
  23 H10'    3.469878  -3.530498  -2.781888 HO    1   GAP  0.4382  0.0000 ****
  24 N9     -1.696155  -0.258620   1.262813 N*    1   GAP -0.1390  0.0000 ****
  25 C8     -2.584459   0.581549   1.893238 CK    1   GAP  0.1714  0.0000 ****
  26 H8     -2.252210   1.236539   2.671358 H5    1   GAP  0.1526  0.0000 ****
  27 N7     -3.783280   0.494507   1.451181 NB    1   GAP -0.5554  0.0000 ****
  28 C6     -4.658517  -1.015694  -0.409182 CA    1   GAP  0.6677  0.0000 ****
  29 N6     -5.953672  -0.665697  -0.378225 N2    1   GAP -0.8735  0.0000 ****
  30 H61    -6.241331   0.111927   0.169797 H     1   GAP  0.4008  0.0000 ****
  31 H62    -6.551887  -0.995486  -1.100244 H     1   GAP  0.4008  0.0000 ****
  32 C5     -3.709216  -0.461543   0.457170 CB    1   GAP  0.0129  0.0000 ****
  33 C4     -2.424041  -0.934376   0.328921 CB    1   GAP  0.4418  0.0000 ****
  34 N3     -2.004789  -1.856143  -0.534156 NC    1   GAP -0.6782  0.0000 ****
  35 C2     -2.986448  -2.293220  -1.291757 CQ    1   GAP  0.5001  0.0000 ****
  36 H2     -2.725981  -3.040636  -2.019527 H5    1   GAP  0.0648  0.0000 ****
  37 N1     -4.262161  -1.935453  -1.282934 NC    1   GAP -0.7139  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3    19 1
    6     5     6 1
    7     5     7 1
    8     5     8 1
    9     8     9 1
   10     8    14 1
```

```
   11      9    10 1
   12      9    11 1
   13      9    12 1
   14     12    13 1
   15     14    15 1
   16     14    16 1
   17     16    17 1
   18     16    18 1
   19     16    24 1
   20     19    20 1
   21     19    21 1
   22     19    22 1
   23     22    23 1
   24     24    25 1
   25     24    33 1
   26     25    26 1
   27     25    27 1
   28     27    32 1
   29     28    29 1
   30     28    32 1
   31     28    37 1
   32     29    30 1
   33     29    31 1
   34     32    33 1
   35     33    34 1
   36     34    35 1
   37     35    36 1
   38     35    37 1
@<TRIPOS>SUBSTRUCTURE
      1 GAP              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

GNA

```
@<TRIPOS>MOLECULE
GNA
    39    40     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'    0.000000   0.000000  -0.000000 N3   1   GNA -0.0679  0.0000 ****
   2 H1'1  -0.372149   0.940540  -0.000000 H    1   GNA  0.2683  0.0000 ****
   3 H1'2  -0.372151  -0.470283  -0.814524 H    1   GNA  0.2683  0.0000 ****
   4 H1'3  -0.372144  -0.470264   0.814532 H    1   GNA  0.2683  0.0000 ****
   5 C2'    1.507376  -0.000000   0.000000 CT   1   GNA  0.3376  0.0000 ****
   6 H2'1   1.835302   0.530722  -0.882699 HP   1   GNA  0.0685  0.0000 ****
   7 C3'    2.093047  -1.418083  -0.000000 CT   1   GNA -0.0833  0.0000 ****
   8 H3'1   1.860282  -1.912064   0.930461 H1   1   GNA  0.0545  0.0000 ****
   9 H3'2   3.173169  -1.353284  -0.055978 H1   1   GNA  0.0545  0.0000 ****
  10 N4'    1.656411  -2.242214  -1.118563 N    1   GNA -0.0870  0.0000 ****
  11 C5'    2.335753  -2.027012  -2.393328 CT   1   GNA -0.1820  0.0000 ****
  12 H5'1   1.665262  -2.295744  -3.192661 H1   1   GNA  0.0902  0.0000 ****
  13 H5'2   2.576470  -0.980542  -2.482883 H1   1   GNA  0.0902  0.0000 ****
  14 C'     3.639722  -2.815302  -2.452241 C    1   GNA  0.6203  0.0000 ****
  15 O1'    4.682808  -2.314574  -2.115868 O    1   GNA -0.5801  0.0000 ****
```

```
  16 C7'    0.778763  -3.257425  -1.046602 C    1   GNA  0.5146  0.0000 ****
  17 O7'    0.614652  -4.043909  -1.952130 O    1   GNA -0.5620  0.0000 ****
  18 C8'   -0.071201  -3.454963   0.207519 CT   1   GNA -0.0347  0.0000 ****
  19 H8'1  -1.074053  -3.647828  -0.148221 H1   1   GNA  0.0860  0.0000 ****
  20 H8'2  -0.099131  -2.606384   0.865363 H1   1   GNA  0.0860  0.0000 ****
  21 C9'    2.020273   0.747366   1.225265 CT   1   GNA  0.0891  0.0000 ****
  22 H9'1   3.103980   0.801223   1.175501 H1   1   GNA  0.0461  0.0000 ****
  23 H9'2   1.755925   0.202261   2.130705 H1   1   GNA  0.0461  0.0000 ****
  24 O10'   1.448284   2.026892   1.234387 OH   1   GNA -0.6783  0.0000 ****
  25 H10'   1.807919   2.533213   1.949252 HO   1   GNA  0.4459  0.0000 ****
  26 N9     0.393781  -4.594065   0.972105 N*   1   GNA -0.1541  0.0000 ****
  27 C8     0.341518  -5.922617   0.619089 CK   1   GNA  0.1833  0.0000 ****
  28 H8    -0.051243  -6.215725  -0.332068 H5   1   GNA  0.1524  0.0000 ****
  29 N7     0.803974  -6.714993   1.512642 NB   1   GNA -0.5620  0.0000 ****
  30 C6     1.778116  -6.105908   3.790308 CA   1   GNA  0.6730  0.0000 ****
  31 N6     2.050565  -7.339154   4.243912 N2   1   GNA -0.8765  0.0000 ****
  32 H61    1.979390  -8.117315   3.629767 H    1   GNA  0.4021  0.0000 ****
  33 H62    2.578638  -7.431150   5.080936 H    1   GNA  0.4021  0.0000 ****
  34 C5     1.197929  -5.876982   2.537392 CB   1   GNA  0.0116  0.0000 ****
  35 C4     0.951607  -4.562632   2.215673 CB   1   GNA  0.4629  0.0000 ****
  36 N3     1.213693  -3.501901   2.974942 NC   1   GNA -0.7199  0.0000 ****
  37 C2     1.759251  -3.848084   4.120007 CQ   1   GNA  0.5327  0.0000 ****
  38 H2     2.004031  -3.043609   4.790101 H5   1   GNA  0.0597  0.0000 ****
  39 N1     2.052751  -5.061197   4.564679 NC   1   GNA -0.7265  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5    21 1
    8     7     8 1
    9     7     9 1
   10     7    10 1
   11    10    11 1
   12    10    16 1
   13    11    12 1
   14    11    13 1
   15    11    14 1
   16    14    15 1
   17    16    17 1
   18    16    18 1
   19    18    19 1
   20    18    20 1
   21    18    26 1
   22    21    22 1
   23    21    23 1
   24    21    24 1
   25    24    25 1
   26    26    27 1
   27    26    35 1
   28    27    28 1
   29    27    29 1
   30    29    34 1
   31    30    31 1
   32    30    34 1
   33    30    39 1
   34    31    32 1
   35    31    33 1
   36    34    35 1
   37    35    36 1
```

```
   38    36    37 1
   39    37    38 1
   40    37    39 1
@<TRIPOS>SUBSTRUCTURE
      1 GNA               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
C' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 C' 0 0 0 0 0
```

GCA

```
@<TRIPOS>MOLECULE
GCA
    38    39     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'   -3.092953   1.875939  -1.401830 N    1   GCA -0.3796  0.0000 ****
   2 H1'1  -3.669103   2.591833  -1.787227 H    1   GCA  0.2738  0.0000 ****
   3 C2'   -1.848722   1.624843  -2.113083 CT   1   GCA  0.0724  0.0000 ****
   4 H2'1  -1.808558   0.568982  -2.340904 H1   1   GCA  0.1339  0.0000 ****
   5 C3'   -0.614803   2.025996  -1.294248 CT   1   GCA -0.1382  0.0000 ****
   6 H3'1  -0.594430   3.096322  -1.160303 H1   1   GCA  0.0801  0.0000 ****
   7 H3'2   0.277515   1.765196  -1.850798 H1   1   GCA  0.0801  0.0000 ****
   8 N4'   -0.518931   1.365121   0.000000 N    1   GCA -0.0476  0.0000 ****
   9 C5'    0.000000  -0.000000   0.000000 CT   1   GCA -0.5030  0.0000 ****
  10 H5'1  -0.403557  -0.523804   0.850575 H1   1   GCA  0.1684  0.0000 ****
  11 H5'2  -0.331675  -0.494469  -0.898085 H1   1   GCA  0.1684  0.0000 ****
  12 C7'   -0.747400   1.942265   1.191975 C    1   GCA  0.5677  0.0000 ****
  13 O7'   -0.446174   1.416833   2.240142 O    1   GCA -0.5857  0.0000 ****
  14 C8'   -1.420571   3.311601   1.269709 CT   1   GCA -0.0382  0.0000 ****
  15 H8'1  -2.164701   3.226002   2.049710 H1   1   GCA  0.0709  0.0000 ****
  16 H8'2  -1.908549   3.620531   0.364135 H1   1   GCA  0.0709  0.0000 ****
  17 C9'   -1.843819   2.397817  -3.426620 CT   1   GCA  0.1197  0.0000 ****
  18 H9'1  -0.943017   2.145984  -3.978831 H1   1   GCA  0.0523  0.0000 ****
  19 H9'2  -1.823057   3.468702  -3.227584 H1   1   GCA  0.0523  0.0000 ****
  20 O10'  -2.994762   2.051758  -4.147740 OH   1   GCA -0.6738  0.0000 ****
  21 H10'  -2.976589   2.467427  -4.998398 HO   1   GCA  0.4408  0.0000 ****
  22 N9    -0.463618   4.337528   1.630357 N*   1   GCA -0.1392  0.0000 ****
  23 C8     0.192134   4.492598   2.829668 CK   1   GCA  0.1702  0.0000 ****
  24 H8     0.044541   3.789953   3.623002 H5   1   GCA  0.1557  0.0000 ****
  25 N7     0.962701   5.514703   2.870779 NB   1   GCA -0.5571  0.0000 ****
  26 C6     1.392859   7.224339   1.026923 CA   1   GCA  0.6665  0.0000 ****
  27 N6     2.245853   8.031004   1.676982 N2   1   GCA -0.8738  0.0000 ****
  28 H61    2.610992   7.754402   2.558991 H    1   GCA  0.4010  0.0000 ****
  29 H62    2.712648   8.741783   1.162367 H    1   GCA  0.4010  0.0000 ****
  30 C5     0.826780   6.091802   1.623359 CB   1   GCA  0.0181  0.0000 ****
  31 C4    -0.050894   5.373124   0.845616 CB   1   GCA  0.4347  0.0000 ****
  32 N3    -0.404461   5.650092  -0.406824 NC   1   GCA -0.6715  0.0000 ****
  33 C2     0.196789   6.731005  -0.852668 CQ   1   GCA  0.4972  0.0000 ****
  34 H2    -0.035659   7.019752  -1.861983 H5   1   GCA  0.0651  0.0000 ****
  35 N1     1.057721   7.518697  -0.225089 NC   1   GCA -0.7129  0.0000 ****
  36 C'     1.524863   0.000000  -0.000000 C    1   GCA  0.8129  0.0000 ****
  37 O1'    2.036542  -1.124430  -0.000000 O2   1   GCA -0.8266  0.0000 ****
  38 OXT    2.065633   1.108429  -0.000000 O2   1   GCA -0.8266  0.0000 ****
@<TRIPOS>BOND
```

```
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3    17 1
    6     5     6 1
    7     5     7 1
    8     5     8 1
    9     8     9 1
   10     8    12 1
   11     9    10 1
   12     9    11 1
   13     9    36 1
   14    12    13 1
   15    12    14 1
   16    14    15 1
   17    14    16 1
   18    14    22 1
   19    17    18 1
   20    17    19 1
   21    17    20 1
   22    20    21 1
   23    22    23 1
   24    22    31 1
   25    23    24 1
   26    23    25 1
   27    25    30 1
   28    26    27 1
   29    26    30 1
   30    26    35 1
   31    27    28 1
   32    27    29 1
   33    30    31 1
   34    31    32 1
   35    32    33 1
   36    33    34 1
   37    33    35 1
   38    36    37 1
   39    36    38 1
@<TRIPOS>SUBSTRUCTURE
      1 GCA               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

MAP

```
@<TRIPOS>MOLECULE
MAP
    49    50     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 C1     3.888841  -2.360831   2.319522 CT   1   MAP -0.1796  0.0000 ****
   2 H1     3.018350  -2.933619   2.619418 HC   1   MAP  0.0586  0.0000 ****
   3 H3     4.634525  -3.052384   1.938797 HC   1   MAP  0.0586  0.0000 ****
```

```
   4 H4      4.300237  -1.845023   3.174709 HC   1   MAP  0.0586  0.0000 ****
   5 C       3.558559  -1.348791   1.244087 C    1   MAP  0.5788  0.0000 ****
   6 O       4.051954  -0.256753   1.220008 O    1   MAP -0.5764  0.0000 ****
   7 N1'     2.658286  -1.774021   0.309197 N    1   MAP -0.3396  0.0000 ****
   8 H1'1    2.517951  -2.759747   0.268759 H    1   MAP  0.2738  0.0000 ****
   9 C2'     2.512277  -1.069824  -0.955646 CT   1   MAP  0.0565  0.0000 ****
  10 H2'1    3.410703  -0.490010  -1.113776 H1   1   MAP  0.1317  0.0000 ****
  11 C3'     1.288324  -0.144835  -0.973096 CT   1   MAP -0.1419  0.0000 ****
  12 H3'1    0.383739  -0.730991  -0.927254 H1   1   MAP  0.0511  0.0000 ****
  13 H3'2    1.264585   0.390541  -1.914799 H1   1   MAP  0.0511  0.0000 ****
  14 N4'     1.282584   0.854551   0.086253 N    1   MAP  0.0039  0.0000 ****
  15 C5'     2.106475   2.040867  -0.129854 CT   1   MAP  0.0403  0.0000 ****
  16 H5'1    2.401438   2.436687   0.827763 H1   1   MAP  0.0436  0.0000 ****
  17 H5'2    2.992408   1.751034  -0.670394 H1   1   MAP  0.0436  0.0000 ****
  18 C'      1.360971   3.072713  -0.969332 C    1   MAP  0.5508  0.0000 ****
  19 O1'     1.478995   3.099056  -2.168198 O    1   MAP -0.5659  0.0000 ****
  20 C7'     0.496122   0.835159   1.175840 C    1   MAP  0.4353  0.0000 ****
  21 O7'     0.374438   1.789833   1.910183 O    1   MAP -0.5726  0.0000 ****
  22 C8'    -0.284550  -0.426263   1.541418 CT   1   MAP -0.0170  0.0000 ****
  23 H8'1   -0.142787  -0.557902   2.605388 H1   1   MAP  0.0335  0.0000 ****
  24 H8'2    0.046395  -1.316442   1.039739 H1   1   MAP  0.0335  0.0000 ****
  25 C9'     2.387422  -2.077500  -2.092261 CT   1   MAP  0.1172  0.0000 ****
  26 H9'1    2.350846  -1.540092  -3.035476 H1   1   MAP  0.0482  0.0000 ****
  27 H9'2    1.458038  -2.637070  -1.992424 H1   1   MAP  0.0482  0.0000 ****
  28 O10'    3.491037  -2.940017  -2.041878 OH   1   MAP -0.6940  0.0000 ****
  29 H10'    3.469878  -3.530498  -2.781888 HO   1   MAP  0.4523  0.0000 ****
  30 N9     -1.696155  -0.258620   1.262813 N*   1   MAP -0.0523  0.0000 ****
  31 C8     -2.584459   0.581549   1.893238 CK   1   MAP  0.1404  0.0000 ****
  32 H8     -2.252210   1.236539   2.671358 H5   1   MAP  0.1674  0.0000 ****
  33 N7     -3.783280   0.494507   1.451181 NB   1   MAP -0.5545  0.0000 ****
  34 C6     -4.658517  -1.015694  -0.409182 CA   1   MAP  0.6713  0.0000 ****
  35 N6     -5.953672  -0.665697  -0.378225 N2   1   MAP -0.8862  0.0000 ****
  36 H61    -6.241331   0.111927   0.169797 H    1   MAP  0.4040  0.0000 ****
  37 H62    -6.551887  -0.995486  -1.100244 H    1   MAP  0.4040  0.0000 ****
  38 C5     -3.709216  -0.461543   0.457170 CB   1   MAP  0.0243  0.0000 ****
  39 C4     -2.424041  -0.934376   0.328921 CB   1   MAP  0.3730  0.0000 ****
  40 N3     -2.004789  -1.856143  -0.534156 NC   1   MAP -0.5797  0.0000 ****
  41 C2     -2.986448  -2.293220  -1.291757 CQ   1   MAP  0.4352  0.0000 ****
  42 H2     -2.725981  -3.040636  -2.019527 H5   1   MAP  0.0749  0.0000 ****
  43 N1     -4.262161  -1.935453  -1.282934 NC   1   MAP -0.6954  0.0000 ****
  44 N       0.570964   3.913920  -0.279606 N    1   MAP -0.4442  0.0000 ****
  45 H       0.403803   3.700114   0.679105 H    1   MAP  0.3013  0.0000 ****
  46 CH3    -0.293905   4.857551  -0.952089 CT   1   MAP -0.1674  0.0000 ****
  47 HH31   -1.128685   4.366843  -1.442432 H1   1   MAP  0.1005  0.0000 ****
  48 HH32   -0.678207   5.557668  -0.222257 H1   1   MAP  0.1005  0.0000 ****
  49 HH33    0.267223   5.400512  -1.698807 H1   1   MAP  0.1005  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    25 1
   12    11    12 1
   13    11    13 1
   14    11    14 1
   15    14    15 1
```

```
   16    14    20 1
   17    15    16 1
   18    15    17 1
   19    15    18 1
   20    18    19 1
   21    18    44 1
   22    20    21 1
   23    20    22 1
   24    22    23 1
   25    22    24 1
   26    22    30 1
   27    25    26 1
   28    25    27 1
   29    25    28 1
   30    28    29 1
   31    30    31 1
   32    30    39 1
   33    31    32 1
   34    31    33 1
   35    33    38 1
   36    34    35 1
   37    34    38 1
   38    34    43 1
   39    35    36 1
   40    35    37 1
   41    38    39 1
   42    39    40 1
   43    40    41 1
   44    41    42 1
   45    41    43 1
   46    44    45 1
   47    44    46 1
   48    46    47 1
   49    46    48 1
   50    46    49 1
@<TRIPOS>SUBSTRUCTURE
      1 MAP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 0 0 0 0 0
```

GTP

```
@<TRIPOS>MOLECULE
GTP
    37    37     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'    -3.098596  -1.629931  -0.637122 N    1   GTP -0.3871  0.0000 ****
   2 H1'1   -3.847978  -1.763996  -1.277424 H    1   GTP  0.2784  0.0000 ****
   3 C2'    -3.349478  -0.638782   0.397211 CT   1   GTP  0.0528  0.0000 ****
   4 H2'1   -4.055355   0.056271  -0.037132 H1   1   GTP  0.1363  0.0000 ****
   5 C3'    -2.104015   0.144775   0.842890 CT   1   GTP -0.1234  0.0000 ****
   6 H3'1   -1.537891  -0.463569   1.523138 H1   1   GTP  0.0809  0.0000 ****
   7 H3'2   -2.454305   1.016378   1.386112 H1   1   GTP  0.0809  0.0000 ****
```

```
   8 N4'    -1.239246    0.566306   -0.248807 N     1   GTP -0.1034   0.0000 ****
   9 C5'    -1.823333    1.401828   -1.291874 CT    1   GTP -0.1424   0.0000 ****
  10 H5'1   -1.199862    1.338948   -2.168107 H1    1   GTP  0.0802   0.0000 ****
  11 H5'2   -2.806879    1.036260   -1.535584 H1    1   GTP  0.0802   0.0000 ****
  12 C'     -1.973998    2.841207   -0.811414 C     1   GTP  0.6199   0.0000 ****
  13 O1'    -3.011624    3.225255   -0.334363 O     1   GTP -0.5806   0.0000 ****
  14 C7'     0.113600    0.464041   -0.230592 C     1   GTP  0.5312   0.0000 ****
  15 O7'     0.800996    0.961161   -1.086483 O     1   GTP -0.5649   0.0000 ****
  16 C8'     0.766505   -0.312212    0.914680 CT    1   GTP -0.0062   0.0000 ****
  17 H8'1    0.372818   -1.315880    0.955700 H1    1   GTP  0.0699   0.0000 ****
  18 H8'2    0.557977    0.180680    1.851259 H1    1   GTP  0.0699   0.0000 ****
  19 C9'    -4.041595   -1.251223    1.610554 CT    1   GTP  0.1284   0.0000 ****
  20 H9'1   -4.210855   -0.477759    2.356256 H1    1   GTP  0.0497   0.0000 ****
  21 H9'2   -3.411778   -2.018791    2.048887 H1    1   GTP  0.0497   0.0000 ****
  22 O10'   -5.257488   -1.789157    1.163294 OH    1   GTP -0.6690   0.0000 ****
  23 H10'   -5.716378   -2.193252    1.886156 HO    1   GTP  0.4382   0.0000 ****
  24 N1      2.199095   -0.369228    0.746125 N*    1   GTP -0.0856   0.0000 ****
  25 C2      2.931817    0.717666    1.150562 C     1   GTP  0.5546   0.0000 ****
  26 O2      2.457802    1.640760    1.749217 O     1   GTP -0.5620   0.0000 ****
  27 N3      4.262929    0.649281    0.840538 NA    1   GTP -0.4513   0.0000 ****
  28 H3      4.813971    1.424838    1.142231 H     1   GTP  0.3319   0.0000 ****
  29 C4      4.918363   -0.329958    0.112682 C     1   GTP  0.5990   0.0000 ****
  30 O4      6.092091   -0.245355   -0.101713 O     1   GTP -0.5614   0.0000 ****
  31 C5      4.055887   -1.425692   -0.339984 CM    1   GTP  0.0097   0.0000 ****
  32 C7      4.680287   -2.529774   -1.145733 CT    1   GTP -0.2784   0.0000 ****
  33 H73     5.127731   -2.138991   -2.052623 HC    1   GTP  0.0921   0.0000 ****
  34 H72     5.468181   -3.019677   -0.584610 HC    1   GTP  0.0921   0.0000 ****
  35 H71     3.939973   -3.272938   -1.419789 HC    1   GTP  0.0921   0.0000 ****
  36 C6      2.767463   -1.371598   -0.014734 CM    1   GTP -0.2376   0.0000 ****
  37 H6      2.077851   -2.136183   -0.317510 H4    1   GTP  0.2348   0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3    19 1
    6     5     6 1
    7     5     7 1
    8     5     8 1
    9     8     9 1
   10     8    14 1
   11     9    10 1
   12     9    11 1
   13     9    12 1
   14    12    13 1
   15    14    15 1
   16    14    16 1
   17    16    17 1
   18    16    18 1
   19    16    24 1
   20    19    20 1
   21    19    21 1
   22    19    22 1
   23    22    23 1
   24    24    25 1
   25    24    36 1
   26    25    26 1
   27    25    27 1
   28    27    28 1
   29    27    29 1
   30    29    30 1
   31    29    31 1
```

```
   32    31    32 1
   33    31    36 1
   34    32    33 1
   35    32    34 1
   36    32    35 1
   37    36    37 1
@<TRIPOS>SUBSTRUCTURE
      1 GTP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

GNT

```
@<TRIPOS>MOLECULE
GNT
    39    39     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'     0.000000   0.000000  -0.000000 N3   1   GNT -0.0679  0.0000 ****
   2 H1'1   -0.372149   0.940540  -0.000000 H    1   GNT  0.2683  0.0000 ****
   3 H1'2   -0.372151  -0.470283  -0.814524 H    1   GNT  0.2683  0.0000 ****
   4 H1'3   -0.372144  -0.470264   0.814532 H    1   GNT  0.2683  0.0000 ****
   5 C2'     1.507376   0.000000  -0.000000 CT   1   GNT  0.3376  0.0000 ****
   6 H2'1    1.793921   0.507528  -0.911221 HP   1   GNT  0.0685  0.0000 ****
   7 C3'     2.143489  -1.399689   0.000000 CT   1   GNT -0.0833  0.0000 ****
   8 H3'1    2.115032  -1.792411   0.999137 H1   1   GNT  0.0545  0.0000 ****
   9 H3'2    3.184252  -1.275903  -0.281027 H1   1   GNT  0.0545  0.0000 ****
  10 N4'     1.505180  -2.347630  -0.900712 N    1   GNT -0.0870  0.0000 ****
  11 C5'     1.433524  -1.996072  -2.314401 CT   1   GNT -0.1820  0.0000 ****
  12 H5'1    0.659949  -2.588207  -2.774198 H1   1   GNT  0.0902  0.0000 ****
  13 H5'2    1.180729  -0.953538  -2.412455 H1   1   GNT  0.0902  0.0000 ****
  14 C'      2.782154  -2.207860  -2.993884 C    1   GNT  0.6203  0.0000 ****
  15 O1'     3.562153  -1.296973  -3.110606 O    1   GNT -0.5801  0.0000 ****
  16 C7'     1.215070  -3.631808  -0.572529 C    1   GNT  0.5146  0.0000 ****
  17 O7'     0.826576  -4.425786  -1.391590 O    1   GNT -0.5620  0.0000 ****
  18 C8'     1.387935  -4.064328   0.884710 CT   1   GNT  0.0068  0.0000 ****
  19 H8'1    0.801018  -3.431954   1.532518 H1   1   GNT  0.0696  0.0000 ****
  20 H8'2    2.425903  -3.981197   1.166344 H1   1   GNT  0.0696  0.0000 ****
  21 C9'     2.072312   0.829105   1.148800 CT   1   GNT  0.0891  0.0000 ****
  22 H9'1    3.158967   0.803132   1.110399 H1   1   GNT  0.0461  0.0000 ****
  23 H9'2    1.752307   0.413373   2.098922 H1   1   GNT  0.0461  0.0000 ****
  24 O10'    1.597366   2.138730   0.984115 OH   1   GNT -0.6783  0.0000 ****
  25 H10'    1.915230   2.687563   1.687072 HO   1   GNT  0.4459  0.0000 ****
  26 N1      0.982076  -5.437928   1.064923 N*   1   GNT -0.0851  0.0000 ****
  27 C2      1.884033  -6.417234   0.734433 C    1   GNT  0.5580  0.0000 ****
  28 O2      3.020652  -6.186268   0.434550 O    1   GNT -0.5585  0.0000 ****
  29 N3      1.387319  -7.690419   0.804634 NA   1   GNT -0.4595  0.0000 ****
  30 H3      2.035369  -8.416452   0.583134 H    1   GNT  0.3331  0.0000 ****
  31 C4      0.089255  -8.083617   1.085166 C    1   GNT  0.6027  0.0000 ****
  32 O4     -0.208038  -9.242222   1.088865 O    1   GNT -0.5614  0.0000 ****
  33 C5     -0.830644  -6.976704   1.364329 CM   1   GNT  0.0100  0.0000 ****
  34 C7     -2.263834  -7.309006   1.670345 CT   1   GNT -0.2726  0.0000 ****
  35 H73    -2.719675  -7.844312   0.845103 HC   1   GNT  0.0912  0.0000 ****
  36 H72    -2.334551  -7.946642   2.544451 HC   1   GNT  0.0912  0.0000 ****
```

```
  37 H71    -2.837502  -6.407686   1.854711 HC   1   GNT  0.0912  0.0000 ****
  38 C6     -0.340205  -5.741202   1.322393 CM   1   GNT -0.2382  0.0000 ****
  39 H6     -0.957644  -4.883753   1.510486 H4   1   GNT  0.2297  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5    21 1
    8     7     8 1
    9     7     9 1
   10     7    10 1
   11    10    11 1
   12    10    16 1
   13    11    12 1
   14    11    13 1
   15    11    14 1
   16    14    15 1
   17    16    17 1
   18    16    18 1
   19    18    19 1
   20    18    20 1
   21    18    26 1
   22    21    22 1
   23    21    23 1
   24    21    24 1
   25    24    25 1
   26    26    27 1
   27    26    38 1
   28    27    28 1
   29    27    29 1
   30    29    30 1
   31    29    31 1
   32    31    32 1
   33    31    33 1
   34    33    34 1
   35    33    38 1
   36    34    35 1
   37    34    36 1
   38    34    37 1
   39    38    39 1
@<TRIPOS>SUBSTRUCTURE
      1 GNT               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
C' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 C' 0 0 0 0 0
```

GCT

```
@<TRIPOS>MOLECULE
GCT
    38    38     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'    -2.529414   0.851945  -2.030472 N    1   GCT -0.3796  0.0000 ****
   2 H1'1   -2.783661   0.026734  -2.524353 H    1   GCT  0.2738  0.0000 ****
   3 C2'    -1.243179   1.416752  -2.406977 CT   1   GCT  0.0724  0.0000 ****
   4 H2'1   -0.654219   0.579836  -2.757283 H1   1   GCT  0.1339  0.0000 ****
   5 C3'    -0.486204   2.098481  -1.255454 CT   1   GCT -0.1382  0.0000 ****
   6 H3'1   -0.902032   3.076247  -1.099325 H1   1   GCT  0.0801  0.0000 ****
   7 H3'2    0.542276   2.220401  -1.579260 H1   1   GCT  0.0801  0.0000 ****
   8 N4'    -0.517724   1.363527   0.000000 N    1   GCT -0.0476  0.0000 ****
   9 C5'     0.000000   0.000000   0.000000 CT   1   GCT -0.5030  0.0000 ****
  10 H5'1   -0.397032  -0.515445   0.858566 H1   1   GCT  0.1684  0.0000 ****
  11 H5'2   -0.324673  -0.507018  -0.893262 H1   1   GCT  0.1684  0.0000 ****
  12 C7'    -0.742178   1.934413   1.210244 C    1   GCT  0.5677  0.0000 ****
  13 O7'    -0.610527   1.319502   2.238224 O    1   GCT -0.5857  0.0000 ****
  14 C8'    -1.178554   3.400173   1.250564 CT   1   GCT -0.0109  0.0000 ****
  15 H8'1   -2.074105   3.537887   0.664862 H1   1   GCT  0.0693  0.0000 ****
  16 H8'2   -0.397613   4.021800   0.841518 H1   1   GCT  0.0693  0.0000 ****
  17 C9'    -1.370592   2.375357  -3.586448 CT   1   GCT  0.1197  0.0000 ****
  18 H9'1   -0.388837   2.772112  -3.834861 H1   1   GCT  0.0523  0.0000 ****
  19 H9'2   -2.019227   3.204519  -3.322387 H1   1   GCT  0.0523  0.0000 ****
  20 O10'   -1.899140   1.641307  -4.658651 OH   1   GCT -0.6738  0.0000 ****
  21 H10'   -2.007475   2.204190  -5.412196 HO   1   GCT  0.4408  0.0000 ****
  22 N1     -1.427023   3.825499   2.607522 N*   1   GCT -0.0880  0.0000 ****
  23 C2     -0.346060   4.193181   3.367863 C    1   GCT  0.5577  0.0000 ****
  24 O2      0.760713   4.302682   2.922616 O    1   GCT -0.5631  0.0000 ****
  25 N3     -0.639807   4.456591   4.678198 NA   1   GCT -0.4533  0.0000 ****
  26 H3      0.132861   4.741569   5.242048 H    1   GCT  0.3320  0.0000 ****
  27 C4     -1.858207   4.317986   5.322066 C    1   GCT  0.5997  0.0000 ****
  28 O4     -1.961867   4.565562   6.487708 O    1   GCT -0.5615  0.0000 ****
  29 C5     -2.949890   3.859175   4.457702 CM   1   GCT  0.0077  0.0000 ****
  30 C7     -4.307608   3.671292   5.073663 CT   1   GCT -0.2831  0.0000 ****
  31 H73    -4.268690   2.944529   5.877199 HC   1   GCT  0.0935  0.0000 ****
  32 H72    -4.671083   4.600230   5.498634 HC   1   GCT  0.0935  0.0000 ****
  33 H71    -5.022291   3.328544   4.333922 HC   1   GCT  0.0935  0.0000 ****
  34 C6     -2.669053   3.629553   3.178189 CM   1   GCT -0.2275  0.0000 ****
  35 H6     -3.418014   3.286665   2.490210 H4   1   GCT  0.2296  0.0000 ****
  36 C'      1.524911  -0.000000   0.000000 C    1   GCT  0.8129  0.0000 ****
  37 O1'     2.036591  -1.124430  -0.000000 O2   1   GCT -0.8266  0.0000 ****
  38 OXT     2.065682   1.108429   0.000000 O2   1   GCT -0.8266  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3    17 1
    6     5     6 1
    7     5     7 1
    8     5     8 1
    9     8     9 1
   10     8    12 1
   11     9    10 1
   12     9    11 1
   13     9    36 1
```

```
   14    12    13 1
   15    12    14 1
   16    14    15 1
   17    14    16 1
   18    14    22 1
   19    17    18 1
   20    17    19 1
   21    17    20 1
   22    20    21 1
   23    22    23 1
   24    22    34 1
   25    23    24 1
   26    23    25 1
   27    25    26 1
   28    25    27 1
   29    27    28 1
   30    27    29 1
   31    29    30 1
   32    29    34 1
   33    30    31 1
   34    30    32 1
   35    30    33 1
   36    34    35 1
   37    36    37 1
   38    36    38 1
@<TRIPOS>SUBSTRUCTURE
      1 GCT               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

MTP

```
@<TRIPOS>MOLECULE
MTP
    49    49     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 C1    -2.294517  -3.685951  -1.627163 CT   1   MTP -0.2117  0.0000 ****
   2 H1    -3.037204  -4.420831  -1.331685 HC   1   MTP  0.0663  0.0000 ****
   3 H2    -2.588060  -3.270281  -2.583803 HC   1   MTP  0.0663  0.0000 ****
   4 H4    -1.340176  -4.183090  -1.722361 HC   1   MTP  0.0663  0.0000 ****
   5 C     -2.194082  -2.628541  -0.549692 C    1   MTP  0.6963  0.0000 ****
   6 O     -1.363114  -2.692738   0.322836 O    1   MTP -0.5971  0.0000 ****
   7 N1'   -3.098596  -1.629931  -0.637122 N    1   MTP -0.5775  0.0000 ****
   8 H1'1  -3.847978  -1.763996  -1.277425 H    1   MTP  0.3483  0.0000 ****
   9 C2'   -3.349478  -0.638782   0.397210 CT   1   MTP  0.0634  0.0000 ****
  10 H2'1  -4.055355   0.056272  -0.037133 H1   1   MTP  0.1441  0.0000 ****
  11 C3'   -2.104015   0.144775   0.842889 CT   1   MTP -0.1693  0.0000 ****
  12 H3'1  -1.537891  -0.463569   1.523137 H1   1   MTP  0.0998  0.0000 ****
  13 H3'2  -2.454304   1.016378   1.386112 H1   1   MTP  0.0998  0.0000 ****
  14 N4'   -1.239246   0.566306  -0.248807 N    1   MTP -0.1318  0.0000 ****
  15 C5'   -1.823333   1.401828  -1.291874 CT   1   MTP -0.1445  0.0000 ****
  16 H5'1  -1.199862   1.338949  -2.168107 H1   1   MTP  0.0645  0.0000 ****
  17 H5'2  -2.806879   1.036260  -1.535584 H1   1   MTP  0.0645  0.0000 ****
```

```
  18 C'    -1.973998   2.841207  -0.811415 C    1   MTP  0.6951  0.0000 ****
  19 O1'   -3.011624   3.225254  -0.334363 O    1   MTP -0.6015  0.0000 ****
  20 C7'    0.113600   0.464041  -0.230592 C    1   MTP  0.5167  0.0000 ****
  21 O7'    0.800996   0.961161  -1.086483 O    1   MTP -0.5280  0.0000 ****
  22 C8'    0.766505  -0.312212   0.914680 CT   1   MTP -0.0520  0.0000 ****
  23 H8'1   0.372818  -1.315880   0.955700 H1   1   MTP  0.1003  0.0000 ****
  24 H8'2   0.557977   0.180680   1.851259 H1   1   MTP  0.1003  0.0000 ****
  25 C9'   -4.041595  -1.251222   1.610553 CT   1   MTP  0.1301  0.0000 ****
  26 H9'1  -4.210855  -0.477758   2.356256 H1   1   MTP  0.0590  0.0000 ****
  27 H9'2  -3.411778  -2.018791   2.048887 H1   1   MTP  0.0590  0.0000 ****
  28 O10'  -5.257488  -1.789156   1.163293 OH   1   MTP -0.6713  0.0000 ****
  29 H10'  -5.716379  -2.193251   1.886156 HO   1   MTP  0.4373  0.0000 ****
  30 N1     2.199095  -0.369228   0.746125 N*   1   MTP -0.0755  0.0000 ****
  31 C2     2.931817   0.717666   1.150562 C    1   MTP  0.5342  0.0000 ****
  32 O2     2.457803   1.640760   1.749217 O    1   MTP -0.5545  0.0000 ****
  33 N3     4.262930   0.649281   0.840538 NA   1   MTP -0.4337  0.0000 ****
  34 H3     4.813971   1.424838   1.142231 H    1   MTP  0.3287  0.0000 ****
  35 C4     4.918363  -0.329957   0.112682 C    1   MTP  0.5866  0.0000 ****
  36 O4     6.092091  -0.245355  -0.101714 O    1   MTP -0.5593  0.0000 ****
  37 C5     4.055888  -1.425692  -0.339984 CM   1   MTP  0.0243  0.0000 ****
  38 C7     4.680287  -2.529774  -1.145733 CT   1   MTP -0.2381  0.0000 ****
  39 H73    5.127731  -2.138991  -2.052623 HC   1   MTP  0.0821  0.0000 ****
  40 H72    5.468182  -3.019676  -0.584610 HC   1   MTP  0.0821  0.0000 ****
  41 H71    3.939973  -3.272938  -1.419789 HC   1   MTP  0.0821  0.0000 ****
  42 C6     2.767463  -1.371598  -0.014734 CM   1   MTP -0.2703  0.0000 ****
  43 H6     2.077851  -2.136183  -0.317510 H4   1   MTP  0.2387  0.0000 ****
  44 N     -0.876798   3.606040  -0.936233 N    1   MTP -0.5053  0.0000 ****
  45 H     -0.026485   3.156773  -1.197597 H    1   MTP  0.3271  0.0000 ****
  46 CH3   -0.819546   4.943626  -0.386512 CT   1   MTP -0.0976  0.0000 ****
  47 HH31  -0.760908   4.933562   0.696658 H1   1   MTP  0.0853  0.0000 ****
  48 HH32   0.055735   5.442353  -0.779939 H1   1   MTP  0.0853  0.0000 ****
  49 HH33  -1.701278   5.495998  -0.677284 H1   1   MTP  0.0853  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    25 1
   12    11    12 1
   13    11    13 1
   14    11    14 1
   15    14    15 1
   16    14    20 1
   17    15    16 1
   18    15    17 1
   19    15    18 1
   20    18    19 1
   21    18    44 1
   22    20    21 1
   23    20    22 1
   24    22    23 1
   25    22    24 1
   26    22    30 1
   27    25    26 1
   28    25    27 1
   29    25    28 1
```

```
   30    28    29 1
   31    30    31 1
   32    30    42 1
   33    31    32 1
   34    31    33 1
   35    33    34 1
   36    33    35 1
   37    35    36 1
   38    35    37 1
   39    37    38 1
   40    37    42 1
   41    38    39 1
   42    38    40 1
   43    38    41 1
   44    42    43 1
   45    44    45 1
   46    44    46 1
   47    46    47 1
   48    46    48 1
   49    46    49 1
@<TRIPOS>SUBSTRUCTURE
      1 MTP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 0 0 0 0 0
```

GCP

```
@<TRIPOS>MOLECULE
GCP
    35    35     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'     2.717500  -1.457508   0.194461 N     1   GCP -0.3871  0.0000 ****
   2 H1'1    2.788920  -2.419137  -0.055916 H     1   GCP  0.2784  0.0000 ****
   3 C2'     2.630417  -0.523116  -0.917345 CT    1   GCP  0.0528  0.0000 ****
   4 H2'1    3.388500   0.232783  -0.767886 H1    1   GCP  0.1363  0.0000 ****
   5 C3'     1.248910   0.134907  -1.024890 CT    1   GCP -0.1234  0.0000 ****
   6 H3'1    0.504790  -0.599531  -1.287396 H1    1   GCP  0.0809  0.0000 ****
   7 H3'2    1.269210   0.861886  -1.828337 H1    1   GCP  0.0809  0.0000 ****
   8 N4'     0.830830   0.844382   0.178566 N     1   GCP -0.1034  0.0000 ****
   9 C5'     1.388083   2.179326   0.377348 CT    1   GCP -0.1424  0.0000 ****
  10 H5'1    1.407954   2.392945   1.433406 H1    1   GCP  0.0802  0.0000 ****
  11 H5'2    2.399076   2.188070   0.005038 H1    1   GCP  0.0802  0.0000 ****
  12 C'      0.595847   3.228539  -0.395664 C     1   GCP  0.6199  0.0000 ****
  13 O1'     0.943714   3.590385  -1.490504 O     1   GCP -0.5806  0.0000 ****
  14 C7'    -0.107239   0.425715   1.042028 C     1   GCP  0.5312  0.0000 ****
  15 O7'    -0.583920   1.152502   1.889173 O     1   GCP -0.5649  0.0000 ****
  16 C8'    -0.596562  -1.021874   1.014113 CT    1   GCP -0.1542  0.0000 ****
  17 H8'1   -0.531795  -1.362822   2.039178 H1    1   GCP  0.0926  0.0000 ****
  18 H8'2   -0.016645  -1.675714   0.396835 H1    1   GCP  0.0926  0.0000 ****
  19 C9'     2.930310  -1.251338  -2.222117 CT    1   GCP  0.1284  0.0000 ****
  20 H9'1    2.924591  -0.531221  -3.035294 H1    1   GCP  0.0497  0.0000 ****
  21 H9'2    2.152720  -1.986351  -2.424248 H1    1   GCP  0.0497  0.0000 ****
  22 O10'    4.181628  -1.873440  -2.105175 OH    1   GCP -0.6690  0.0000 ****
```

```
  23 H10'   4.406673  -2.291285  -2.924595 HO   1   GCP  0.4382  0.0000 ****
  24 N1    -1.984179  -1.106290   0.576026 N*   1   GCP -0.0388  0.0000 ****
  25 C2    -2.267068  -1.655975  -0.689648 C    1   GCP  0.7490  0.0000 ****
  26 O2    -1.351038  -2.022043  -1.380657 O    1   GCP -0.5921  0.0000 ****
  27 N3    -3.567503  -1.740644  -1.060912 NC   1   GCP -0.7700  0.0000 ****
  28 C4    -4.512582  -1.323108  -0.274271 CA   1   GCP  0.8543  0.0000 ****
  29 N4    -5.780838  -1.459083  -0.697812 N2   1   GCP -0.9738  0.0000 ****
  30 H41   -6.533134  -1.019636  -0.222212 H    1   GCP  0.4235  0.0000 ****
  31 H42   -5.926363  -1.766910  -1.632749 H    1   GCP  0.4235  0.0000 ****
  32 C5    -4.265886  -0.749157   1.019776 CM   1   GCP -0.5173  0.0000 ****
  33 H5    -5.058468  -0.412818   1.657783 HA   1   GCP  0.2029  0.0000 ****
  34 C6    -2.980078  -0.667278   1.385305 CM   1   GCP -0.0185  0.0000 ****
  35 H6    -2.667692  -0.247377   2.320001 H4   1   GCP  0.1899  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3    19 1
    6     5     6 1
    7     5     7 1
    8     5     8 1
    9     8     9 1
   10     8    14 1
   11     9    10 1
   12     9    11 1
   13     9    12 1
   14    12    13 1
   15    14    15 1
   16    14    16 1
   17    16    17 1
   18    16    18 1
   19    16    24 1
   20    19    20 1
   21    19    21 1
   22    19    22 1
   23    22    23 1
   24    24    25 1
   25    24    34 1
   26    25    26 1
   27    25    27 1
   28    27    28 1
   29    28    29 1
   30    28    32 1
   31    29    30 1
   32    29    31 1
   33    32    33 1
   34    32    34 1
   35    34    35 1
@<TRIPOS>SUBSTRUCTURE
      1 GCP              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

GNC

```
@<TRIPOS>MOLECULE
GNC
    37    37     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'     0.000000   0.000000  -0.000000 N3   1   GNC -0.0679  0.0000 ****
   2 H1'1   -0.372149   0.940540  -0.000000 H    1   GNC  0.2683  0.0000 ****
   3 H1'2   -0.372151  -0.470283  -0.814524 H    1   GNC  0.2683  0.0000 ****
   4 H1'3   -0.372144  -0.470264   0.814532 H    1   GNC  0.2683  0.0000 ****
   5 C2'     1.507376  -0.000000   0.000000 CT   1   GNC  0.3376  0.0000 ****
   6 H2'1    1.833250   0.534506  -0.881204 HP   1   GNC  0.0685  0.0000 ****
   7 C3'     2.094850  -1.417038   0.000000 CT   1   GNC -0.0833  0.0000 ****
   8 H3'1    1.868310  -1.915294   0.928678 H1   1   GNC  0.0545  0.0000 ****
   9 H3'2    3.174494  -1.348210  -0.063790 H1   1   GNC  0.0545  0.0000 ****
  10 N4'     1.655874  -2.244745  -1.117431 N    1   GNC -0.0870  0.0000 ****
  11 C5'     2.327958  -2.027648  -2.395434 CT   1   GNC -0.1820  0.0000 ****
  12 H5'1    1.656952  -2.305509  -3.191569 H1   1   GNC  0.0902  0.0000 ****
  13 H5'2    2.557570  -0.979129  -2.488622 H1   1   GNC  0.0902  0.0000 ****
  14 C'      3.639927  -2.801990  -2.467311 C    1   GNC  0.6203  0.0000 ****
  15 O1'     4.688139  -2.279399  -2.186654 O    1   GNC -0.5801  0.0000 ****
  16 C7'     0.783308  -3.261095  -1.036904 C    1   GNC  0.5146  0.0000 ****
  17 O7'     0.631240  -4.062069  -1.936023 O    1   GNC -0.5620  0.0000 ****
  18 C8'    -0.095758  -3.432499   0.201480 CT   1   GNC -0.1478  0.0000 ****
  19 H8'1   -1.101927  -3.550372  -0.179215 H1   1   GNC  0.1032  0.0000 ****
  20 H8'2   -0.078678  -2.603269   0.877445 H1   1   GNC  0.1032  0.0000 ****
  21 C9'     2.018809   0.743539   1.228112 CT   1   GNC  0.0891  0.0000 ****
  22 H9'1    3.103039   0.791350   1.183189 H1   1   GNC  0.0461  0.0000 ****
  23 H9'2    1.747766   0.199052   2.131337 H1   1   GNC  0.0461  0.0000 ****
  24 O10'    1.455016   2.027503   1.235743 OH   1   GNC -0.6783  0.0000 ****
  25 H10'    1.799371   2.521511   1.966541 HO   1   GNC  0.4459  0.0000 ****
  26 N1      0.267856  -4.628625   0.950876 N*   1   GNC -0.0588  0.0000 ****
  27 C2      0.898955  -4.483583   2.201791 C    1   GNC  0.7677  0.0000 ****
  28 O2      1.137076  -3.374256   2.605897 O    1   GNC -0.6116  0.0000 ****
  29 N3      1.206124  -5.612925   2.884705 NC   1   GNC -0.7639  0.0000 ****
  30 C4      0.929717  -6.783087   2.394161 CA   1   GNC  0.8431  0.0000 ****
  31 N4      1.241958  -7.859097   3.136494 N2   1   GNC -0.9719  0.0000 ****
  32 H41     1.205773  -8.775500   2.756818 H    1   GNC  0.4235  0.0000 ****
  33 H42     1.767425  -7.711138   3.968380 H    1   GNC  0.4235  0.0000 ****
  34 C5      0.294687  -6.974159   1.119381 CM   1   GNC -0.5099  0.0000 ****
  35 H5      0.070586  -7.947538   0.731212 HA   1   GNC  0.2007  0.0000 ****
  36 C6     -0.009016  -5.856781   0.446441 CM   1   GNC -0.0133  0.0000 ****
  37 H6     -0.472308  -5.868348  -0.519366 H4   1   GNC  0.1904  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5    21 1
    8     7     8 1
    9     7     9 1
   10     7    10 1
   11    10    11 1
   12    10    16 1
   13    11    12 1
   14    11    13 1
```

```
   15    11    14 1
   16    14    15 1
   17    16    17 1
   18    16    18 1
   19    18    19 1
   20    18    20 1
   21    18    26 1
   22    21    22 1
   23    21    23 1
   24    21    24 1
   25    24    25 1
   26    26    27 1
   27    26    36 1
   28    27    28 1
   29    27    29 1
   30    29    30 1
   31    30    31 1
   32    30    34 1
   33    31    32 1
   34    31    33 1
   35    34    35 1
   36    34    36 1
   37    36    37 1
@<TRIPOS>SUBSTRUCTURE
      1 GNC               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
C' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 C' 0 0 0 0 0
```

GCC

```
@<TRIPOS>MOLECULE
GCC
    36    36     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'   -3.099827   1.846246  -1.417676 N    1   GCC -0.3796  0.0000 ****
   2 H1'1  -3.671574   2.574750  -1.784973 H    1   GCC  0.2738  0.0000 ****
   3 C2'   -1.848261   1.613523  -2.122096 CT   1   GCC  0.0724  0.0000 ****
   4 H2'1  -1.797784   0.559705  -2.357321 H1   1   GCC  0.1339  0.0000 ****
   5 C3'   -0.623438   2.024668  -1.295119 CT   1   GCC -0.1382  0.0000 ****
   6 H3'1  -0.609107   3.093376  -1.154832 H1   1   GCC  0.0801  0.0000 ****
   7 H3'2   0.273367   1.768719  -1.847087 H1   1   GCC  0.0801  0.0000 ****
   8 N4'   -0.528153   1.361314   0.000000 N    1   GCC -0.0476  0.0000 ****
   9 C5'    0.000000  -0.000000   0.000000 CT   1   GCC -0.5030  0.0000 ****
  10 H5'1  -0.398625  -0.526479   0.851590 H1   1   GCC  0.1684  0.0000 ****
  11 H5'2  -0.330444  -0.496262  -0.897402 H1   1   GCC  0.1684  0.0000 ****
  12 C7'   -0.766535   1.937298   1.188376 C    1   GCC  0.5677  0.0000 ****
  13 O7'   -0.448304   1.419482   2.238945 O    1   GCC -0.5857  0.0000 ****
  14 C8'   -1.494072   3.278962   1.268118 CT   1   GCC -0.1640  0.0000 ****
  15 H8'1  -2.281823   3.131485   1.995372 H1   1   GCC  0.0932  0.0000 ****
  16 H8'2  -1.932254   3.602882   0.347295 H1   1   GCC  0.0932  0.0000 ****
  17 C9'   -1.843699   2.397174  -3.429208 CT   1   GCC  0.1197  0.0000 ****
  18 H9'1  -0.933167   2.165351  -3.974230 H1   1   GCC  0.0523  0.0000 ****
  19 H9'2  -1.842978   3.466050  -3.221242 H1   1   GCC  0.0523  0.0000 ****
  20 O10'  -2.980946   2.036702  -4.166237 OH   1   GCC -0.6738  0.0000 ****
```

```
  21 H10'  -2.969980   2.478239  -5.003861 HO   1   GCC  0.4408  0.0000 ****
  22 N1    -0.609300   4.337001   1.739561 N*   1   GCC -0.0389  0.0000 ****
  23 C2    -0.199002   5.335840   0.835079 C    1   GCC  0.7445  0.0000 ****
  24 O2    -0.576436   5.274309  -0.307010 O    1   GCC -0.5879  0.0000 ****
  25 N3     0.606440   6.317903   1.307256 NC   1   GCC -0.7716  0.0000 ****
  26 C4     0.985900   6.327676   2.549121 CA   1   GCC  0.8616  0.0000 ****
  27 N4     1.765827   7.344614   2.953967 N2   1   GCC -0.9778  0.0000 ****
  28 H41    2.217869   7.327563   3.837497 H    1   GCC  0.4244  0.0000 ****
  29 H42    2.103523   7.973587   2.260860 H    1   GCC  0.4244  0.0000 ****
  30 C5     0.596715   5.323905   3.500865 CM   1   GCC -0.5198  0.0000 ****
  31 H5     0.916436   5.349404   4.523358 HA   1   GCC  0.2025  0.0000 ****
  32 C6    -0.200141   4.354734   3.032524 CM   1   GCC -0.0194  0.0000 ****
  33 H6    -0.547291   3.543921   3.640494 H4   1   GCC  0.1940  0.0000 ****
  34 C'     1.525134   0.000000  -0.000000 C    1   GCC  0.8129  0.0000 ****
  35 O1'    2.036814  -1.124430  -0.000000 O2   1   GCC -0.8266  0.0000 ****
  36 OXT    2.065904   1.108429  -0.000000 O2   1   GCC -0.8266  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3    17 1
    6     5     6 1
    7     5     7 1
    8     5     8 1
    9     8     9 1
   10     8    12 1
   11     9    10 1
   12     9    11 1
   13     9    34 1
   14    12    13 1
   15    12    14 1
   16    14    15 1
   17    14    16 1
   18    14    22 1
   19    17    18 1
   20    17    19 1
   21    17    20 1
   22    20    21 1
   23    22    23 1
   24    22    32 1
   25    23    24 1
   26    23    25 1
   27    25    26 1
   28    26    27 1
   29    26    30 1
   30    27    28 1
   31    27    29 1
   32    30    31 1
   33    30    32 1
   34    32    33 1
   35    34    35 1
   36    34    36 1
@<TRIPOS>SUBSTRUCTURE
     1 GCC              1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

MCP

```
@<TRIPOS>MOLECULE
MCP
    47    47     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 C1      3.679805  -2.260148   2.273309 CT   1   MCP -0.1310  0.0000 ****
   2 H1      2.926850  -3.040248   2.254910 HC   1   MCP  0.0460  0.0000 ****
   3 H2      4.621422  -2.690784   1.946091 HC   1   MCP  0.0460  0.0000 ****
   4 H3      3.800560  -1.891206   3.281362 HC   1   MCP  0.0460  0.0000 ****
   5 C       3.323224  -1.104279   1.363367 C    1   MCP  0.5575  0.0000 ****
   6 O       3.558932   0.034592   1.656503 O    1   MCP -0.5770  0.0000 ****
   7 N1'     2.717500  -1.457508   0.194461 N    1   MCP -0.3361  0.0000 ****
   8 H1'1    2.788920  -2.419137  -0.055916 H    1   MCP  0.2690  0.0000 ****
   9 C2'     2.630417  -0.523116  -0.917345 CT   1   MCP  0.1463  0.0000 ****
  10 H2'1    3.388500   0.232783  -0.767886 H1   1   MCP  0.1139  0.0000 ****
  11 C3'     1.248910   0.134907  -1.024890 CT   1   MCP -0.1738  0.0000 ****
  12 H3'1    0.504790  -0.599531  -1.287396 H1   1   MCP  0.0712  0.0000 ****
  13 H3'2    1.269210   0.861886  -1.828337 H1   1   MCP  0.0712  0.0000 ****
  14 N4'     0.830830   0.844382   0.178566 N    1   MCP  0.0073  0.0000 ****
  15 C5'     1.388083   2.179326   0.377348 CT   1   MCP -0.0141  0.0000 ****
  16 H5'1    1.407954   2.392945   1.433406 H1   1   MCP  0.0549  0.0000 ****
  17 H5'2    2.399076   2.188070   0.005038 H1   1   MCP  0.0549  0.0000 ****
  18 C'      0.595847   3.228539  -0.395664 C    1   MCP  0.5472  0.0000 ****
  19 O1'     0.943714   3.590385  -1.490504 O    1   MCP -0.5600  0.0000 ****
  20 C7'    -0.107239   0.425715   1.042028 C    1   MCP  0.4267  0.0000 ****
  21 O7'    -0.583920   1.152502   1.889173 O    1   MCP -0.5674  0.0000 ****
  22 C8'    -0.596562  -1.021874   1.014113 CT   1   MCP -0.1183  0.0000 ****
  23 H8'1   -0.531795  -1.362822   2.039178 H1   1   MCP  0.0559  0.0000 ****
  24 H8'2   -0.016645  -1.675714   0.396835 H1   1   MCP  0.0559  0.0000 ****
  25 C9'     2.930310  -1.251338  -2.222117 CT   1   MCP  0.0338  0.0000 ****
  26 H9'1    2.924591  -0.531221  -3.035294 H1   1   MCP  0.0785  0.0000 ****
  27 H9'2    2.152720  -1.986351  -2.424248 H1   1   MCP  0.0785  0.0000 ****
  28 O10'    4.181628  -1.873440  -2.105175 OH   1   MCP -0.6968  0.0000 ****
  29 H10'    4.406673  -2.291285  -2.924595 HO   1   MCP  0.4539  0.0000 ****
  30 N1     -1.984179  -1.106290   0.576026 N*   1   MCP -0.0010  0.0000 ****
  31 C2     -2.267068  -1.655975  -0.689648 C    1   MCP  0.7905  0.0000 ****
  32 O2     -1.351038  -2.022043  -1.380657 O    1   MCP -0.6190  0.0000 ****
  33 N3     -3.567503  -1.740644  -1.060912 NC   1   MCP -0.7781  0.0000 ****
  34 C4     -4.512582  -1.323108  -0.274271 CA   1   MCP  0.8462  0.0000 ****
  35 N4     -5.780838  -1.459083  -0.697812 N2   1   MCP -0.9721  0.0000 ****
  36 H41    -6.533134  -1.019636  -0.222212 H    1   MCP  0.4244  0.0000 ****
  37 H42    -5.926363  -1.766910  -1.632749 H    1   MCP  0.4244  0.0000 ****
  38 C5     -4.265886  -0.749157   1.019776 CM   1   MCP -0.5185  0.0000 ****
  39 H5     -5.058468  -0.412818   1.657783 HA   1   MCP  0.2030  0.0000 ****
  40 C6     -2.980078  -0.667278   1.385305 CM   1   MCP -0.0348  0.0000 ****
  41 H6     -2.667692  -0.247377   2.320001 H4   1   MCP  0.2055  0.0000 ****
  42 N      -0.501691   3.694204   0.227672 N    1   MCP -0.4405  0.0000 ****
  43 H      -0.796491   3.215623   1.050367 H    1   MCP  0.2932  0.0000 ****
  44 CH3    -1.416689   4.596588  -0.433545 CT   1   MCP -0.1192  0.0000 ****
  45 HH31   -1.983923   4.101447  -1.215571 H1   1   MCP  0.0853  0.0000 ****
  46 HH32   -2.104773   4.994056   0.301041 H1   1   MCP  0.0853  0.0000 ****
  47 HH33   -0.868289   5.414008  -0.879122 H1   1   MCP  0.0853  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
```

```
     5      5      6 1
     6      5      7 1
     7      7      8 1
     8      7      9 1
     9      9     10 1
    10      9     11 1
    11      9     25 1
    12     11     12 1
    13     11     13 1
    14     11     14 1
    15     14     15 1
    16     14     20 1
    17     15     16 1
    18     15     17 1
    19     15     18 1
    20     18     19 1
    21     18     42 1
    22     20     21 1
    23     20     22 1
    24     22     23 1
    25     22     24 1
    26     22     30 1
    27     25     26 1
    28     25     27 1
    29     25     28 1
    30     28     29 1
    31     30     31 1
    32     30     40 1
    33     31     32 1
    34     31     33 1
    35     33     34 1
    36     34     35 1
    37     34     38 1
    38     35     36 1
    39     35     37 1
    40     38     39 1
    41     38     40 1
    42     40     41 1
    43     42     43 1
    44     42     44 1
    45     44     45 1
    46     44     46 1
    47     44     47 1
@<TRIPOS>SUBSTRUCTURE
      1 MCP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 0 0 0 0 0
```

GGP

```
@<TRIPOS>MOLECULE
GGP
    38    39     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'     3.885223  -0.880910  -0.808808 N     1   GGP -0.3871  0.0000 ****
   2 H1'1    4.692324  -0.821494  -1.390197 H     1   GGP  0.2784  0.0000 ****
   3 C2'     3.269966   0.403563  -0.501697 CT    1   GGP  0.0528  0.0000 ****
   4 H2'1    3.627542   0.743034   0.463926 H1    1   GGP  0.1363  0.0000 ****
   5 C3'     1.722821   0.356073  -0.484518 CT    1   GGP -0.1234  0.0000 ****
   6 H3'1    1.388455  -0.566907  -0.917081 H1    1   GGP  0.0809  0.0000 ****
   7 H3'2    1.327562   1.143279  -1.113853 H1    1   GGP  0.0809  0.0000 ****
   8 N4'     1.148909   0.539995   0.839740 N     1   GGP -0.1034  0.0000 ****
   9 C5'     1.193173   1.896219   1.365522 CT    1   GGP -0.1424  0.0000 ****
  10 H5'1    1.269349   1.861910   2.440148 H1    1   GGP  0.0802  0.0000 ****
  11 H5'2    2.066925   2.395857   0.973812 H1    1   GGP  0.0802  0.0000 ****
  12 C'     -0.022711   2.704404   0.913884 C     1   GGP  0.6199  0.0000 ****
  13 O1'    -0.060654   3.200901  -0.191064 O     1   GGP -0.5806  0.0000 ****
  14 C7'     0.331682  -0.315533   1.497025 C     1   GGP  0.5312  0.0000 ****
  15 O7'    -0.229491   0.010256   2.514325 O     1   GGP -0.5649  0.0000 ****
  16 C8'     0.084256  -1.731059   0.973989 CT    1   GGP -0.3588  0.0000 ****
  17 H8'1    0.323946  -2.395413   1.790117 H1    1   GGP  0.1686  0.0000 ****
  18 H8'2    0.695268  -2.010694   0.137551 H1    1   GGP  0.1686  0.0000 ****
  19 C9'     3.729343   1.392596  -1.566309 CT    1   GGP  0.1284  0.0000 ****
  20 H9'1    3.365225   2.382666  -1.309302 H1    1   GGP  0.0497  0.0000 ****
  21 H9'2    3.300096   1.116687  -2.527664 H1    1   GGP  0.0497  0.0000 ****
  22 O10'    5.131334   1.367779  -1.625421 OH    1   GGP -0.6690  0.0000 ****
  23 H10'    5.437384   1.965740  -2.292987 HO    1   GGP  0.4382  0.0000 ****
  24 N9     -1.318472  -1.899626   0.647736 N*    1   GGP  0.0925  0.0000 ****
  25 C8     -2.151295  -2.929493   1.019778 CK    1   GGP  0.0810  0.0000 ****
  26 H8     -1.810756  -3.699205   1.681946 H5    1   GGP  0.1655  0.0000 ****
  27 N7     -3.319112  -2.853317   0.504598 NB    1   GGP -0.5385  0.0000 ****
  28 C6     -4.289069  -1.079564  -1.071869 C     1   GGP  0.4329  0.0000 ****
  29 O6     -5.412616  -1.396765  -1.323070 O     1   GGP -0.5353  0.0000 ****
  30 C5     -3.286072  -1.707024  -0.258674 CB    1   GGP  0.2010  0.0000 ****
  31 C4     -2.056538  -1.116212  -0.172703 CB    1   GGP  0.0655  0.0000 ****
  32 N3     -1.620990   0.032575  -0.741352 NC    1   GGP -0.3538  0.0000 ****
  33 C2     -2.523556   0.622336  -1.451609 CA    1   GGP  0.4952  0.0000 ****
  34 N2     -2.229830   1.791259  -2.076673 N2    1   GGP -0.9092  0.0000 ****
  35 H21    -1.438697   2.273128  -1.694739 H     1   GGP  0.4099  0.0000 ****
  36 H22    -2.988062   2.385710  -2.327274 H     1   GGP  0.4099  0.0000 ****
  37 N1     -3.772306   0.116500  -1.633233 NA    1   GGP -0.3450  0.0000 ****
  38 H1     -4.414685   0.588749  -2.232978 H     1   GGP  0.3137  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3    19 1
    6     5     6 1
    7     5     7 1
    8     5     8 1
    9     8     9 1
   10     8    14 1
   11     9    10 1
   12     9    11 1
   13     9    12 1
```

```
   14    12    13 1
   15    14    15 1
   16    14    16 1
   17    16    17 1
   18    16    18 1
   19    16    24 1
   20    19    20 1
   21    19    21 1
   22    19    22 1
   23    22    23 1
   24    24    25 1
   25    24    31 1
   26    25    26 1
   27    25    27 1
   28    27    30 1
   29    28    29 1
   30    28    30 1
   31    28    37 1
   32    30    31 1
   33    31    32 1
   34    32    33 1
   35    33    34 1
   36    33    37 1
   37    34    35 1
   38    34    36 1
   39    37    38 1
@<TRIPOS>SUBSTRUCTURE
      1 GGP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

GNG

```
@<TRIPOS>MOLECULE
GNG
    40    41     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'    0.000000   0.000000  -0.000000 N3   1   GNG -0.0679  0.0000 ****
   2 H1'1  -0.372149   0.940540  -0.000000 H    1   GNG  0.2683  0.0000 ****
   3 H1'2  -0.372151  -0.470283  -0.814524 H    1   GNG  0.2683  0.0000 ****
   4 H1'3  -0.372144  -0.470264   0.814532 H    1   GNG  0.2683  0.0000 ****
   5 C2'    1.507376  -0.000000   0.000000 CT   1   GNG  0.3376  0.0000 ****
   6 H2'1   1.859200   0.541465  -0.870956 HP   1   GNG  0.0685  0.0000 ****
   7 C3'    2.122471  -1.420516  -0.000000 CT   1   GNG -0.0833  0.0000 ****
   8 H3'1   1.358780  -2.140998   0.220192 H1   1   GNG  0.0545  0.0000 ****
   9 H3'2   2.850738  -1.501736   0.796879 H1   1   GNG  0.0545  0.0000 ****
  10 N4'    2.806114  -1.759431  -1.238802 N    1   GNG -0.0870  0.0000 ****
  11 C5'    4.093915  -1.114611  -1.447430 CT   1   GNG -0.1820  0.0000 ****
  12 H5'1   4.258019  -0.974729  -2.503510 H1   1   GNG  0.0902  0.0000 ****
  13 H5'2   4.082859  -0.146317  -0.969098 H1   1   GNG  0.0902  0.0000 ****
  14 C'     5.224675  -1.916774  -0.804444 C    1   GNG  0.6203  0.0000 ****
  15 O1'    5.445504  -1.832517   0.384251 O    1   GNG -0.5801  0.0000 ****
  16 C7'    2.535524  -2.803226  -2.056778 C    1   GNG  0.5146  0.0000 ****
```

```
  17 O7'     3.274157  -3.096001  -2.964885 O    1   GNG -0.5620  0.0000 ****
  18 C8'     1.281814  -3.656573  -1.860538 CT   1   GNG -0.3590  0.0000 ****
  19 H8'1    0.766924  -3.650548  -2.809082 H1   1   GNG  0.1808  0.0000 ****
  20 H8'2    0.600949  -3.285147  -1.119171 H1   1   GNG  0.1808  0.0000 ****
  21 C9'     1.960922   0.742656   1.251148 CT   1   GNG  0.0891  0.0000 ****
  22 H9'1    3.041717   0.844064   1.229432 H1   1   GNG  0.0461  0.0000 ****
  23 H9'2    1.696300   0.162963   2.133482 H1   1   GNG  0.0461  0.0000 ****
  24 O10'    1.334538   1.998280   1.278344 OH   1   GNG -0.6783  0.0000 ****
  25 H10'    1.591751   2.471051   2.057572 HO   1   GNG  0.4459  0.0000 ****
  26 N9      1.656789  -5.023667  -1.555662 N*   1   GNG  0.0604  0.0000 ****
  27 C8      1.178959  -6.176564  -2.134578 CK   1   GNG  0.0966  0.0000 ****
  28 H8      0.496144  -6.135073  -2.958551 H5   1   GNG  0.1596  0.0000 ****
  29 N7      1.630678  -7.244108  -1.594848 NB   1   GNG -0.5410  0.0000 ****
  30 C6      3.271740  -7.511572   0.356470 C    1   GNG  0.4206  0.0000 ****
  31 O6      3.413603  -8.681415   0.549914 O    1   GNG -0.5339  0.0000 ****
  32 C5      2.466422  -6.798686  -0.594451 CB   1   GNG  0.2116  0.0000 ****
  33 C4      2.486193  -5.432275  -0.567367 CB   1   GNG  0.0660  0.0000 ****
  34 N3      3.195186  -4.605619   0.236837 NC   1   GNG -0.3192  0.0000 ****
  35 C2      3.946556  -5.234987   1.077222 CA   1   GNG  0.4407  0.0000 ****
  36 N2      4.721299  -4.532969   1.943424 N2   1   GNG -0.8765  0.0000 ****
  37 H21     4.892542  -3.585672   1.665455 H    1   GNG  0.4011  0.0000 ****
  38 H22     5.512743  -4.993186   2.334455 H    1   GNG  0.4011  0.0000 ****
  39 N1      3.989653  -6.591107   1.162750 NA   1   GNG -0.3243  0.0000 ****
  40 H1      4.550843  -7.024708   1.864476 H    1   GNG  0.3125  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     5    21 1
    8     7     8 1
    9     7     9 1
   10     7    10 1
   11    10    11 1
   12    10    16 1
   13    11    12 1
   14    11    13 1
   15    11    14 1
   16    14    15 1
   17    16    17 1
   18    16    18 1
   19    18    19 1
   20    18    20 1
   21    18    26 1
   22    21    22 1
   23    21    23 1
   24    21    24 1
   25    24    25 1
   26    26    27 1
   27    26    33 1
   28    27    28 1
   29    27    29 1
   30    29    32 1
   31    30    31 1
   32    30    32 1
   33    30    39 1
   34    32    33 1
   35    33    34 1
   36    34    35 1
   37    35    36 1
```

```
   38    35    39 1
   39    36    37 1
   40    36    38 1
   41    39    40 1
@<TRIPOS>SUBSTRUCTURE
      1 GNG               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
C' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 C' 0 0 0 0 0
```

GCG

```
@<TRIPOS>MOLECULE
GCG
    39    40     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
   1 N1'   -2.967890   2.379092  -2.284478 N    1   GCG -0.3796  0.0000 ****
   2 H1'1  -3.406790   2.348313  -3.178555 H    1   GCG  0.2738  0.0000 ****
   3 C2'   -1.889871   1.414513  -2.110720 CT   1   GCG  0.0724  0.0000 ****
   4 H2'1  -2.280208   0.537709  -1.606357 H1   1   GCG  0.1339  0.0000 ****
   5 C3'   -0.689136   1.971916  -1.308382 CT   1   GCG -0.1382  0.0000 ****
   6 H3'1  -0.783380   3.036598  -1.216876 H1   1   GCG  0.0801  0.0000 ****
   7 H3'2   0.227624   1.797618  -1.857158 H1   1   GCG  0.0801  0.0000 ****
   8 N4'   -0.526617   1.356622  -0.000000 N    1   GCG -0.0476  0.0000 ****
   9 C5'    0.000000  -0.000000   0.000000 CT   1   GCG -0.5030  0.0000 ****
  10 H5'1  -0.396333  -0.538526   0.845406 H1   1   GCG  0.1684  0.0000 ****
  11 H5'2  -0.315180  -0.498534  -0.904789 H1   1   GCG  0.1684  0.0000 ****
  12 C7'   -0.523101   1.985187   1.198627 C    1   GCG  0.5677  0.0000 ****
  13 O7'   -0.204978   1.407021   2.208813 O    1   GCG -0.5857  0.0000 ****
  14 C8'   -0.920253   3.456914   1.320235 CT   1   GCG -0.3124  0.0000 ****
  15 H8'1  -1.703475   3.492915   2.061955 H1   1   GCG  0.1489  0.0000 ****
  16 H8'2  -1.307070   3.890551   0.418297 H1   1   GCG  0.1489  0.0000 ****
  17 C9'   -1.417699   1.006679  -3.501165 CT   1   GCG  0.1197  0.0000 ****
  18 H9'1  -0.680373   0.215390  -3.405851 H1   1   GCG  0.0523  0.0000 ****
  19 H9'2  -0.937988   1.855318  -3.985181 H1   1   GCG  0.0523  0.0000 ****
  20 O10'  -2.528794   0.577347  -4.243345 OH   1   GCG -0.6738  0.0000 ****
  21 H10'  -2.258784   0.333116  -5.117595 HO   1   GCG  0.4408  0.0000 ****
  22 N9     0.203046   4.233690   1.807389 N*   1   GCG  0.0894  0.0000 ****
  23 C8     0.211072   5.149352   2.834099 CK   1   GCG  0.0757  0.0000 ****
  24 H8    -0.662604   5.311947   3.431714 H5   1   GCG  0.1679  0.0000 ****
  25 N7     1.332735   5.746379   2.977069 NB   1   GCG -0.5384  0.0000 ****
  26 C6     3.508356   5.460320   1.651509 C    1   GCG  0.4327  0.0000 ****
  27 O6     4.308753   6.222144   2.104350 O    1   GCG -0.5358  0.0000 ****
  28 C5     2.138216   5.207838   1.997867 CB   1   GCG  0.2036  0.0000 ****
  29 C4     1.447019   4.275452   1.275989 CB   1   GCG  0.0617  0.0000 ****
  30 N3     1.876062   3.499727   0.252783 NC   1   GCG -0.3583  0.0000 ****
  31 C2     3.115941   3.696160  -0.048918 CA   1   GCG  0.4921  0.0000 ****
  32 N2     3.685140   2.981204  -1.053015 N2   1   GCG -0.9048  0.0000 ****
  33 H21    3.197662   2.136408  -1.282466 H    1   GCG  0.4085  0.0000 ****
  34 H22    4.676823   2.893746  -1.054811 H    1   GCG  0.4085  0.0000 ****
  35 N1     3.895632   4.615646   0.579462 NA   1   GCG -0.3428  0.0000 ****
  36 H1     4.833699   4.761078   0.272470 H    1   GCG  0.3130  0.0000 ****
  37 C'     1.528239   0.000000  -0.000000 C    1   GCG  0.8129  0.0000 ****
  38 O1'    2.039919  -1.124430  -0.000000 O2   1   GCG -0.8266  0.0000 ****
```

```
  39 OXT    2.069009   1.108429   0.000000 O2   1   GCG -0.8266  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     3     4 1
    4     3     5 1
    5     3    17 1
    6     5     6 1
    7     5     7 1
    8     5     8 1
    9     8     9 1
   10     8    12 1
   11     9    10 1
   12     9    11 1
   13     9    37 1
   14    12    13 1
   15    12    14 1
   16    14    15 1
   17    14    16 1
   18    14    22 1
   19    17    18 1
   20    17    19 1
   21    17    20 1
   22    20    21 1
   23    22    23 1
   24    22    29 1
   25    23    24 1
   26    23    25 1
   27    25    28 1
   28    26    27 1
   29    26    28 1
   30    26    35 1
   31    28    29 1
   32    29    30 1
   33    30    31 1
   34    31    32 1
   35    31    35 1
   36    32    33 1
   37    32    34 1
   38    35    36 1
   39    37    38 1
   40    37    39 1
@<TRIPOS>SUBSTRUCTURE
      1 GCG               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
N1' 1
0 0
@<TRIPOS>RESIDUECONNECT
1 N1' 0 0 0 0 0
```

MGP

```
@<TRIPOS>MOLECULE
MGP
    50    51     1     0     1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
```

```
   1 C1      4.797332  -3.055821  -0.240181 CT   1   MGP -0.2105  0.0000 ****
   2 H2      4.398571  -3.973878   0.166687 HC   1   MGP  0.0680  0.0000 ****
   3 H3      4.983696  -3.174554  -1.301012 HC   1   MGP  0.0680  0.0000 ****
   4 H4      5.742392  -2.845956   0.251674 HC   1   MGP  0.0680  0.0000 ****
   5 C       3.832568  -1.927072   0.049755 C    1   MGP  0.6718  0.0000 ****
   6 O       3.073211  -1.962992   0.982695 O    1   MGP -0.6062  0.0000 ****
   7 N1'     3.885223  -0.880910  -0.808808 N    1   MGP -0.5466  0.0000 ****
   8 H1'1    4.692324  -0.821494  -1.390197 H    1   MGP  0.3188  0.0000 ****
   9 C2'     3.269966   0.403563  -0.501697 CT   1   MGP  0.0879  0.0000 ****
  10 H2'1    3.627542   0.743034   0.463926 H1   1   MGP  0.1240  0.0000 ****
  11 C3'     1.722821   0.356073  -0.484518 CT   1   MGP -0.0561  0.0000 ****
  12 H3'1    1.388455  -0.566907  -0.917081 H1   1   MGP  0.0476  0.0000 ****
  13 H3'2    1.327562   1.143279  -1.113853 H1   1   MGP  0.0476  0.0000 ****
  14 N4'     1.148909   0.539995   0.839740 N    1   MGP -0.1749  0.0000 ****
  15 C5'     1.193173   1.896219   1.365522 CT   1   MGP -0.1384  0.0000 ****
  16 H5'1    1.269349   1.861910   2.440148 H1   1   MGP  0.0884  0.0000 ****
  17 H5'2    2.066925   2.395857   0.973812 H1   1   MGP  0.0884  0.0000 ****
  18 C'     -0.022711   2.704404   0.913884 C    1   MGP  0.5906  0.0000 ****
  19 O1'    -0.060654   3.200901  -0.191064 O    1   MGP -0.5690  0.0000 ****
  20 C7'     0.331682  -0.315533   1.497025 C    1   MGP  0.5935  0.0000 ****
  21 O7'    -0.229491   0.010256   2.514325 O    1   MGP -0.5839  0.0000 ****
  22 C8'     0.084256  -1.731059   0.973989 CT   1   MGP -0.1737  0.0000 ****
  23 H8'1    0.323946  -2.395413   1.790117 H1   1   MGP  0.1426  0.0000 ****
  24 H8'2    0.695268  -2.010694   0.137551 H1   1   MGP  0.1426  0.0000 ****
  25 C9'     3.729343   1.392596  -1.566309 CT   1   MGP  0.1164  0.0000 ****
  26 H9'1    3.365225   2.382666  -1.309302 H1   1   MGP  0.0568  0.0000 ****
  27 H9'2    3.300096   1.116687  -2.527664 H1   1   MGP  0.0568  0.0000 ****
  28 O10'    5.131334   1.367779  -1.625421 OH   1   MGP -0.6703  0.0000 ****
  29 H10'    5.437384   1.965740  -2.292987 HO   1   MGP  0.4425  0.0000 ****
  30 N9     -1.318472  -1.899626   0.647736 N*   1   MGP -0.0583  0.0000 ****
  31 C8     -2.151295  -2.929493   1.019778 CK   1   MGP  0.1204  0.0000 ****
  32 H8     -1.810756  -3.699205   1.681946 H5   1   MGP  0.1610  0.0000 ****
  33 N7     -3.319112  -2.853317   0.504598 NB   1   MGP -0.5544  0.0000 ****
  34 C6     -4.289069  -1.079564  -1.071869 C    1   MGP  0.3876  0.0000 ****
  35 O6     -5.412616  -1.396765  -1.323070 O    1   MGP -0.5301  0.0000 ****
  36 C5     -3.286072  -1.707024  -0.258674 CB   1   MGP  0.2799  0.0000 ****
  37 C4     -2.056538  -1.116212  -0.172703 CB   1   MGP -0.0067  0.0000 ****
  38 N3     -1.620990   0.032575  -0.741352 NC   1   MGP -0.2318  0.0000 ****
  39 C2     -2.523556   0.622336  -1.451609 CA   1   MGP  0.3686  0.0000 ****
  40 N2     -2.229830   1.791259  -2.076673 N2   1   MGP -0.8561  0.0000 ****
  41 H21    -1.438697   2.273128  -1.694739 H    1   MGP  0.4002  0.0000 ****
  42 H22    -2.988062   2.385710  -2.327274 H    1   MGP  0.4002  0.0000 ****
  43 N1     -3.772306   0.116500  -1.633233 NA   1   MGP -0.2892  0.0000 ****
  44 H1     -4.414685   0.588749  -2.232978 H    1   MGP  0.3076  0.0000 ****
  45 N      -0.997952   2.830308   1.823108 N    1   MGP -0.5123  0.0000 ****
  46 H      -0.983172   2.174325   2.574093 H    1   MGP  0.3597  0.0000 ****
  47 CH3    -2.250717   3.484390   1.517669 CT   1   MGP -0.0655  0.0000 ****
  48 HH31   -2.913524   2.846151   0.942320 H1   1   MGP  0.0763  0.0000 ****
  49 HH32   -2.742007   3.747994   2.444593 H1   1   MGP  0.0763  0.0000 ****
  50 HH33   -2.057496   4.385125   0.954315 H1   1   MGP  0.0763  0.0000 ****
@<TRIPOS>BOND
    1     1     2 1
    2     1     3 1
    3     1     4 1
    4     1     5 1
    5     5     6 1
    6     5     7 1
    7     7     8 1
    8     7     9 1
    9     9    10 1
   10     9    11 1
   11     9    25 1
```

```
   12    11    12 1
   13    11    13 1
   14    11    14 1
   15    14    15 1
   16    14    20 1
   17    15    16 1
   18    15    17 1
   19    15    18 1
   20    18    19 1
   21    18    45 1
   22    20    21 1
   23    20    22 1
   24    22    23 1
   25    22    24 1
   26    22    30 1
   27    25    26 1
   28    25    27 1
   29    25    28 1
   30    28    29 1
   31    30    31 1
   32    30    37 1
   33    31    32 1
   34    31    33 1
   35    33    36 1
   36    34    35 1
   37    34    36 1
   38    34    43 1
   39    36    37 1
   40    37    38 1
   41    38    39 1
   42    39    40 1
   43    39    43 1
   44    40    41 1
   45    40    42 1
   46    43    44 1
   47    45    46 1
   48    45    47 1
   49    47    48 1
   50    47    49 1
   51    47    50 1
@<TRIPOS>SUBSTRUCTURE
      1 MGP               1 ****               0 ****  ****
@<TRIPOS>HEADTAIL
0 0
0 0
@<TRIPOS>RESIDUECONNECT
1 0 0 0 0 0 0
```

frcmod.pna

```
PNA frcmod file: frcmod.pna
MASS

BOND
CQ-NA        502.0    1.324        by analogy to CQ-NC, parm99.dat

ANGL
C -CT-N*    50.00     110.13       FF cst by analogy to CT-CT-N*; Eq.val. from
QM geo.opt.
CQ-NA-C     70.0      118.60       by analogy to CA-NC-CQ, parm99.dat
H5-CQ-NA    50.0      115.45       by analogy to H5-CQ-NC, parm99.dat
NC-CQ-NA    70.0      129.10       by analogy to NC-CQ-NC, parm99.dat
CQ-NA-H     50.0      118.00       by analogy to CA-NA-H, parm99.dat

DIHEDRAL
X -CQ-NA-X    4    6.00        180.0         2.0      by analogy to X -CA-NA-X,
parm99.dat

NONB
```

8.3. Modified NAB program

8.3.1. fd_helix.nab file

```
// put_tmp is a helper function to simplify the code somewhat -
// all it does is output the proper elements to a temporary
// pdb file.

int put_tmp( float x, float y, float z, int count, int rescount,
string atomname, string resname, file outfile, int nres ) {

// check if 5' end

if( rescount == 1 || rescount == ( nres / 2 ) + 1 )
      if( atomname =~ "P" )
                  return 1;

if( atomname == "O1'" )
      atomname = "O4'";

// I have no idea why this is necessary.. would be just as easy
// to change O1' to O4' in all data files, but they must have had
// some reason for naming it O1' in the data.

fprintf(outfile, "ATOM  %5d %-4s %3s%6d    %8.3lf%8.3lf%8.3lf\n",
      count, atomname, resname, rescount, x, y, z );

return 1;

};

// fd_helix is a function designed to build helices of varying types
```

```
// from fiber diffraction data.  The current available types are:

// arna           Right Handed A-RNA (Arnott)
// aprna          Right Handed A-PRIME RNA (Arnott)
// lbdna          Right Handed B-DNA (Langridge)
// abdna          Right Handed B-DNA (Arnott)
// sbdna          Left  Handed B-DNA (Sasisekharan)
// adna           Right Handed A-DNA (Arnott)

// fd_helix will create a temporary file in your working directory
// called nab_tmp.pdb.  This file contains the helix _before_ the
// addition of hydrogens.  This is ordinarily deleted, but could be
// examined if things go wrong.


molecule fd_helix( string helix_type, string seq, string acid_type )
{

// ------------------------------------------------------------------ //

//    Data Segment

molecule m;

// These arrays are used to store the r, phi, zz values of the
// various atom types within each residue:

float ade_r[ hashed ];
float ade_phi[ hashed ];
float ade_zz[ hashed ];

float gua_r[ hashed ];
float gua_phi[ hashed ];
float gua_zz[ hashed ];

float thy_r[ hashed ];
float thy_phi[ hashed ];
float thy_zz[ hashed ];

float ura_r[ hashed ];
float ura_phi[ hashed ];
float ura_zz[ hashed ];

float cyt_r[ hashed ];
float cyt_phi[ hashed ];
float cyt_zz[ hashed ];

float apn_r[ hashed ];
float apn_phi[ hashed ];
float apn_zz[ hashed ];

float tpn_r[ hashed ];
float tpn_phi[ hashed ];
float tpn_zz[ hashed ];

float cpn_r[ hashed ];
float cpn_phi[ hashed ];
float cpn_zz[ hashed ];

float gpn_r[ hashed ];
float gpn_phi[ hashed ];
float gpn_zz[ hashed ];
```

```
float dim_r[ hashed ];
float dim_phi[ hashed ];
float dim_zz[ hashed ];

float nbc_r[ hashed ];
float nbc_phi[ hashed ];
float nbc_zz[ hashed ];


// These are simply temp variables used to aid in filling the arrays.

float temp_r, temp_phi, temp_zz;

// These are used to calculate and output coordinates to the temporary pdb file.

float x, y, z, yyr, xrad;

// height values are angstroms, rotation; values are degrees.

float current_height, current_rotation;
float height_increment, rotation_increment;

float hxht[ hashed ];
float hxrep[ hashed ];

string tempname, cseq, fullseq, buffer, restype;
string temp, amberhome;
string resout;

int nres, nresh, i, hxmul, count, chain, begin, end;

file infile, outfile;

// ------------------------------------------------------------

//    Code Segment

hxht[ "arna" ] = 3.30; hxht[ "aprna" ] = 3.00; hxht[ "lbdna" ] = 3.38;
hxht[ "abdna" ] = 3.38; hxht[ "sbdna" ] = -3.38; hxht[ "adna" ] = 2.56;
hxht[ "apna" ] = 3.30;

hxrep["arna"] = 23.2; hxrep["aprna"] = 30.0; hxrep["lbdna"] = 36.0;
hxrep["abdna"] = 36.0; hxrep["sbdna"] = 36.0; hxrep["adna"] = 32.7;
hxrep["apna"] = 23.2;

temp = wc_complement( seq, acid_type + "amber94.rlb", acid_type );
cseq = "";
for( i = length( temp ); i >= 1; i-- ) {
      cseq += substr( temp, i, 1) ;
}
fullseq = seq + cseq;

nresh = length( seq );

nres = length( fullseq );

if( !( amberhome = getenv( "AMBERHOME" ) ) ){
    fprintf( stderr, "AMBERHOME not defined.\n" );
    exit( 1 );
}
```

```
temp = amberhome + "/dat/fd_data/" + helix_type + ".dat";
printf("Data file: %s\n", temp);

infile = fopen( temp, "r" );
if( infile == NULL ){
      fprintf( stderr, "Unable to open data file %s; exiting\n", temp );
      exit(1);
}

outfile = fopen( "nab_tmp.pdb", "w" );

// Read lines from data file and store r, phi, zz values as appropriate
// residue type.

while( buffer = getline( infile ) ) {
      sscanf( buffer, "%s %lf %lf %lf %s", tempname, temp_r, temp_phi,
            temp_zz, restype );

      if( restype == "A" || restype == "a" ) {
            ade_r[ tempname ] = temp_r;
            ade_phi[ tempname ] = temp_phi;
            ade_zz[ tempname ] = temp_zz;
      }
      else if( restype == "G" || restype == "g" ) {
            gua_r[ tempname ] = temp_r;
            gua_phi[ tempname ] = temp_phi;
            gua_zz[ tempname ] = temp_zz;
      }
      else if( restype == "T" || restype == "t" ) {
            thy_r[ tempname ] = temp_r;
            thy_phi[ tempname ] = temp_phi;
            thy_zz[ tempname ] = temp_zz;
      }
      else if( restype == "U" || restype == "u" ) {
            ura_r[ tempname ] = temp_r;
            ura_phi[ tempname ] = temp_phi;
            ura_zz[ tempname ] = temp_zz;
      }
      else if( restype == "C" || restype == "c" ) {
            cyt_r[ tempname ] = temp_r;
            cyt_phi[ tempname ] = temp_phi;
            cyt_zz[ tempname ] = temp_zz;
      }
        else if( restype == "X" || restype == "x" ) {
                apn_r[ tempname ] = temp_r;
                apn_phi[ tempname ] = temp_phi;
                apn_zz[ tempname ] = temp_zz;
        }
       else if( restype == "Y" || restype == "y" ) {
                tpn_r[ tempname ] = temp_r;
                tpn_phi[ tempname ] = temp_phi;
                tpn_zz[ tempname ] = temp_zz;
        }
       else if( restype == "J" || restype == "j" ) {
                cpn_r[ tempname ] = temp_r;
                cpn_phi[ tempname ] = temp_phi;
                cpn_zz[ tempname ] = temp_zz;
        }
       else if( restype == "K" || restype == "k" ) {
                gpn_r[ tempname ] = temp_r;
                gpn_phi[ tempname ] = temp_phi;
                gpn_zz[ tempname ] = temp_zz;
```

```
        }
       else if( restype == "W" || restype == "w" ) {
                dim_r[ tempname ] = temp_r;
                dim_phi[ tempname ] = temp_phi;
                dim_zz[ tempname ] = temp_zz;
        }
       else if( restype == "Q" || restype == "q" ) {
                nbc_r[ tempname ] = temp_r;
                nbc_phi[ tempname ] = temp_phi;
                nbc_zz[ tempname ] = temp_zz;
        }


}

height_increment = hxht[ helix_type ];
rotation_increment = hxrep[ helix_type ];
current_height = 0;
current_rotation = 0;
count = 0;

//    Here we build the actual molecule - it is output to a temporary
//    pdb file in order to allow the addition of hydrogens as a final
//    stage.

for( chain = 1; chain <= 2; chain++ ) {

      if(chain == 1) {
            begin = 1;
            end = nresh;
            hxmul = -1;
      }

      else if( chain == 2) {
            begin = nresh + 1;
            end = nres;
            hxmul = 1;
      }
      for( i = begin; i <= end; i++ ) {
            restype = substr( fullseq, i, 1 );
            if( restype == "A" || restype == "a" ) {
                  resout = "DA"; if( acid_type == "rna" ) resout = "A";
                  for( tempname in ade_r ) {
                        count++;
                        yyr = (hxmul * ade_phi[ tempname ] + current_rotation);
                        xrad = ade_r[ tempname ];
                        x = xrad * cos(yyr);
                        y = xrad * sin(yyr);
                        z = hxmul * ade_zz[ tempname ] + current_height;
                        put_tmp( x, y, z, count, i, tempname, resout, outfile,
nres);
                  }
            }

            else if( restype == "G" || restype == "g" ) {
                  resout = "DG"; if( acid_type == "rna" ) resout = "G";
                  for( tempname in gua_r ) {
                        count++;
                        yyr = (hxmul * gua_phi[ tempname ] + current_rotation);
                        xrad = gua_r[ tempname ];
                        x = xrad * cos(yyr);
                        y = xrad * sin(yyr);
```

```
                        z = hxmul * gua_zz[ tempname ] + current_height;
                        put_tmp( x, y, z, count, i, tempname, resout, outfile,
nres);
                  }

            }

            else if( restype == "T" || restype == "t" ) {
                  resout = "DT";
                  for( tempname in thy_r ) {
                        count++;
                        yyr = (hxmul * thy_phi[ tempname ] + current_rotation);
                        xrad = thy_r[ tempname ];
                        x = xrad * cos(yyr);
                        y = xrad * sin(yyr);
                        z = hxmul * thy_zz[ tempname ] + current_height;
                        put_tmp( x, y, z, count, i, tempname, resout, outfile,
nres);
                  }

            }

            else if( restype == "U" || restype == "u" ) {
                  resout = "U";
                  for( tempname in ura_r ) {
                        count++;
                        yyr = (hxmul * ura_phi[ tempname ] + current_rotation);
                        xrad = ura_r[ tempname ];
                        x = xrad * cos(yyr);
                        y = xrad * sin(yyr);
                        z = hxmul * ura_zz[ tempname ] + current_height;
                        put_tmp( x, y, z, count, i, tempname, resout, outfile,
nres);
                  }

            }

            else if( restype == "C" || restype == "c" ) {
                  resout = "DC"; if( acid_type == "rna" ) resout = "C";
                  for( tempname in cyt_r ) {
                        count++;
                        yyr = (hxmul * cyt_phi[ tempname ] + current_rotation);
                        xrad = cyt_r[ tempname ];
                        x = xrad * cos(yyr);
                        y = xrad * sin(yyr);
                        z = hxmul * cyt_zz[ tempname ] + current_height;
                        put_tmp( x, y, z, count, i, tempname, resout, outfile,
nres);
                  }

            }

                else if( restype == "X" || restype == "x" ) {
                        resout = "APN"; if( acid_type == "rna" ) resout = "C";
                        for( tempname in apn_r ) {
                                count++;
                                yyr = (hxmul * apn_phi[ tempname ] +
current_rotation);
                                xrad = apn_r[ tempname ];
                                x = xrad * cos(yyr);
                                y = xrad * sin(yyr);
                                z = hxmul * apn_zz[ tempname ] + current_height;
```

```
                                put_tmp( x, y, z, count, i, tempname, resout,
outfile, nres);
                        }

                }
            else if( restype == "Y" || restype == "y" ) {
                        resout = "TPN"; if( acid_type == "rna" ) resout = "C";
                        for( tempname in tpn_r ) {
                                count++;
                                yyr = (hxmul * tpn_phi[ tempname ] +
current_rotation);
                                xrad = tpn_r[ tempname ];
                                x = xrad * cos(yyr);
                                y = xrad * sin(yyr);
                                z = hxmul * tpn_zz[ tempname ] + current_height;
                                put_tmp( x, y, z, count, i, tempname, resout,
outfile, nres);
                        }

                }
            else if( restype == "J" || restype == "j" ) {
                        resout = "CPN"; if( acid_type == "rna" ) resout = "C";
                        for( tempname in cpn_r ) {
                                count++;
                                yyr = (hxmul * cpn_phi[ tempname ] +
current_rotation);
                                xrad = cpn_r[ tempname ];
                                x = xrad * cos(yyr);
                                y = xrad * sin(yyr);
                                z = hxmul * cpn_zz[ tempname ] + current_height;
                                put_tmp( x, y, z, count, i, tempname, resout,
outfile, nres);
                        }

                }
            else if( restype == "K" || restype == "k" ) {
                        resout = "GPN"; if( acid_type == "rna" ) resout = "C";
                        for( tempname in gpn_r ) {
                                count++;
                                yyr = (hxmul * gpn_phi[ tempname ] +
current_rotation);
                                xrad = gpn_r[ tempname ];
                                x = xrad * cos(yyr);
                                y = xrad * sin(yyr);
                                z = hxmul * gpn_zz[ tempname ] + current_height;
                                put_tmp( x, y, z, count, i, tempname, resout,
outfile, nres);
                        }


            }
            else if( restype == "W" || restype == "w" ) {
                        resout = "DIM"; if( acid_type == "rna" ) resout = "C";
                        for( tempname in dim_r ) {
                                count++;
                                yyr = (hxmul * dim_phi[ tempname ] +
current_rotation);
                                xrad = dim_r[ tempname ];
                                x = xrad * cos(yyr);
                                y = xrad * sin(yyr);
                                z = hxmul * dim_zz[ tempname ] + current_height;
```

```
                                put_tmp( x, y, z, count, i, tempname, resout,
outfile, nres);
                        }


                }
            else if( restype == "Q" || restype == "q" ) {
                        resout = "NBC"; if( acid_type == "rna" ) resout = "C";
                        for( tempname in nbc_r ) {
                                count++;
                                yyr = (hxmul * nbc_phi[ tempname ] +
current_rotation);
                                xrad = nbc_r[ tempname ];
                                x = xrad * cos(yyr);
                                y = xrad * sin(yyr);
                                z = hxmul * nbc_zz[ tempname ] + current_height;
                                put_tmp( x, y, z, count, i, tempname, resout,
outfile, nres);
                        }


                }

            // Increase unit twist and height

            current_height += height_increment;
            current_rotation += rotation_increment;
      }


height_increment = -height_increment;
rotation_increment = -rotation_increment;

current_rotation += rotation_increment;
current_height += height_increment;

if(chain == 1)
      fprintf( outfile, "TER\n" );  // Need TER card for getpdb_prm to
                              // function properly

}

fclose( infile );

// Close outfile in order to flush output stream.
// Otherwise nab_tmp.pdb would be incomplete when
// read by getpdb_prm

fclose( outfile );

m = getpdb_prm( "nab_tmp.pdb", "leaprc.pna", "", 0 );
unlink( "nab_tmp.pdb" );

return m;

};
```

8.3.2. wc_complement.nab

```
// wc_complement() - create W/C string to seq.
string      wc_complement( string seq, string rlib, string rlt )
//    (note the rlib is unused: just there for backwards compatibility)
{
      string      acbase, base, wcbase, wcseq;
      int   i, len;

      if( rlt == "dna" )
            acbase = "t";
      else if( rlt == "rna" )
            acbase = "u";
      else if( rlt == "pna" )
                acbase = "t";
      else{
            fprintf( stderr,
            "wc_complement: rlt (%s) is not dna/rna, has no W/C complement\n",
                  rlt );
            exit( 1 );
      }

      len = length( seq );
      wcseq = NULL;
      for( i = 1; i <= len; i = i + 1 ){
            base = substr( seq, i, 1 );
            if( base == "a" || base == "A" )
                  wcbase = acbase;
            else if( base == "c" || base == "C" )
                  wcbase = "g";
            else if( base == "g" || base == "G" )
                  wcbase = "c";
            else if( base == "t" || base == "T" )
                  wcbase = "a";
            else if( base == "u" || base == "U" )
                  wcbase = "a";
            else if( base == "x" || base == "APN" )
                        wcbase = "acbase";
            else if( base == "y" || base == "TPN" )
                        wcbase = "a";
            else if( base == "j" || base == "CPN" )
                        wcbase = "g";
            else if( base == "k" || base == "GPN" )
                        wcbase = "c";
            else if( base == "w" || base == "DIM" )
                        wcbase = "a";
            else if( base == "q" || base == "NBC" )
                        wcbase = "a";

            else{
                  fprintf( stderr, "wc_complement: unknown base %s\n",
                        base );
                  exit( 1 );
            }
            wcseq = wcseq + wcbase;
      }
      return( wcseq );
};
```

8.3.3. pna.amber94.rlb file

```
#
#       bpna single bases
#
pdb      pna.amber94/pna.amber94.pdb
bnd     pna.amber94/pna.amber94.bnd
qr      pna.amber94/pna.amber94.qr
chi     pna.amber94/pna.amber94.chi
type    pna
atomtype        all
```

8.3.4. pna.amber94.pdb file

```
ATOM      1 P    ADE     1      -2.946  -5.462  -1.138
ATOM      2 O1P  ADE     1      -2.555  -6.846  -0.805
ATOM      3 O2P  ADE     1      -4.059  -4.892  -0.340
ATOM      4 O5'  ADE     1      -1.679  -4.498  -1.044
ATOM      5 C5'  ADE     1      -0.625  -4.632  -2.031
ATOM      6 H5'1 ADE     1      -0.795  -3.931  -2.848
ATOM      7 H5'2 ADE     1      -0.617  -5.647  -2.425
ATOM      8 C4'  ADE     1       0.712  -4.336  -1.394
ATOM      9 H4'  ADE     1       1.480  -4.833  -1.986
ATOM     10 O4'  ADE     1       0.889  -2.884  -1.330
ATOM     11 C1'  ADE     1       0.731  -2.441   0.004
ATOM     12 H1'  ADE     1       1.684  -2.029   0.336
ATOM     13 N9   ADE     1      -0.275  -1.345   0.002
ATOM     14 C8   ADE     1      -1.640  -1.447   0.001
ATOM     15 H8   ADE     1      -2.167  -2.401   0.007
ATOM     16 N7   ADE     1      -2.248  -0.298  -0.007
ATOM     17 C5   ADE     1      -1.215   0.623   0.003
ATOM     18 C6   ADE     1      -1.198   2.033  -0.001
ATOM     19 N6   ADE     1      -2.314   2.780  -0.005
ATOM     20 H61  ADE     1      -2.246   3.807  -0.008
ATOM     21 H62  ADE     1      -3.238   2.324  -0.005
ATOM     22 N1   ADE     1       0.000   2.637   0.000
ATOM     23 C2   ADE     1       1.098   1.887   0.005
ATOM     24 H2   ADE     1       2.038   2.438   0.011
ATOM     25 N3   ADE     1       1.216   0.576   0.003
ATOM     26 C4   ADE     1      -0.000   0.000  -0.000
ATOM     27 C3'  ADE     1       0.894  -4.801   0.050
ATOM     28 H3'  ADE     1       0.352  -5.728   0.235
ATOM     29 C2'  ADE     1       0.330  -3.640   0.853
ATOM     30 H2'1 ADE     1      -0.750  -3.724   0.966
ATOM     31 H2'2 ADE     1       0.763  -3.602   1.853
ATOM     32 O3'  ADE     1       2.267  -4.981   0.388
ATOM      1 P    GUA     1      -2.846  -5.524  -1.132
ATOM      2 O1P  GUA     1      -2.426  -6.899  -0.800
ATOM      3 O2P  GUA     1      -3.969  -4.977  -0.332
ATOM      4 O5'  GUA     1      -1.599  -4.534  -1.040
ATOM      5 C5'  GUA     1      -0.545  -4.646  -2.029
ATOM      6 H5'1 GUA     1      -0.731  -3.949  -2.846
ATOM      7 H5'2 GUA     1      -0.516  -5.661  -2.424
ATOM      8 C4'  GUA     1       0.787  -4.322  -1.395
ATOM      9 H4'  GUA     1       1.564  -4.803  -1.989
ATOM     10 O4'  GUA     1       0.935  -2.868  -1.331
ATOM     11 C1'  GUA     1       0.770  -2.428   0.003
ATOM     12 H1'  GUA     1       1.715  -1.996   0.333
ATOM     13 N9   GUA     1      -0.259  -1.353   0.003
```

```
ATOM     14 C8   GUA     1      -1.634  -1.459   0.003
ATOM     15 H8   GUA     1      -2.154  -2.417   0.003
ATOM     16 N7   GUA     1      -2.259  -0.310   0.005
ATOM     17 C5   GUA     1      -1.230   0.634   0.002
ATOM     18 C6   GUA     1      -1.287   2.046   0.000
ATOM     19 O6   GUA     1      -2.273   2.783   0.000
ATOM     20 N1   GUA     1      -0.000   2.612  -0.000
ATOM     21 H1   GUA     1       0.067   3.641  -0.007
ATOM     22 C2   GUA     1       1.180   1.898   0.008
ATOM     23 N2   GUA     1       2.306   2.619   0.011
ATOM     24 H21  GUA     1       2.258   3.648   0.008
ATOM     25 H22  GUA     1       3.220   2.145   0.017
ATOM     26 N3   GUA     1       1.235   0.568   0.002
ATOM     27 C4   GUA     1       0.000  -0.000   0.000
ATOM     28 C3'  GUA     1       0.982  -4.784   0.048
ATOM     29 H3'  GUA     1       0.459  -5.722   0.235
ATOM     30 C2'  GUA     1       0.396  -3.635   0.853
ATOM     31 H2'1 GUA     1      -0.683  -3.741   0.968
ATOM     32 H2'2 GUA     1       0.829  -3.588   1.852
ATOM     33 O3'  GUA     1       2.359  -4.936   0.384
ATOM      1 P    THY     1      -0.049  -4.059  -1.130
ATOM      2 O1P  THY     1       0.760  -5.248  -0.796
ATOM      3 O2P  THY     1      -1.284  -3.869  -0.332
ATOM      4 O5'  THY     1       0.849  -2.745  -1.038
ATOM      5 C5'  THY     1       1.890  -2.540  -2.026
ATOM      6 H5'1 THY     1       1.507  -1.930  -2.844
ATOM      7 H5'2 THY     1       2.219  -3.501  -2.419
ATOM      8 C4'  THY     1       3.066  -1.836  -1.390
ATOM      9 H4'  THY     1       3.952  -2.066  -1.983
ATOM     10 O4'  THY     1       2.776  -0.403  -1.328
ATOM     11 C1'  THY     1       2.486  -0.030   0.005
ATOM     12 H1'  THY     1       3.261   0.662   0.336
ATOM     13 N1   THY     1       1.186   0.692   0.003
ATOM     14 C6   THY     1      -0.000  -0.000  -0.000
ATOM     15 H6   THY     1       0.015  -1.090  -0.005
ATOM     16 C5   THY     1      -1.181   0.648   0.003
ATOM     17 C7   THY     1      -2.504  -0.058  -0.006
ATOM     18 H71  THY     1      -2.630  -0.606   0.928
ATOM     19 H72  THY     1      -3.306   0.673  -0.109
ATOM     20 H73  THY     1      -2.539  -0.756  -0.843
ATOM     21 C4   THY     1      -1.234   2.090  -0.001
ATOM     22 O4   THY     1      -2.272   2.762  -0.003
ATOM     23 N3   THY     1      -0.000   2.699  -0.000
ATOM     24 H3   THY     1       0.007   3.729  -0.003
ATOM     25 C2   THY     1       1.230   2.071   0.004
ATOM     26 O2   THY     1       2.275   2.697   0.007
ATOM     27 C3'  THY     1       3.387  -2.218   0.053
ATOM     28 H3'  THY     1       3.165  -3.268   0.241
ATOM     29 C2'  THY     1       2.485  -1.294   0.856
ATOM     30 H2'1 THY     1       1.487  -1.714   0.970
ATOM     31 H2'2 THY     1       2.884  -1.120   1.856
ATOM     32 O3'  THY     1       4.747  -1.955   0.391
ATOM      1 P    CYT     1      -0.006  -4.082  -1.120
ATOM      2 O1P  CYT     1       0.817  -5.261  -0.785
ATOM      3 O2P  CYT     1      -1.243  -3.906  -0.321
ATOM      4 O5'  CYT     1       0.876  -2.757  -1.031
ATOM      5 C5'  CYT     1       1.913  -2.541  -2.021
ATOM      6 H5'1 CYT     1       1.521  -1.937  -2.840
ATOM      7 H5'2 CYT     1       2.253  -3.499  -2.413
ATOM      8 C4'  CYT     1       3.081  -1.821  -1.388
ATOM      9 H4'  CYT     1       3.968  -2.042  -1.982
ATOM     10 O4'  CYT     1       2.773  -0.392  -1.329
```

```
ATOM     11 C1'  CYT     1       2.481  -0.021   0.005
ATOM     12 H1'  CYT     1       3.247   0.682   0.333
ATOM     13 N1   CYT     1       1.171   0.686   0.002
ATOM     14 C6   CYT     1      -0.000   0.000  -0.000
ATOM     15 H6   CYT     1       0.020  -1.090  -0.004
ATOM     16 C5   CYT     1      -1.199   0.644   0.002
ATOM     17 H5   CYT     1      -2.142   0.097   0.009
ATOM     18 C4   CYT     1      -1.161   2.073  -0.005
ATOM     19 N4   CYT     1      -2.286   2.770  -0.010
ATOM     20 H41  CYT     1      -3.194   2.283  -0.013
ATOM     21 H42  CYT     1      -2.255   3.800  -0.010
ATOM     22 N3   CYT     1      -0.000   2.741  -0.000
ATOM     23 C2   CYT     1       1.183   2.079  -0.000
ATOM     24 O2   CYT     1       2.271   2.668   0.004
ATOM     25 C3'  CYT     1       3.408  -2.197   0.055
ATOM     26 H3'  CYT     1       3.199  -3.250   0.245
ATOM     27 C2'  CYT     1       2.496  -1.282   0.858
ATOM     28 H2'1 CYT     1       1.503  -1.715   0.974
ATOM     29 H2'2 CYT     1       2.894  -1.101   1.857
ATOM     30 O3'  CYT     1       4.765  -1.917   0.391
ATOM      1 P    APU     1      -0.049  -4.059  -1.130
ATOM      2 O1P  APU     1       0.760  -5.248  -0.796
ATOM      3 O2P  APU     1      -1.284  -3.869  -0.332
ATOM      4 O5'  APU     1       0.849  -2.745  -1.038
ATOM      5 C5'  APU     1       1.890  -2.540  -2.026
ATOM      6 H5'1 APU     1       1.507  -1.930  -2.844
ATOM      7 H5'2 APU     1       2.219  -3.501  -2.419
ATOM      8 C4'  APU     1       3.066  -1.836  -1.390
ATOM      9 H4'  APU     1       3.952  -2.066  -1.983
ATOM     10 O4'  APU     1       2.776  -0.403  -1.328
ATOM     11 C1'  APU     1       2.486  -0.030   0.005
ATOM     12 H1'  APU     1       3.261   0.662   0.336
ATOM     13 O1'  APU     1       1.186   0.692   0.003
ATOM     14 HO1' APU     1      -0.000  -0.000  -0.000
ATOM     15 C3'  APU     1       3.387  -2.218   0.053
ATOM     16 H3'  APU     1       3.165  -3.268   0.241
ATOM     17 C2'  APU     1       2.485  -1.294   0.856
ATOM     18 H2'1 APU     1       1.487  -1.714   0.970
ATOM     19 H2'2 APU     1       2.884  -1.120   1.856
ATOM     20 O3'  APU     1       4.747  -1.955   0.391
ATOM      1  N1' APN     1       3.621  -1.450  -0.069
ATOM      2 H1'1 APN     1       4.246  -1.272   0.683
ATOM      3  C2' APN     1       3.458  -0.420  -1.077
ATOM      4 H2'1 APN     1       3.628  -0.848  -2.057
ATOM      5 H2'2 APN     1       4.227   0.322  -0.906
ATOM      6  C3' APN     1       2.080   0.241  -1.080
ATOM      7 H3'1 APN     1       1.324  -0.471  -1.363
ATOM      8 H3'2 APN     1       2.071   1.016  -1.836
ATOM      9  N4' APN     1       1.741   0.868   0.196
ATOM     10  C5' APN     1       2.299   2.197   0.430
ATOM     11 H5'1 APN     1       2.392   2.357   1.493
ATOM     12 H5'2 APN     1       3.276   2.251  -0.025
ATOM     13  C'  APN     1       1.450   3.289  -0.217
ATOM     14  O1' APN     1       1.768   3.771  -1.273
ATOM     15  C7' APN     1       0.858   0.403   1.100
ATOM     16  O7' APN     1       0.506   1.058   2.054
ATOM     17  C8' APN     1       0.288  -1.011   0.978
ATOM     18 H8'1 APN     1       0.638  -1.548   1.851
ATOM     19 H8'2 APN     1       0.593  -1.558   0.104
ATOM     20  N9  APN     1      -1.153  -0.974   1.005
ATOM     21  C8  APN     1      -1.982  -0.899   2.098
ATOM     22  H8  APN     1      -1.577  -0.897   3.089
```

```
ATOM     23  N7  APN     1      -3.226  -0.831   1.805
ATOM     24  C6  APN     1      -4.279  -0.782  -0.521
ATOM     25  N6  APN     1      -5.576  -0.734  -0.172
ATOM     26  H61 APN     1      -5.821  -0.568   0.777
ATOM     27  H62 APN     1      -6.246  -0.541  -0.880
ATOM     28  C5  APN     1      -3.248  -0.843   0.423
ATOM     29  C4  APN     1      -1.972  -0.921  -0.084
ATOM     30  N3  APN     1      -1.628  -0.930  -1.367
ATOM     31  C2  APN     1      -2.681  -0.862  -2.150
ATOM     32  H2  APN     1      -2.486  -0.862  -3.207
ATOM     33  N1  APN     1      -3.961  -0.790  -1.811
ATOM      1  C'  CAP     1       0.000   0.000   0.000
ATOM      2  O1' CAP     1      -0.512  -1.124   0.000
ATOM      3  OXT CAP     1      -0.541   1.108   0.000
ATOM      4  N1' CAP     1       4.659   2.044   1.148
ATOM      5 H1'1 CAP     1       5.197   1.232   0.950
ATOM      6  C2' CAP     1       3.404   1.871   1.856
ATOM      7 H2'1 CAP     1       3.389   2.519   2.723
ATOM      8 H2'2 CAP     1       3.374   0.849   2.212
ATOM      9  C3' CAP     1       2.164   2.173   1.015
ATOM     10 H3'1 CAP     1       2.133   3.217   0.756
ATOM     11 H3'2 CAP     1       1.284   1.978   1.616
ATOM     12  N4' CAP     1       2.069   1.344  -0.185
ATOM     13  C5' CAP     1       1.527   0.000   0.000
ATOM     14 H5'1 CAP     1       1.916  -0.644  -0.773
ATOM     15 H5'2 CAP     1       1.839  -0.372   0.964
ATOM     16  C7' CAP     1       2.293   1.748  -1.450
ATOM     17  O7' CAP     1       2.033   1.053  -2.405
ATOM     18  C8' CAP     1       2.936   3.107  -1.732
ATOM     19 H8'1 CAP     1       3.884   2.893  -2.208
ATOM     20 H8'2 CAP     1       3.126   3.716  -0.867
ATOM     21  N9  CAP     1       2.119   3.866  -2.647
ATOM     22  C8  CAP     1       2.068   3.777  -4.016
ATOM     23  H8  CAP     1       2.712   3.103  -4.543
ATOM     24  N7  CAP     1       1.204   4.551  -4.557
ATOM     25  C6  CAP     1      -0.402   6.153  -3.385
ATOM     26  N6  CAP     1      -1.010   6.679  -4.463
ATOM     27  H61 CAP     1      -0.863   6.269  -5.356
ATOM     28  H62 CAP     1      -1.821   7.236  -4.319
ATOM     29  C5  CAP     1       0.614   5.198  -3.487
ATOM     30  C4  CAP     1       1.164   4.775  -2.299
ATOM     31  N3  CAP     1       0.818   5.173  -1.081
ATOM     32  C2  CAP     1      -0.148   6.063  -1.118
ATOM     33  H2  CAP     1      -0.487   6.434  -0.168
ATOM     34  N1  CAP     1      -0.767   6.567  -2.177
ATOM      1  C'  CCP     1       0.000   0.000   0.000
ATOM      2  O1' CCP     1      -0.512  -1.124   0.000
ATOM      3  OXT CCP     1      -0.541   1.108   0.000
ATOM      4  N1' CCP     1       4.548   2.222   1.204
ATOM      5 H1'1 CCP     1       5.114   1.414   1.093
ATOM      6  C2' CCP     1       3.247   2.062   1.829
ATOM      7 H2'1 CCP     1       3.151   2.775   2.637
ATOM      8 H2'2 CCP     1       3.221   1.070   2.263
ATOM      9  C3' CCP     1       2.066   2.253   0.881
ATOM     10 H3'1 CCP     1       2.020   3.268   0.530
ATOM     11 H3'2 CCP     1       1.151   2.079   1.435
ATOM     12  N4' CCP     1       2.080   1.327  -0.251
ATOM     13  C5' CCP     1       1.528   0.000   0.000
ATOM     14 H5'1 CCP     1       1.912  -0.684  -0.741
ATOM     15 H5'2 CCP     1       1.839  -0.327   0.981
ATOM     16  C7' CCP     1       2.319   1.666  -1.533
ATOM     17  O7' CCP     1       2.067   0.925  -2.456
```

```
ATOM     18  C8' CCP     1       3.021   2.983  -1.865
ATOM     19 H8'1 CCP     1       4.035   2.720  -2.140
ATOM     20 H8'2 CCP     1       3.064   3.687  -1.056
ATOM     21  N1  CCP     1       2.387   3.609  -3.006
ATOM     22  C2  CCP     1       1.123   4.184  -2.781
ATOM     23  O2  CCP     1       0.673   4.116  -1.668
ATOM     24  N3  CCP     1       0.498   4.768  -3.831
ATOM     25  C4  CCP     1       1.022   4.716  -5.017
ATOM     26  N4  CCP     1       0.366   5.334  -6.017
ATOM     27  H41 CCP     1       0.615   5.181  -6.965
ATOM     28  H42 CCP     1      -0.543   5.689  -5.820
ATOM     29  C5  CCP     1       2.263   4.051  -5.307
ATOM     30  H5  CCP     1       2.665   3.994  -6.298
ATOM     31  C6  CCP     1       2.891   3.508  -4.254
ATOM     32  H6  CCP     1       3.822   2.984  -4.349
ATOM      1  C'  CGP     1       0.000   0.000   0.000
ATOM      2  O1' CGP     1      -0.512  -1.124   0.000
ATOM      3  OXT CGP     1      -0.541   1.108   0.000
ATOM      4  N1' CGP     1       4.733   2.038   1.113
ATOM      5 H1'1 CGP     1       5.293   1.308   0.738
ATOM      6  C2' CGP     1       3.569   1.675   1.899
ATOM      7 H2'1 CGP     1       3.614   2.178   2.857
ATOM      8 H2'2 CGP     1       3.630   0.610   2.086
ATOM      9  C3' CGP     1       2.237   2.023   1.239
ATOM     10 H3'1 CGP     1       2.152   3.088   1.129
ATOM     11 H3'2 CGP     1       1.428   1.718   1.889
ATOM     12  N4' CGP     1       2.057   1.357  -0.047
ATOM     13  C5' CGP     1       1.529   0.000   0.000
ATOM     14 H5'1 CGP     1       1.923  -0.570  -0.827
ATOM     15 H5'2 CGP     1       1.845  -0.462   0.924
ATOM     16  C7' CGP     1       2.035   1.955  -1.260
ATOM     17  O7' CGP     1       1.713   1.358  -2.258
ATOM     18  C8' CGP     1       2.465   3.417  -1.408
ATOM     19 H8'1 CGP     1       3.222   3.422  -2.180
ATOM     20 H8'2 CGP     1       2.885   3.852  -0.518
ATOM     21  N9  CGP     1       1.356   4.220  -1.882
ATOM     22  C8  CGP     1       1.365   5.151  -2.897
ATOM     23  H8  CGP     1       2.241   5.306  -3.493
ATOM     24  N7  CGP     1       0.255   5.768  -3.030
ATOM     25  C6  CGP     1      -1.923   5.510  -1.702
ATOM     26  O6  CGP     1      -2.709   6.291  -2.147
ATOM     27  C5  CGP     1      -0.560   5.233  -2.056
ATOM     28  C4  CGP     1       0.114   4.280  -1.346
ATOM     29  N3  CGP     1      -0.327   3.501  -0.332
ATOM     30  C2  CGP     1      -1.560   3.725  -0.016
ATOM     31  N2  CGP     1      -2.134   3.020   0.987
ATOM     32  H21 CGP     1      -1.655   2.176   1.234
ATOM     33  H22 CGP     1      -3.127   2.961   1.014
ATOM     34  N1  CGP     1      -2.322   4.667  -0.633
ATOM     35  H1  CGP     1      -3.253   4.834  -0.316
ATOM      1  N1' CPN     1       3.313  -1.543   0.213
ATOM      2 H1'1 CPN     1       3.703  -1.508   1.125
ATOM      3  C2' CPN     1       3.410  -0.355  -0.616
ATOM      4 H2'1 CPN     1       3.831  -0.625  -1.576
ATOM      5 H2'2 CPN     1       4.106   0.317  -0.131
ATOM      6  C3' CPN     1       2.079   0.351  -0.862
ATOM      7 H3'1 CPN     1       1.419  -0.270  -1.440
ATOM      8 H3'2 CPN     1       2.266   1.236  -1.458
ATOM      9  N4' CPN     1       1.418   0.779   0.370
ATOM     10  C5' CPN     1       1.868   2.049   0.931
ATOM     11 H5'1 CPN     1       1.604   2.082   1.976
ATOM     12 H5'2 CPN     1       2.942   2.113   0.832
```

```
ATOM     13  C'  CPN     1       1.286   3.249   0.184
ATOM     14  O1' CPN     1       1.962   3.877  -0.589
ATOM     15  C7' CPN     1       0.290   0.243   0.878
ATOM     16  O7' CPN     1      -0.353   0.791   1.744
ATOM     17  C8' CPN     1      -0.169  -1.147   0.438
ATOM     18 H8'1 CPN     1       0.068  -1.813   1.258
ATOM     19 H8'2 CPN     1       0.309  -1.521  -0.447
ATOM     20  N1  CPN     1      -1.603  -1.156   0.242
ATOM     21  C2  CPN     1      -2.082  -0.490  -0.899
ATOM     22  O2  CPN     1      -1.272   0.028  -1.621
ATOM     23  N3  CPN     1      -3.422  -0.471  -1.105
ATOM     24  C4  CPN     1      -4.228  -0.981  -0.225
ATOM     25  N4  CPN     1      -5.545  -0.953  -0.495
ATOM     26  H41 CPN     1      -6.217  -1.160   0.206
ATOM     27  H42 CPN     1      -5.840  -0.435  -1.292
ATOM     28  C5  CPN     1      -3.781  -1.582   1.003
ATOM     29  H5  CPN     1      -4.464  -1.973   1.730
ATOM     30  C6  CPN     1      -2.454  -1.629   1.179
ATOM     31  H6  CPN     1      -2.006  -2.050   2.057
ATOM      1  C'  CTP     1       0.000   0.000   0.000
ATOM      2  O1' CTP     1      -0.512  -1.124   0.000
ATOM      3  OXT CTP     1      -0.541   1.108   0.000
ATOM      4  N1' CTP     1       4.595   2.187   1.178
ATOM      5 H1'1 CTP     1       5.185   1.393   1.088
ATOM      6  C2' CTP     1       3.310   2.006   1.830
ATOM      7 H2'1 CTP     1       3.231   2.691   2.665
ATOM      8 H2'2 CTP     1       3.294   1.000   2.230
ATOM      9  C3' CTP     1       2.105   2.230   0.919
ATOM     10 H3'1 CTP     1       2.064   3.255   0.597
ATOM     11 H3'2 CTP     1       1.205   2.051   1.494
ATOM     12  N4' CTP     1       2.076   1.333  -0.235
ATOM     13  C5' CTP     1       1.528   0.000   0.000
ATOM     14 H5'1 CTP     1       1.915  -0.674  -0.748
ATOM     15 H5'2 CTP     1       1.839  -0.336   0.977
ATOM     16  C7' CTP     1       2.306   1.688  -1.514
ATOM     17  O7' CTP     1       2.032   0.967  -2.446
ATOM     18  C8' CTP     1       3.017   3.004  -1.830
ATOM     19 H8'1 CTP     1       4.031   2.741  -2.101
ATOM     20 H8'2 CTP     1       3.063   3.701  -1.016
ATOM     21  N1  CTP     1       2.401   3.654  -2.970
ATOM     22  C2  CTP     1       1.246   4.351  -2.728
ATOM     23  O2  CTP     1       0.772   4.453  -1.631
ATOM     24  N3  CTP     1       0.681   4.920  -3.834
ATOM     25  H3  CTP     1      -0.150   5.450  -3.671
ATOM     26  C4  CTP     1       1.094   4.802  -5.153
ATOM     27  O4  CTP     1       0.491   5.348  -6.029
ATOM     28  C5  CTP     1       2.285   3.968  -5.341
ATOM     29  C6  CTP     1       2.849   3.439  -4.258
ATOM     30  H6  CTP     1       3.711   2.806  -4.326
ATOM     31  C7  CTP     1       2.782   3.753  -6.743
ATOM     32  H71 CTP     1       2.018   3.288  -7.356
ATOM     33  H72 CTP     1       3.037   4.698  -7.210
ATOM     34  H73 CTP     1       3.660   3.117  -6.748
ATOM      1  N1' GPN     1       4.122  -0.011  -0.354
ATOM      2 H1'1 GPN     1       4.686   0.388   0.360
ATOM      3  C2' GPN     1       3.375   0.883  -1.218
ATOM      4 H2'1 GPN     1       3.603   0.654  -2.252
ATOM      5 H2'2 GPN     1       3.728   1.888  -1.020
ATOM      6  C3' GPN     1       1.861   0.807  -1.034
ATOM      7 H3'1 GPN     1       1.509  -0.169  -1.309
ATOM      8 H3'2 GPN     1       1.386   1.504  -1.712
ATOM      9  N4' GPN     1       1.442   1.143   0.323
```

```
ATOM     10  C5' GPN     1       1.275   2.561   0.612
ATOM     11 H5'1 GPN     1       1.508   2.749   1.648
ATOM     12 H5'2 GPN     1       1.958   3.123  -0.007
ATOM     13  C'  GPN     1      -0.137   3.034   0.262
ATOM     14  O1' GPN     1      -0.437   3.292  -0.881
ATOM     15  C7' GPN     1       0.909   0.291   1.228
ATOM     16  O7' GPN     1       0.435   0.681   2.266
ATOM     17  C8' GPN     1       0.931  -1.219   0.974
ATOM     18 H8'1 GPN     1       1.388  -1.656   1.851
ATOM     19 H8'2 GPN     1       1.499  -1.519   0.110
ATOM     20  N9  GPN     1      -0.420  -1.732   0.886
ATOM     21  C8  GPN     1      -0.938  -2.848   1.506
ATOM     22  H8  GPN     1      -0.351  -3.414   2.200
ATOM     23  N7  GPN     1      -2.143  -3.104   1.168
ATOM     24  C6  GPN     1      -3.681  -1.852  -0.458
ATOM     25  O6  GPN     1      -4.721  -2.438  -0.491
ATOM     26  C5  GPN     1      -2.471  -2.112   0.269
ATOM     27  C4  GPN     1      -1.413  -1.265   0.092
ATOM     28  N3  GPN     1      -1.328  -0.155  -0.676
ATOM     29  C2  GPN     1      -2.418   0.102  -1.321
ATOM     30  N2  GPN     1      -2.477   1.183  -2.135
ATOM     31  H21 GPN     1      -1.767   1.871  -1.970
ATOM     32  H22 GPN     1      -3.371   1.552  -2.367
ATOM     33  N1  GPN     1      -3.526  -0.682  -1.245
ATOM     34  H1  GPN     1      -4.324  -0.466  -1.804
ATOM      1  N1' NAP     1       0.000   0.000   0.000
ATOM      2 H1'1 NAP     1      -0.372  -0.941   0.000
ATOM      3 H1'2 NAP     1      -0.372   0.470   0.815
ATOM      4 H1'3 NAP     1      -0.372   0.470  -0.815
ATOM      5  C2' NAP     1       1.451   0.000   0.000
ATOM      6 H2'1 NAP     1       1.809   0.729  -0.716
ATOM      7 H2'2 NAP     1       1.772  -0.977  -0.339
ATOM      8  C3' NAP     1       2.077   0.319   1.357
ATOM      9 H3'1 NAP     1       1.854   1.331   1.644
ATOM     10 H3'2 NAP     1       3.155   0.256   1.263
ATOM     11  N4' NAP     1       1.673  -0.613   2.409
ATOM     12  C5' NAP     1       2.392  -1.884   2.456
ATOM     13 H5'1 NAP     1       1.756  -2.630   2.906
ATOM     14 H5'2 NAP     1       2.636  -2.184   1.449
ATOM     15  C'  NAP     1       3.713  -1.759   3.213
ATOM     16  O1' NAP     1       4.753  -1.642   2.619
ATOM     17  C7' NAP     1       0.814  -0.353   3.413
ATOM     18  O7' NAP     1       0.656  -1.115   4.340
ATOM     19  C8' NAP     1      -0.041   0.915   3.400
ATOM     20 H8'1 NAP     1      -1.068   0.578   3.334
ATOM     21 H8'2 NAP     1       0.144   1.587   2.583
ATOM     22  N9  NAP     1       0.129   1.645   4.632
ATOM     23  C8  NAP     1      -0.484   1.427   5.841
ATOM     24  H8  NAP     1      -1.217   0.655   5.949
ATOM     25  N7  NAP     1      -0.093   2.213   6.773
ATOM     26  C6  NAP     1       1.678   4.047   6.623
ATOM     27  N6  NAP     1       1.615   4.514   7.882
ATOM     28  H61 NAP     1       1.102   4.007   8.566
ATOM     29  H62 NAP     1       2.320   5.146   8.184
ATOM     30  C5  NAP     1       0.862   3.006   6.165
ATOM     31  C4  NAP     1       1.016   2.659   4.844
ATOM     32  N3  NAP     1       1.863   3.200   3.975
ATOM     33  C2  NAP     1       2.574   4.158   4.526
ATOM     34  H2  NAP     1       3.286   4.647   3.886
ATOM     35  N1  NAP     1       2.533   4.605   5.774
ATOM      1  N1' NCP     1       0.000   0.000   0.000
ATOM      2 H1'1 NCP     1      -0.372  -0.941   0.000
```

```
ATOM      3 H1'2 NCP     1      -0.372   0.470   0.815
ATOM      4 H1'3 NCP     1      -0.372   0.470  -0.815
ATOM      5  C2' NCP     1       1.452   0.000   0.000
ATOM      6 H2'1 NCP     1       1.807   0.707  -0.739
ATOM      7 H2'2 NCP     1       1.772  -0.987  -0.311
ATOM      8  C3' NCP     1       2.081   0.359   1.344
ATOM      9 H3'1 NCP     1       1.859   1.376   1.609
ATOM     10 H3'2 NCP     1       3.158   0.297   1.245
ATOM     11  N4' NCP     1       1.683  -0.543   2.425
ATOM     12  C5' NCP     1       2.433  -1.791   2.526
ATOM     13 H5'1 NCP     1       1.846  -2.507   3.080
ATOM     14 H5'2 NCP     1       2.613  -2.171   1.532
ATOM     15  C'  NCP     1       3.802  -1.592   3.175
ATOM     16  O1' NCP     1       4.802  -1.581   2.506
ATOM     17  C7' NCP     1       0.880  -0.217   3.457
ATOM     18  O7' NCP     1       0.791  -0.905   4.448
ATOM     19  C8' NCP     1      -0.037   1.002   3.365
ATOM     20 H8'1 NCP     1      -1.035   0.614   3.201
ATOM     21 H8'2 NCP     1       0.194   1.681   2.567
ATOM     22  N1  NCP     1      -0.029   1.724   4.620
ATOM     23  C2  NCP     1       1.136   2.454   4.913
ATOM     24  O2  NCP     1       2.026   2.428   4.105
ATOM     25  N3  NCP     1       1.179   3.132   6.086
ATOM     26  C4  NCP     1       0.205   3.031   6.938
ATOM     27  N4  NCP     1       0.294   3.745   8.075
ATOM     28  H41 NCP     1      -0.320   3.576   8.836
ATOM     29  H42 NCP     1       1.154   4.212   8.258
ATOM     30  C5  NCP     1      -0.958   2.214   6.718
ATOM     31  H5  NCP     1      -1.738   2.122   7.446
ATOM     32  C6  NCP     1      -1.007   1.578   5.540
ATOM     33  H6  NCP     1      -1.824   0.939   5.267
ATOM      1  N1' NGP     1       0.000   0.000   0.000
ATOM      2 H1'1 NGP     1      -0.372  -0.941   0.000
ATOM      3 H1'2 NGP     1      -0.372   0.470   0.815
ATOM      4 H1'3 NGP     1      -0.372   0.470  -0.815
ATOM      5  C2' NGP     1       1.451   0.000   0.000
ATOM      6 H2'1 NGP     1       1.809   0.668  -0.774
ATOM      7 H2'2 NGP     1       1.770  -1.002  -0.259
ATOM      8  C3' NGP     1       2.074   0.428   1.327
ATOM      9 H3'1 NGP     1       1.817   1.449   1.538
ATOM     10 H3'2 NGP     1       3.152   0.398   1.239
ATOM     11  N4' NGP     1       1.689  -0.440   2.435
ATOM     12  C5' NGP     1       2.476  -1.652   2.607
ATOM     13 H5'1 NGP     1       1.855  -2.432   3.018
ATOM     14 H5'2 NGP     1       2.840  -1.969   1.640
ATOM     15  C'  NGP     1       3.703  -1.396   3.483
ATOM     16  O1' NGP     1       4.698  -0.886   3.020
ATOM     17  C7' NGP     1       0.899  -0.098   3.479
ATOM     18  O7' NGP     1       0.765  -0.818   4.437
ATOM     19  C8' NGP     1       0.108   1.213   3.453
ATOM     20 H8'1 NGP     1      -0.919   0.939   3.649
ATOM     21 H8'2 NGP     1       0.145   1.739   2.515
ATOM     22  N9  NGP     1       0.540   2.080   4.529
ATOM     23  C8  NGP     1      -0.250   2.777   5.417
ATOM     24  H8  NGP     1      -1.315   2.662   5.412
ATOM     25  N7  NGP     1       0.413   3.533   6.204
ATOM     26  C6  NGP     1       2.946   3.911   6.350
ATOM     27  O6  NGP     1       3.140   4.708   7.217
ATOM     28  C5  NGP     1       1.729   3.345   5.839
ATOM     29  C4  NGP     1       1.812   2.452   4.808
ATOM     30  N3  NGP     1       2.910   1.974   4.180
ATOM     31  C2  NGP     1       4.014   2.460   4.644
```

```
ATOM     32  N2  NGP     1       5.195   2.073   4.106
ATOM     33  H21 NGP     1       5.155   1.226   3.572
ATOM     34  H22 NGP     1       6.021   2.187   4.649
ATOM     35  N1  NGP     1       4.056   3.374   5.649
ATOM     36  H1  NGP     1       4.933   3.754   5.935
ATOM      1  N1' NTP     1       0.000   0.000   0.000
ATOM      2 H1'1 NTP     1      -0.372  -0.941   0.000
ATOM      3 H1'2 NTP     1      -0.372   0.470   0.815
ATOM      4 H1'3 NTP     1      -0.372   0.470  -0.815
ATOM      5  C2' NTP     1       1.452   0.000   0.000
ATOM      6 H2'1 NTP     1       1.811   0.710  -0.734
ATOM      7 H2'2 NTP     1       1.772  -0.986  -0.314
ATOM      8  C3' NTP     1       2.081   0.353   1.346
ATOM      9 H3'1 NTP     1       1.845   1.367   1.616
ATOM     10 H3'2 NTP     1       3.158   0.304   1.243
ATOM     11  N4' NTP     1       1.700  -0.561   2.422
ATOM     12  C5' NTP     1       2.457  -1.807   2.505
ATOM     13 H5'1 NTP     1       1.863  -2.543   3.023
ATOM     14 H5'2 NTP     1       2.663  -2.156   1.504
ATOM     15  C'  NTP     1       3.809  -1.612   3.190
ATOM     16  O1' NTP     1       4.817  -1.511   2.541
ATOM     17  C7' NTP     1       0.878  -0.269   3.448
ATOM     18  O7' NTP     1       0.793  -0.972   4.429
ATOM     19  C8' NTP     1      -0.057   0.938   3.365
ATOM     20 H8'1 NTP     1      -1.043   0.540   3.162
ATOM     21 H8'2 NTP     1       0.181   1.643   2.592
ATOM     22  N1  NTP     1      -0.104   1.632   4.637
ATOM     23  C2  NTP     1       0.939   2.477   4.914
ATOM     24  O2  NTP     1       1.838   2.665   4.143
ATOM     25  N3  NTP     1       0.871   3.086   6.135
ATOM     26  H3  NTP     1       1.613   3.721   6.343
ATOM     27  C4  NTP     1      -0.072   2.884   7.132
ATOM     28  O4  NTP     1       0.000   3.480   8.165
ATOM     29  C5  NTP     1      -1.106   1.900   6.795
ATOM     30  C6  NTP     1      -1.053   1.330   5.594
ATOM     31  H6  NTP     1      -1.767   0.591   5.291
ATOM     32  C7  NTP     1      -2.148   1.589   7.832
ATOM     33  H71 NTP     1      -1.690   1.211   8.739
ATOM     34  H72 NTP     1      -2.702   2.482   8.102
ATOM     35  H73 NTP     1      -2.848   0.847   7.466
ATOM      1  N1' TPN     1      -3.164  -2.077  -0.367
ATOM      2 H1'1 TPN     1      -3.506  -2.157  -1.296
ATOM      3  C2' TPN     1      -3.574  -0.919   0.408
ATOM      4 H2'1 TPN     1      -4.043  -1.249   1.326
ATOM      5 H2'2 TPN     1      -4.324  -0.395  -0.172
ATOM      6  C3' TPN     1      -2.435   0.029   0.773
ATOM      7 H3'1 TPN     1      -1.736  -0.458   1.430
ATOM      8 H3'2 TPN     1      -2.848   0.863   1.327
ATOM      9  N4' TPN     1      -1.740   0.577  -0.391
ATOM     10  C5' TPN     1      -2.365   1.732  -1.029
ATOM     11 H5'1 TPN     1      -2.030   1.790  -2.053
ATOM     12 H5'2 TPN     1      -3.437   1.599  -1.013
ATOM     13  C'  TPN     1      -2.071   3.030  -0.281
ATOM     14  O1' TPN     1      -2.886   3.507   0.466
ATOM     15  C7' TPN     1      -0.494   0.253  -0.788
ATOM     16  O7' TPN     1       0.117   0.907  -1.602
ATOM     17  C8' TPN     1       0.167  -1.027  -0.277
ATOM     18 H8'1 TPN     1       0.113  -1.739  -1.091
ATOM     19 H8'2 TPN     1      -0.303  -1.469   0.580
ATOM     20  N1  TPN     1       1.566  -0.793   0.023
ATOM     21  C2  TPN     1       1.843  -0.187   1.221
ATOM     22  O2  TPN     1       0.986   0.112   2.005
```

```
ATOM     23  N3  TPN     1       3.170   0.045   1.450
ATOM     24  H3  TPN     1       3.390   0.466   2.328
ATOM     25  C4  TPN     1       4.231  -0.185   0.587
ATOM     26  O4  TPN     1       5.349   0.081   0.916
ATOM     27  C5  TPN     1       3.845  -0.752  -0.709
ATOM     28  C6  TPN     1       2.555  -1.003  -0.917
ATOM     29  H6  TPN     1       2.204  -1.401  -1.848
ATOM     30  C7  TPN     1       4.928  -0.998  -1.722
ATOM     31  H71 TPN     1       5.446  -0.077  -1.964
ATOM     32  H72 TPN     1       5.668  -1.689  -1.335
ATOM     33  H73 TPN     1       4.515  -1.411  -2.635
ATOM      1  C4" DIM     1       4.038   2.897  -0.899
ATOM      2  C5" DIM     1       5.423   3.134  -0.514
ATOM      3  C6" DIM     1       5.906   4.407  -0.572
ATOM      4  C2" DIM     1       3.834   5.358  -1.353
ATOM      5  H3" DIM     1       2.391   3.938  -1.567
ATOM      6  H6" DIM     1       6.939   4.582  -0.287
ATOM      7  C7" DIM     1       6.378   2.073  -0.052
ATOM      8  N2L DIM     1       5.877   0.811  -0.017
ATOM      9  C0L DIM     1       6.675  -0.323   0.404
ATOM     10 H0L1 DIM     1       7.574  -0.400  -0.224
ATOM     11 H0L2 DIM     1       7.018  -0.097   1.419
ATOM     12  C7L DIM     1       5.941  -1.653   0.377
ATOM     13  C6L DIM     1       4.568  -1.748   0.309
ATOM     14  C8L DIM     1       6.713  -2.845   0.450
ATOM     15  C5L DIM     1       3.911  -3.009   0.314
ATOM     16  H6L DIM     1       3.966  -0.846   0.247
ATOM     17  C9L DIM     1       6.111  -4.080   0.464
ATOM     18  H8L DIM     1       7.797  -2.771   0.496
ATOM     19  C4L DIM     1       2.497  -3.120   0.241
ATOM     20 C10L DIM     1       4.698  -4.201   0.397
ATOM     21  H9L DIM     1       6.714  -4.983   0.520
ATOM     22  C3L DIM     1       1.871  -4.349   0.248
ATOM     23  H4L DIM     1       1.904  -2.213   0.175
ATOM     24 C11L DIM     1       4.033  -5.455   0.402
ATOM     25 C12L DIM     1       2.662  -5.527   0.327
ATOM     26 H11L DIM     1       4.627  -6.364   0.459
ATOM     27 H12L DIM     1       2.170  -6.496   0.324
ATOM     28  C2L DIM     1       0.360  -4.495   0.208
ATOM     29 H2L1 DIM     1       0.104  -5.416  -0.334
ATOM     30 H2L2 DIM     1      -0.044  -4.602   1.220
ATOM     31  N1L DIM     1      -0.329  -3.366  -0.387
ATOM     32  C7  DIM     1      -1.680  -3.324  -0.506
ATOM     33  C5  DIM     1      -2.233  -2.089  -1.159
ATOM     34  C4  DIM     1      -1.431  -0.975  -1.646
ATOM     35  C6  DIM     1      -3.583  -2.033  -1.313
ATOM     36  H6  DIM     1      -4.179  -2.871  -0.969
ATOM     37  C2  DIM     1      -3.560   0.132  -2.382
ATOM     38  H3  DIM     1      -1.667   0.848  -2.576
ATOM     39  C8' DIM     1      -5.679  -1.036  -2.106
ATOM     40 H8'1 DIM     1      -6.163  -1.539  -1.265
ATOM     41 H8'2 DIM     1      -6.045  -0.009  -2.161
ATOM     42  C7' DIM     1      -5.957  -1.769  -3.437
ATOM     43  N4' DIM     1      -7.266  -1.955  -3.781
ATOM     44  C3' DIM     1      -8.416  -1.477  -3.004
ATOM     45 H3'1 DIM     1      -9.162  -1.123  -3.719
ATOM     46 H3'2 DIM     1      -8.130  -0.617  -2.396
ATOM     47  C5' DIM     1      -7.549  -2.630  -5.055
ATOM     48 H5'1 DIM     1      -6.675  -3.217  -5.323
ATOM     49 H5'2 DIM     1      -8.401  -3.301  -4.916
ATOM     50  C'  DIM     1      -7.974  -1.615  -6.130
ATOM     51  C2' DIM     1      -9.024  -2.575  -2.104
```

```
ATOM     52 H2'1 DIM     1      -8.293  -2.908  -1.364
ATOM     53 H2'2 DIM     1      -9.266  -3.453  -2.713
ATOM     54  N1' DIM     1     -10.223  -2.125  -1.425
ATOM     55 H1'1 DIM     1     -11.071  -2.033  -1.971
ATOM     56  CC1 DIM     1       5.742   6.838  -1.014
ATOM     57  HC1 DIM     1       5.683   7.233  -2.030
ATOM     58  HC2 DIM     1       5.181   7.503  -0.354
ATOM     59  HC3 DIM     1       6.782   6.789  -0.695
ATOM     60  O2" DIM     1       3.161   6.308  -1.708
ATOM     61  O4" DIM     1       3.445   1.819  -0.908
ATOM     62  N3" DIM     1       3.358   4.059  -1.290
ATOM     63  O7" DIM     1       7.532   2.370   0.263
ATOM     64  H2L DIM     1       4.914   0.685  -0.320
ATOM     65  H1L DIM     1       0.196  -2.579  -0.762
ATOM     66  O7  DIM     1      -2.440  -4.214  -0.117
ATOM     67  O4  DIM     1      -0.207  -0.869  -1.607
ATOM     68  N3  DIM     1      -2.191   0.061  -2.210
ATOM     69  O2  DIM     1      -4.125   1.068  -2.917
ATOM     70  N1  DIM     1      -4.249  -0.973  -1.853
ATOM     71  O7' DIM     1      -5.017  -2.157  -4.125
ATOM     72  O1' DIM     1      -9.075  -1.072  -6.047
ATOM     73  N1" DIM     1       5.176   5.487  -0.967
ATOM      1  C4" NBC     1       4.195   3.338  -0.909
ATOM      2  C5" NBC     1       4.837   3.370   0.399
ATOM      3  C6" NBC     1       4.623   4.449   1.203
ATOM      4  C2" NBC     1       3.178   5.569  -0.369
ATOM      5  H3" NBC     1       2.943   4.485  -2.075
ATOM      6  H6" NBC     1       5.102   4.473   2.176
ATOM      7  C7" NBC     1       5.731   2.295   0.942
ATOM      8  N2L NBC     1       5.940   1.237   0.116
ATOM      9  C0L NBC     1       6.781   0.119   0.495
ATOM     10 H0L1 NBC     1       7.687   0.105  -0.128
ATOM     11 H0L2 NBC     1       7.108   0.315   1.521
ATOM     12  C7L NBC     1       6.102  -1.237   0.405
ATOM     13  C6L NBC     1       4.735  -1.385   0.322
ATOM     14  C8L NBC     1       6.922  -2.399   0.434
ATOM     15  C5L NBC     1       4.130  -2.670   0.268
ATOM     16  H6L NBC     1       4.097  -0.506   0.294
ATOM     17  C9L NBC     1       6.370  -3.657   0.390
ATOM     18  H8L NBC     1       8.001  -2.283   0.492
ATOM     19  C4L NBC     1       2.722  -2.835   0.179
ATOM     20 C10L NBC     1       4.964  -3.832   0.307
ATOM     21  H9L NBC     1       7.009  -4.536   0.413
ATOM     22  C3L NBC     1       2.146  -4.088   0.129
ATOM     23  H4L NBC     1       2.093  -1.951   0.147
ATOM     24 C11L NBC     1       4.351  -5.111   0.253
ATOM     25 C12L NBC     1       2.984  -5.235   0.164
ATOM     26 H11L NBC     1       4.981  -5.997   0.276
ATOM     27 H12L NBC     1       2.532  -6.222   0.116
ATOM     28  C2L NBC     1       0.643  -4.293   0.071
ATOM     29 H2L1 NBC     1       0.428  -5.199  -0.512
ATOM     30 H2L2 NBC     1       0.237  -4.459   1.074
ATOM     31  N1L NBC     1      -0.088  -3.168  -0.481
ATOM     32  C7  NBC     1      -1.129  -2.591   0.170
ATOM     33  C5  NBC     1      -1.786  -1.447  -0.546
ATOM     34  C4  NBC     1      -1.379  -0.937  -1.849
ATOM     35  C6  NBC     1      -2.848  -0.867   0.074
ATOM     36  H6  NBC     1      -3.165  -1.260   1.035
ATOM     37  C2  NBC     1      -3.239   0.740  -1.684
ATOM     38  H3  NBC     1      -1.911   0.517  -3.207
ATOM     39  C8' NBC     1      -4.740   0.685   0.233
ATOM     40 H8'1 NBC     1      -4.587   0.663   1.315
```

```
ATOM     41 H8'2 NBC     1      -4.886   1.721  -0.081
ATOM     42  C7' NBC     1      -5.958  -0.170  -0.182
ATOM     43  N4' NBC     1      -7.155   0.133   0.398
ATOM     44  C3' NBC     1      -7.377   1.229   1.349
ATOM     45 H3'1 NBC     1      -8.345   1.673   1.109
ATOM     46 H3'2 NBC     1      -6.628   2.011   1.205
ATOM     47  C5' NBC     1      -8.329  -0.655  -0.002
ATOM     48 H5'1 NBC     1      -7.975  -1.603  -0.399
ATOM     49 H5'2 NBC     1      -8.943  -0.846   0.883
ATOM     50  C'  NBC     1      -9.215   0.138  -0.977
ATOM     51  C2' NBC     1      -7.360   0.756   2.818
ATOM     52 H2'1 NBC     1      -6.382   0.343   3.075
ATOM     53 H2'2 NBC     1      -8.081  -0.061   2.947
ATOM     54  N1' NBC     1      -7.687   1.824   3.743
ATOM     55 H1'1 NBC     1      -8.653   2.125   3.798
ATOM     56  O2" NBC     1       2.472   6.507  -0.690
ATOM     57  O4" NBC     1       4.281   2.449  -1.755
ATOM     58  N3" NBC     1       3.407   4.466  -1.174
ATOM     59  O7" NBC     1       6.223   2.404   2.067
ATOM     60  H2L NBC     1       5.496   1.267  -0.798
ATOM     61  H1L NBC     1       0.149  -2.798  -1.399
ATOM     62  O7  NBC     1      -1.539  -2.949   1.278
ATOM     63  O4  NBC     1      -0.461  -1.351  -2.554
ATOM     64  N3  NBC     1      -2.153   0.145  -2.296
ATOM     65  O2  NBC     1      -3.869   1.660  -2.174
ATOM     66  N1  NBC     1      -3.536   0.202  -0.420
ATOM     67  O7' NBC     1      -5.813  -1.077  -0.997
ATOM     68  O1' NBC     1      -9.866   1.096  -0.560
ATOM     69  N1" NBC     1       3.839   5.511   0.867
ATOM     70  C1A NBC     1       3.682   6.620   1.820
ATOM     71 H1A1 NBC     1       4.102   6.339   2.763
ATOM     72  C2A NBC     1       4.404   7.871   1.293
ATOM     73 H2A1 NBC     1       4.397   8.658   2.049
ATOM     74 H2A2 NBC     1       3.890   8.254   0.401
ATOM     75  N3A NBC     1       5.798   7.606   0.987
ATOM     76 H3A1 NBC     1       5.881   7.317   0.033
ATOM     77 H1A2 NBC     1       2.642   6.836   1.944
ATOM     78  CCC NBC     1       6.588   8.827   1.199
ATOM     79  C2C NBC     1       8.061   8.726   1.647
ATOM     80  OCC NBC     1       6.062   9.956   1.014
ATOM     81  N0C NBC     1       8.805   9.971   1.355
ATOM     82 H2C1 NBC     1       8.049   8.566   2.731
ATOM     83 H2C2 NBC     1       8.542   7.848   1.200
ATOM     84  C1C NBC     1       9.448   9.969   0.005
ATOM     85  C3C NBC     1       9.691  10.451   2.460
ATOM     86  C9C NBC     1       9.775  11.389  -0.501
ATOM     87 H1C1 NBC     1      10.367   9.368  -0.001
ATOM     88 H1C2 NBC     1       8.744   9.491  -0.686
ATOM     89  C4C NBC     1       9.010  11.521   3.338
ATOM     90 H3C1 NBC     1      10.604  10.851   2.016
ATOM     91 H3C2 NBC     1      10.010   9.623   3.101
ATOM     92  N8C NBC     1       8.623  12.350  -0.348
ATOM     93 H9C1 NBC     1      10.638  11.800   0.027
ATOM     94 H9C2 NBC     1      10.063  11.321  -1.554
ATOM     95  N5C NBC     1       8.354  12.604   2.522
ATOM     96 H4C1 NBC     1       9.754  11.948   4.020
ATOM     97 H4C2 NBC     1       8.224  11.072   3.953
ATOM     98  C7C NBC     1       8.979  13.658   0.297
ATOM     99 H8C1 NBC     1       8.238  12.553  -1.270
ATOM    100  C6C NBC     1       9.320  13.499   1.791
ATOM    101 H5C1 NBC     1       7.796  13.170   3.159
ATOM    102 H7C1 NBC     1       8.114  14.317   0.172
```

```
ATOM    103 H7C2 NBC     1       9.826  14.130  -0.213
ATOM    104 H6C1 NBC     1       9.328  14.491   2.250
ATOM    105 H6C2 NBC     1      10.327  13.095   1.915
ATOM    106  Zn  NBC     1       7.301  11.490   1.059
```

8.3.5. pna.amber94.bnd

```
# ADE
    5    6
    5    7
    8    9
   11   12
   14   15
   19   20
   19   21
   23   24
   27   28
   29   30
   29   31
    1    2
    1    3
    1    4
    4    5
    5    8
    8   10
    8   27
   10   11
   11   13
   11   29
   13   14
   13   26
   14   16
   16   17
   17   18
   17   26
   18   19
   18   22
   22   23
   23   25
   25   26
   27   29
   27   32
# GUA
    5    6
    5    7
    8    9
   11   12
   14   15
   20   21
   23   24
   23   25
   28   29
   30   31
   30   32
    1    2
    1    3
    1    4
    4    5
```

```
     5    8
     8   10
     8   28
    10   11
    11   13
    11   30
    13   14
    13   27
    14   16
    16   17
    17   18
    17   27
    18   19
    18   20
    20   22
    22   23
    22   26
    26   27
    28   30
    28   33
# THY
     5    6
     5    7
     8    9
    11   12
    14   15
    17   18
    17   19
    17   20
    23   24
    27   28
    29   30
    29   31
     1    2
     1    3
     1    4
     4    5
     5    8
     8   10
     8   27
    10   11
    11   13
    11   29
    13   14
    13   25
    14   16
    16   17
    16   21
    21   22
    21   23
    23   25
    25   26
    27   29
    27   32
# CYT
     5    6
     5    7
     8    9
    11   12
    14   15
    16   17
    19   20
```

```
   19   21
   25   26
   27   28
   27   29
    1    2
    1    3
    1    4
    4    5
    5    8
    8   10
    8   25
   10   11
   11   13
   11   27
   13   14
   13   23
   14   16
   16   18
   18   19
   18   22
   22   23
   23   24
   25   27
   25   30
# APU
    5    6
    5    7
    8    9
   11   12
   15   16
   17   18
   17   19
    1    2
    1    3
    1    4
    4    5
    5    8
    8   10
    8   15
   10   11
   11   13
   11   29
   13   14
   15   17
   15   20
```

8.3.6. pna.amber94.qr

```
ATOM      1 P    ADE     1       1.166   2.000
ATOM      2 O1P  ADE     1      -0.776   1.400
ATOM      3 O2P  ADE     1      -0.776   1.400
ATOM      4 O5'  ADE     1      -0.495   1.400
ATOM      5 C5'  ADE     1      -0.007   1.700
ATOM      6 H5'1 ADE     1       0.075   1.000
ATOM      7 H5'2 ADE     1       0.075   1.000
ATOM      8 C4'  ADE     1       0.163   1.700
ATOM      9 H4'  ADE     1       0.118   1.000
ATOM     10 O4'  ADE     1      -0.369   1.400
```

```
ATOM     11 C1'  ADE     1       0.043   1.700
ATOM     12 H1'  ADE     1       0.184   1.000
ATOM     13 N9   ADE     1      -0.027   1.500
ATOM     14 C8   ADE     1       0.161   1.700
ATOM     15 H8   ADE     1       0.188   1.000
ATOM     16 N7   ADE     1      -0.618   1.500
ATOM     17 C5   ADE     1       0.072   1.700
ATOM     18 C6   ADE     1       0.690   1.700
ATOM     19 N6   ADE     1      -0.912   1.500
ATOM     20 H61  ADE     1       0.417   1.000
ATOM     21 H62  ADE     1       0.417   1.000
ATOM     22 N1   ADE     1      -0.762   1.500
ATOM     23 C2   ADE     1       0.572   1.700
ATOM     24 H2   ADE     1       0.060   1.000
ATOM     25 N3   ADE     1      -0.742   1.500
ATOM     26 C4   ADE     1       0.380   1.700
ATOM     27 C3'  ADE     1       0.071   1.700
ATOM     28 H3'  ADE     1       0.099   1.000
ATOM     29 C2'  ADE     1      -0.085   1.700
ATOM     30 H2'1 ADE     1       0.072   1.000
ATOM     31 H2'2 ADE     1       0.072   1.000
ATOM     32 O3'  ADE     1      -0.523   1.400
ATOM      1 P    GUA     1       1.166   2.000
ATOM      2 O1P  GUA     1      -0.776   1.400
ATOM      3 O2P  GUA     1      -0.776   1.400
ATOM      4 O5'  GUA     1      -0.495   1.400
ATOM      5 C5'  GUA     1      -0.007   1.700
ATOM      6 H5'1 GUA     1       0.075   1.000
ATOM      7 H5'2 GUA     1       0.075   1.000
ATOM      8 C4'  GUA     1       0.163   1.700
ATOM      9 H4'  GUA     1       0.118   1.000
ATOM     10 O4'  GUA     1      -0.369   1.400
ATOM     11 C1'  GUA     1       0.036   1.700
ATOM     12 H1'  GUA     1       0.175   1.000
ATOM     13 N9   GUA     1       0.058   1.500
ATOM     14 C8   GUA     1       0.074   1.700
ATOM     15 H8   GUA     1       0.200   1.000
ATOM     16 N7   GUA     1      -0.573   1.500
ATOM     17 C5   GUA     1       0.199   1.700
ATOM     18 C6   GUA     1       0.492   1.700
ATOM     19 O6   GUA     1      -0.570   1.400
ATOM     20 N1   GUA     1      -0.505   1.500
ATOM     21 H1   GUA     1       0.352   1.000
ATOM     22 C2   GUA     1       0.743   1.700
ATOM     23 N2   GUA     1      -0.923   1.500
ATOM     24 H21  GUA     1       0.423   1.000
ATOM     25 H22  GUA     1       0.423   1.000
ATOM     26 N3   GUA     1      -0.664   1.500
ATOM     27 C4   GUA     1       0.181   1.700
ATOM     28 C3'  GUA     1       0.071   1.700
ATOM     29 H3'  GUA     1       0.099   1.000
ATOM     30 C2'  GUA     1      -0.085   1.700
ATOM     31 H2'1 GUA     1       0.072   1.000
ATOM     32 H2'2 GUA     1       0.072   1.000
ATOM     33 O3'  GUA     1      -0.523   1.400
ATOM      1 P    THY     1       1.166   2.000
ATOM      2 O1P  THY     1      -0.776   1.400
ATOM      3 O2P  THY     1      -0.776   1.400
ATOM      4 O5'  THY     1      -0.495   1.400
ATOM      5 C5'  THY     1      -0.007   1.700
ATOM      6 H5'1 THY     1       0.075   1.000
ATOM      7 H5'2 THY     1       0.075   1.000
```

```
ATOM      8 C4'  THY     1       0.163   1.700
ATOM      9 H4'  THY     1       0.118   1.000
ATOM     10 O4'  THY     1      -0.369   1.400
ATOM     11 C1'  THY     1       0.068   1.700
ATOM     12 H1'  THY     1       0.180   1.000
ATOM     13 N1   THY     1      -0.024   1.500
ATOM     14 C6   THY     1      -0.221   1.700
ATOM     15 H6   THY     1       0.261   1.000
ATOM     16 C5   THY     1       0.003   1.700
ATOM     17 C7   THY     1      -0.227   1.700
ATOM     18 H71  THY     1       0.077   1.000
ATOM     19 H72  THY     1       0.077   1.000
ATOM     20 H73  THY     1       0.077   1.000
ATOM     21 C4   THY     1       0.519   1.700
ATOM     22 O4   THY     1      -0.556   1.400
ATOM     23 N3   THY     1      -0.434   1.500
ATOM     24 H3   THY     1       0.342   1.000
ATOM     25 C2   THY     1       0.568   1.700
ATOM     26 O2   THY     1      -0.588   1.400
ATOM     27 C3'  THY     1       0.071   1.700
ATOM     28 H3'  THY     1       0.099   1.000
ATOM     29 C2'  THY     1      -0.085   1.700
ATOM     30 H2'1 THY     1       0.072   1.000
ATOM     31 H2'2 THY     1       0.072   1.000
ATOM     32 O3'  THY     1      -0.523   1.400
ATOM      1 P    CYT     1       1.166   2.000
ATOM      2 O1P  CYT     1      -0.776   1.400
ATOM      3 O2P  CYT     1      -0.776   1.400
ATOM      4 O5'  CYT     1      -0.495   1.400
ATOM      5 C5'  CYT     1      -0.007   1.700
ATOM      6 H5'1 CYT     1       0.075   1.000
ATOM      7 H5'2 CYT     1       0.075   1.000
ATOM      8 C4'  CYT     1       0.163   1.700
ATOM      9 H4'  CYT     1       0.118   1.000
ATOM     10 O4'  CYT     1      -0.369   1.400
ATOM     11 C1'  CYT     1      -0.012   1.700
ATOM     12 H1'  CYT     1       0.196   1.000
ATOM     13 N1   CYT     1      -0.034   1.500
ATOM     14 C6   CYT     1      -0.018   1.700
ATOM     15 H6   CYT     1       0.229   1.000
ATOM     16 C5   CYT     1      -0.522   1.700
ATOM     17 H5   CYT     1       0.186   1.000
ATOM     18 C4   CYT     1       0.844   1.700
ATOM     19 N4   CYT     1      -0.977   1.500
ATOM     20 H41  CYT     1       0.431   1.000
ATOM     21 H42  CYT     1       0.431   1.000
ATOM     22 N3   CYT     1      -0.775   1.500
ATOM     23 C2   CYT     1       0.796   1.700
ATOM     24 O2   CYT     1      -0.655   1.400
ATOM     25 C3'  CYT     1       0.071   1.700
ATOM     26 H3'  CYT     1       0.099   1.000
ATOM     27 C2'  CYT     1      -0.085   1.700
ATOM     28 H2'1 CYT     1       0.072   1.000
ATOM     29 H2'2 CYT     1       0.072   1.000
ATOM     30 O3'  CYT     1      -0.523   1.400
ATOM      1 P    APU     1       1.166   2.000
ATOM      2 O1P  APU     1      -0.776   1.400
ATOM      3 O2P  APU     1      -0.776   1.400
ATOM      4 O5'  APU     1      -0.495   1.400
ATOM      5 C5'  APU     1      -0.007   1.700
ATOM      6 H5'1 APU     1       0.075   1.000
ATOM      7 H5'2 APU     1       0.075   1.000
```

```
ATOM      8 C4'  APU     1       0.163   1.700
ATOM      9 H4'  APU     1       0.118   1.000
ATOM     10 O4'  APU     1      -0.369   1.400
ATOM     11 C1'  APU     1       0.068   1.700
ATOM     12 H1'  APU     1       0.180   1.000
ATOM     13 O1'  APU     1      -0.600   1.400
ATOM     14 HO1' APU     1       0.452   1.000
ATOM     27 C3'  APU     1       0.071   1.700
ATOM     28 H3'  APU     1       0.099   1.000
ATOM     29 C2'  APU     1      -0.085   1.700
ATOM     30 H2'1 APU     1       0.072   1.000
ATOM     31 H2'2 APU     1       0.072   1.000
ATOM     32 O3'  APU     1      -0.523   1.400
```

8.3.7. pna.amber94.chi

```
# ADE
N9 O4' H1' C2'
C1' H2'2 H2'1 C3'
C2' O3' H3' C4'
C3' O4' H4' C5'
C4' H5'1 H5'2 O5'
N9 C1' C8 C6
C8 N9 N7 H8
C4 N9 N3 C5
C5 N7 C4 C6
C2 N3 N1 H2
C6 C5 N1 N6
N6 C6 H61 H62
# CYT
N1 O4' H1' C2'
C1' H2'2 H2'1 C3'
C2' O3' H3' C4'
C3' O4' H4' C5'
C4' H5'1 H5'2 O5'
N1 C1' C6 C2
C2 N1 N3 O2
C4 N3 C5 N4
N4 C4 H41 H42
C5 C4 C6 H5
C6 C5 N1 H6
# GUA
N9 O4' H1' C2'
C1' H2'2 H2'1 C3'
C2' O3' H3' C4'
C3' O4' H4' C5'
C4' H5'1 H5'2 O5'
N9 C1' C8 C6
C8 N9 N7 H8
C4 N9 N3 C5
C5 N7 C4 C6
C2 N3 N1 N2
N2 C2 H21 H22
C6 C5 N1 O6
# THY
N1 O4' H1' C2'
C1' H2'2 H2'1 C3'
C2' O3' H3' C4'
```

```
C3' O4' H4' C5'
C4' H5'1 H5'2 O5'
N1 C1' C6 C2
C2 N1 N3 O2
N3 C2 C4 H3
C4 N3 C5 O4
C5 C4 C6 C7
C6 C5 N1 H6
# APU
```

8.3.8. apna.dat file

```
P       8.92      69.20      -4.0600   A
O1P     7.69      73.90      -4.5300   A
O2P     9.96      67.90      -5.1000   A
O3'     9.58      41.43      -5.4300   A
O5'     8.70      60.10      -3.3300   A
C1'     8.59      35.9       -2.04     A
C2'     8.81      33.4       -3.50     A
C3'     8.91      42.2       -4.17     A
C4'     9.77      46.6       -3.13     A
C5'     9.91      55.4       -3.1900   A
O1'     9.2200    44.00      -1.8600   A
N6      3.19      14.8       -0.13     A
C6      4.49      16.7       -0.43     A
C5      5.22      29.8       -0.91     A
N7      5.18      44.8       -1.19     A
C8      6.40      45.7       -1.61     A
N9      7.16      36.1       -1.62     A
C4      6.52      26.7       -1.17     A
N3      7.25      17.3       -1.02     A
C2      6.65       8.4       -0.56     A
N1      5.40       5.3       -0.26     A
P       8.92      69.20      -4.0600   G
O1P     7.69      73.90      -4.5300   G
O2P     9.96      67.90      -5.1000   G
O3'     9.58      41.43      -5.4300   G
O5'     8.70      60.10      -3.3300   G
C1'     8.59      35.9       -2.04     G
C2'     8.81      33.4       -3.50     G
C3'     8.91      42.2       -4.17     G
C4'     9.77      46.6       -3.13     G
C5'     9.91      55.4       -3.1900   G
O1'     9.2200    44.00      -1.8600   G
N2      7.71       0.7       -0.37     G
O6      3.27      13.2       -0.11     G
C6      4.44      16.4       -0.41     G
C5      5.20      29.4       -0.90     G
N7      5.14      44.6       -1.17     G
C8      6.38      45.7       -1.60     G
N9      7.16      36.1       -1.62     G
C4      6.52      26.6       -1.17     G
N3      7.30      17.4       -1.03     G
C2      6.78       8.1       -0.57     G
N1      5.46       5.3       -0.27     G
P       8.92      69.20      -4.0600   T
O1P     7.69      73.90      -4.5300   T
O2P     9.96      67.90      -5.1000   T
```

O3'	9.58	41.43	-5.4300	T	
O5'	8.70	60.10	-3.3300	T	
C1'	8.59	35.9	-2.04	T	
C2'	8.81	33.4	-3.50	T	
C3'	8.91	42.2	-4.17	T	
C4'	9.77	46.6	-3.13	T	
C5'	9.91	55.4	-3.1900	T	
O1'	9.2200	44.00	-1.8600	T	
O4	3.30	35.5	-0.47	T	
C4	4.48	36.4	-0.83	T	
N3	5.39	24.3	-0.82	T	
C2	6.71	26.0	-1.20	T	
O2	7.56	19.0	-1.15	T	
N1	7.16	36.1	-1.62	T	
C6	6.55	46.4	-1.67	T	
C5	5.30	49.3	-1.30	T	
C7	4.98	65.7	-1.33	T	
P	8.92	69.20	-4.0600	C	
O1P	7.69	73.90	-4.5300	C	
O2P	9.96	67.90	-5.1000	C	
O3'	9.58	41.43	-5.4300	C	
O5'	8.70	60.10	-3.3300	C	
C1'	8.59	35.9	-2.04	C	
C2'	8.81	33.4	-3.50	C	
C3'	8.91	42.2	-4.17	C	
C4'	9.77	46.6	-3.13	C	
C5'	9.91	55.4	-3.1900	C	
O1'	9.2200	44.00	-1.8600	C	
N4	3.30	34.5	-0.46	C	
C4	4.57	35.5	-0.84	C	
N3	5.41	23.4	-0.80	C	
C2	6.69	25.9	-1.19	C	
O2	7.61	19.2	-1.17	C	
N1	7.16	36.1	-1.62	C	
C6	6.55	46.3	-1.66	C	
C5	5.28	49.1	-1.29	C	
N6	3.19	14.8	-0.1300		X
C6	4.49	16.7	-0.4300		X
C5	5.22	29.8	-0.9100		X
N7	5.18	44.8	-1.1900		X
C8	6.40	45.7	-1.6100		X
N9	7.16	36.1	-1.6200		X
C4	6.52	26.7	-1.1700		X
N3	7.25	17.3	-1.0200		X
C2	6.65	8.4	-0.5600		X
N1	5.40	5.3	-0.2600		X
O4	3.30	35.5	-0.4700		Y
C4	4.48	36.4	-0.8300		Y
N3	5.39	24.3	-0.8200		Y
C2	6.71	26.0	-1.2000		Y
O2	7.56	19.0	-1.1500		Y
N1	7.16	36.1	-1.6200		Y
C6	6.55	46.4	-1.6700		Y
C5	5.30	49.3	-1.3000		Y
C7	4.98	65.7	-1.3300		Y
N4	3.30	34.5	-0.4600		J
C4	4.57	35.5	-0.8400		J
N3	5.41	23.4	-0.8000		J
C2	6.69	25.9	-1.1900		J
O2	7.61	19.2	-1.1700		J
N1	7.16	36.1	-1.6200		J
C6	6.55	46.3	-1.6600		J

```
C5      5.28      49.1       -1.2900        J
N2      7.71       0.7       -0.3700        K
O6      3.27      13.2       -0.1100        K
C6      4.44      16.4       -0.4100        K
C5      5.20      29.4       -0.9000        K
N7      5.14      44.6       -1.1700        K
C8      6.38      45.7       -1.6000        K
N9      7.16      36.1       -1.6200        K
C4      6.52      26.6       -1.1700        K
N3      7.30      17.4       -1.0300        K
C2      6.78       8.1       -0.5700        K
N1      5.46       5.3       -0.2700        K
O4      3.30      35.5       -0.4700        W
C4      4.48      36.4       -0.8300        W
N3      5.39      24.3       -0.8200        W
C2      6.71      26.0       -1.2000        W
O2      7.56      19.0       -1.1500        W
N1      7.16      36.1       -1.6200        W
C6      6.55      46.4       -1.6700        W
C5      5.30      49.3       -1.3000        W
C7      4.98      65.7       -1.3300        W
O4      3.30      35.5       -0.4700        Q
C4      4.48      36.4       -0.8300        Q
N3      5.39      24.3       -0.8200        Q
C2      6.71      26.0       -1.2000        Q
O2      7.56      19.0       -1.1500        Q
N1      7.16      36.1       -1.6200        Q
C6      6.55      46.4       -1.6700        Q
C5      5.30      49.3       -1.3000        Q
C7      4.98      65.7       -1.3300        Q
```

8.3.9. leaprc.pna

```
logFile leap.log
#
# ----- leaprc for loading the ff10 force field
# ----- NOTE: this is designed for PDB format 3!
#       ff10 =  ff99SB for proteins; ff99bsc0 for DNA; ff99sbsc_chiOL3 for RNA
#       phosphoaa parms from N. Homeyer, A.H.C. Horn, H. Lanig, H. Sticht
#           J. Mol. Model. 2006, 12, 281-289.  OP vdW parameters modified
#           by T. Steinbrecher and J. Latzer.
#
#     load atom type hybridizations
#
addAtomTypes {
        { "H"   "H" "sp3" }
      { "HO"  "H" "sp3" }
      { "HS"  "H" "sp3" }
      { "H1"  "H" "sp3" }
      { "H2"  "H" "sp3" }
      { "H3"  "H" "sp3" }
      { "H4"  "H" "sp3" }
      { "H5"  "H" "sp3" }
      { "HW"  "H" "sp3" }
      { "HC"  "H" "sp3" }
      { "HA"  "H" "sp3" }
      { "HP"  "H" "sp3" }
      { "OH"  "O" "sp3" }
```

```
      { "OS"  "O" "sp3" }
      { "O"   "O" "sp2" }
      { "O2"  "O" "sp2" }
      { "OW"  "O" "sp3" }
      { "CT"  "C" "sp3" }
      { "CH"  "C" "sp3" }
      { "C2"  "C" "sp3" }
      { "C3"  "C" "sp3" }
      { "C"   "C" "sp2" }
      { "C*"  "C" "sp2" }
      { "CA"  "C" "sp2" }
      { "CB"  "C" "sp2" }
      { "CC"  "C" "sp2" }
      { "CN"  "C" "sp2" }
      { "CM"  "C" "sp2" }
      { "CK"  "C" "sp2" }
      { "CQ"  "C" "sp2" }
      { "CD"  "C" "sp2" }
      { "CE"  "C" "sp2" }
      { "CF"  "C" "sp2" }
      { "CP"  "C" "sp2" }
      { "CI"  "C" "sp2" }
      { "CJ"  "C" "sp2" }
      { "CW"  "C" "sp2" }
      { "CV"  "C" "sp2" }
      { "CR"  "C" "sp2" }
      { "CA"  "C" "sp2" }
      { "CY"  "C" "sp2" }
      { "C0"  "C" "sp2" }
      { "MG"  "Mg" "sp3" }
      { "N"   "N" "sp2" }
      { "NA"  "N" "sp2" }
      { "N2"  "N" "sp2" }
      { "N*"  "N" "sp2" }
      { "NP"  "N" "sp2" }
      { "NQ"  "N" "sp2" }
      { "NB"  "N" "sp2" }
      { "NC"  "N" "sp2" }
      { "NT"  "N" "sp3" }
      { "N3"  "N" "sp3" }
      { "S"   "S" "sp3" }
      { "SH"  "S" "sp3" }
      { "P"   "P" "sp3" }
      { "LP"  ""  "sp3" }
      { "F"   "F" "sp3" }
      { "CL"  "Cl" "sp3" }
      { "BR"  "Br" "sp3" }
      { "I"   "I"  "sp3" }
      { "FE"  "Fe" "sp3" }
      { "EP"  ""  "sp3" }


}



#
#     Load the main parameter set.
#
parm10 = loadamberparams parm10.dat
FRCMOD = loadAmberParams frcmod.pna
```

```
source leaprc.ff99bsc0
frcmod14SB = loadamberparams frcmod.ff14SB
source leaprc.gaff
#
#     Load main chain and terminating amino acid libraries, nucleic acids
#
loadOff amino10.lib
loadOff aminoct10.lib
loadOff aminont10.lib
loadOff phosphoaa10.lib
loadOff nucleic10.lib
loadOff RNA_CI.lib


#
#       Load water and ions
#
loadOff ions94.lib
loadOff solvents.lib
loadOff atomic_ions.lib
loadOff solvents.lib
HOH = TP3
WAT = TP3

# PNA residue

APN  = loadmol3 tripos1-mol3.mol2
NAP = loadmol3 tripos2-mol3.mol2
CAP = loadmol3 tripos3-mol3.mol2
MAPN = loadmol3 tripos4-mol3.mol2

CPN  = loadmol3 tripos5-mol3.mol2
NCP = loadmol3 tripos6-mol3.mol2
CCP = loadmol3 tripos7-mol3.mol2
MCPN = loadmol3 tripos8-mol3.mol2

GPN  = loadmol3 tripos9-mol3.mol2
NGP = loadmol3 tripos10-mol3.mol2
CGP = loadmol3 tripos11-mol3.mol2
MGPN = loadmol3 tripos12-mol3.mol2

IPN  = loadmol3 tripos13-mol3.mol2
NIP = loadmol3 tripos14-mol3.mol2
CIP = loadmol3 tripos15-mol3.mol2
MIPN = loadmol3 tripos16-mol3.mol2

TPN  = loadmol3 tripos17-mol3.mol2
NTP = loadmol3 tripos18-mol3.mol2
CTP = loadmol3 tripos19-mol3.mol2
MTP = loadmol3 tripos20-mol3.mol2

UPN  = loadmol3 tripos21-mol3.mol2
NUP = loadmol3 tripos22-mol3.mol2
CUP = loadmol3 tripos23-mol3.mol2
MUPN = loadmol3 tripos24-mol3.mol2

DIM = loadmol3 dimeric_charged.mol2

NBC = loadmol3 NBC.mol2

charge APN
charge CPN
```

```
charge GPN
charge IPN
charge TPN
charge UPN

charge NAP
charge NCP
charge NGP
charge NIP
charge NTP
charge NUP

charge CAP
charge CCP
charge CGP
charge CIP
charge CTP
charge CUP

charge MAPN
charge MCPN
charge MGPN
charge MIPN
charge MTPN
charge MUPN

charge DIM

charge NBC

#
#     Define the PDB name map for the amino acids and nucleic acids
#
addPdbResMap {
  { 0 "ALA" "NALA" } { 1 "ALA" "CALA" }
  { 0 "ARG" "NARG" } { 1 "ARG" "CARG" }
  { 0 "ASN" "NASN" } { 1 "ASN" "CASN" }
  { 0 "ASP" "NASP" } { 1 "ASP" "CASP" }
  { 0 "CYS" "NCYS" } { 1 "CYS" "CCYS" }
  { 0 "CYX" "NCYX" } { 1 "CYX" "CCYX" }
  { 0 "GLN" "NGLN" } { 1 "GLN" "CGLN" }
  { 0 "GLU" "NGLU" } { 1 "GLU" "CGLU" }
  { 0 "GLY" "NGLY" } { 1 "GLY" "CGLY" }
  { 0 "HID" "NHID" } { 1 "HID" "CHID" }
  { 0 "HIE" "NHIE" } { 1 "HIE" "CHIE" }
  { 0 "HIP" "NHIP" } { 1 "HIP" "CHIP" }
  { 0 "ILE" "NILE" } { 1 "ILE" "CILE" }
  { 0 "LEU" "NLEU" } { 1 "LEU" "CLEU" }
  { 0 "LYS" "NLYS" } { 1 "LYS" "CLYS" }
  { 0 "MET" "NMET" } { 1 "MET" "CMET" }
  { 0 "PHE" "NPHE" } { 1 "PHE" "CPHE" }
  { 0 "PRO" "NPRO" } { 1 "PRO" "CPRO" }
  { 0 "SER" "NSER" } { 1 "SER" "CSER" }
  { 0 "THR" "NTHR" } { 1 "THR" "CTHR" }
  { 0 "TRP" "NTRP" } { 1 "TRP" "CTRP" }
  { 0 "TYR" "NTYR" } { 1 "TYR" "CTYR" }
  { 0 "VAL" "NVAL" } { 1 "VAL" "CVAL" }
  { 0 "HIS" "NHIS" } { 1 "HIS" "CHIS" }
  { 0 "G" "G5"  } { 1 "G" "G3"  }
  { 0 "A" "A5"  } { 1 "A" "A3"  }
  { 0 "C" "C5"  } { 1 "C" "C3"  }
  { 0 "U" "U5"  } { 1 "U" "U3"  }
```

```
  { 0 "DG" "DG5"  } { 1 "DG" "DG3"  }
  { 0 "DA" "DA5"  } { 1 "DA" "DA3"  }
  { 0 "DC" "DC5"  } { 1 "DC" "DC3"  }
  { 0 "DT" "DT5"  } { 1 "DT" "DT3"  }
#  some old Amber residue names for RNA:
  { 0  "RA5" "A5" } { 1 "RA3" "A3"} {"RA" "A" }
  { 0  "RC5" "C5" } { 1 "RC3" "C3"} {"RC" "C" }
  { 0  "RG5" "G5" } { 1 "RG3" "G3"} {"RG" "G" }
  { 0  "RU5" "U5" } { 1 "RU3" "U3"} {"RU" "U" }
#  some really old Amber residue names, assuming DNA:
  { 0 "GUA" "DG5"  } { 1 "GUA" "DG3"  } { "GUA" "DG" }
  { 0 "ADE" "DA5"  } { 1 "ADE" "DA3"  } { "ADE" "DA" }
  { 0 "CYT" "DC5"  } { 1 "CYT" "DC3"  } { "CYT" "DC" }
  { 0 "THY" "DT5"  } { 1 "THY" "DT3"  } { "THY" "DT" }
#  uncomment out the following if you have this old style RNA files:
# { 0 "GUA" "G5"  } { 1 "GUA" "G3"  } { "GUA" "G" }
# { 0 "ADE" "A5"  } { 1 "ADE" "A3"  } { "ADE" "A" }
# { 0 "CYT" "C5"  } { 1 "CYT" "C3"  } { "CYT" "C" }
# { 0 "THY" "T5"  } { 1 "THY" "T3"  } { "THY" "T" }

   { 0 "APN" "NAP" } { 1 "APN" "CAP" }
   { 0 "CPN" "NCP" } { 1 "CPN" "CCP" }
   { 0 "GPN" "NGP" } { 1 "GPN" "CGP" }
   { 0 "IPN" "NIP" } { 1 "IPN" "CIP" }
   { 0 "TPN" "NTP" } { 1 "TPN" "CTP" }
   { 0 "UPN" "NUP" } { 1 "UPN" "CUP" }


}

#  try to be good about reading in really old atom names as well:
addPdbAtomMap {
  { "O5*" "O5'" }
  { "C5*" "C5'" }
  { "C4*" "C4'" }
  { "O4*" "O4'" }
  { "C3*" "C3'" }
  { "O3*" "O3'" }
  { "C2*" "C2'" }
  { "O2*" "O2'" }
  { "C1*" "C1'" }
  { "C5M" "C7"  }
  { "H1*" "H1'" }
  { "H2*1" "H2'" }
  { "H2*2" "H2''" }
  { "H2'1" "H2'" }
  { "H2'2" "H2''" }
  { "H3*" "H3'" }
  { "H4*" "H4'" }
  { "H5*1" "H5'" }
  { "H5*2" "H5''" }
  { "H5'1" "H5'" }
  { "H5'2" "H5''" }
  { "HO'2" "HO2'" }
  { "H5T"  "HO5'" }
  { "H3T"  "HO3'" }
  { "O1'" "O4'" }
  { "OA"  "OP1" }
  { "OB"  "OP2" }
  { "O1P" "OP1" }
  { "O2P" "OP2" }
}
```

```
#
# assume that most often proteins use HIE
#
NHIS = NHIE
HIS = HIE
CHIS = CHIE
```

8.3.10. script.sh

```
logfile 0PNA-from-mol3-files.log

verbosity 2

addAtomTypes {
      { "H"   "H" "sp3" }
      { "HO"  "H" "sp3" }
      { "HS"  "H" "sp3" }
      { "H1"  "H" "sp3" }
      { "H2"  "H" "sp3" }
      { "H3"  "H" "sp3" }
      { "H4"  "H" "sp3" }
      { "H5"  "H" "sp3" }
      { "HW"  "H" "sp3" }
      { "HC"  "H" "sp3" }
      { "HA"  "H" "sp3" }
      { "HP"  "H" "sp3" }
      { "OH"  "O" "sp3" }
      { "OS"  "O" "sp3" }
      { "O"   "O" "sp2" }
      { "O2"  "O" "sp2" }
      { "OW"  "O" "sp3" }
      { "CT"  "C" "sp3" }
      { "CH"  "C" "sp3" }
      { "C2"  "C" "sp3" }
      { "C3"  "C" "sp3" }
      { "C"   "C" "sp2" }
      { "C*"  "C" "sp2" }
      { "CA"  "C" "sp2" }
      { "CB"  "C" "sp2" }
      { "CC"  "C" "sp2" }
      { "CN"  "C" "sp2" }
      { "CM"  "C" "sp2" }
      { "CK"  "C" "sp2" }
      { "CQ"  "C" "sp2" }
      { "CD"  "C" "sp2" }
      { "CE"  "C" "sp2" }
      { "CF"  "C" "sp2" }
      { "CP"  "C" "sp2" }
      { "CI"  "C" "sp2" }
      { "CJ"  "C" "sp2" }
      { "CW"  "C" "sp2" }
```

```
      { "CV"  "C" "sp2" }
      { "CR"  "C" "sp2" }
      { "CA"  "C" "sp2" }
      { "CY"  "C" "sp2" }
      { "C0"  "C" "sp2" }
      { "MG"  "Mg" "sp3" }
      { "N"   "N" "sp2" }
      { "NA"  "N" "sp2" }
      { "N2"  "N" "sp2" }
      { "N*"  "N" "sp2" }
      { "NP"  "N" "sp2" }
      { "NQ"  "N" "sp2" }
      { "NB"  "N" "sp2" }
      { "NC"  "N" "sp2" }
      { "NT"  "N" "sp3" }
      { "N3"  "N" "sp3" }
      { "S"   "S" "sp3" }
      { "SH"  "S" "sp3" }
      { "P"   "P" "sp3" }
      { "LP"  ""  "sp3" }
      { "F"   "F" "sp3" }
      { "CL"  "Cl" "sp3" }
      { "BR"  "Br" "sp3" }
      { "I"   "I"  "sp3" }
      { "FE"  "Fe" "sp3" }
      { "EP"  ""  "sp3" }
}

addPdbResMap {
   { 0 "APN" "NAP" } { 1 "APN" "CAP" }
   { 0 "CPN" "NCP" } { 1 "CPN" "CCP" }
   { 0 "GPN" "NGP" } { 1 "GPN" "CGP" }
   { 0 "IPN" "NIP" } { 1 "IPN" "CIP" }
   { 0 "TPN" "NTP" } { 1 "TPN" "CTP" }
   { 0 "UPN" "NUP" } { 1 "UPN" "CUP" }
}

#parm99 = loadamberparams parm99.dat
#frcmod99SB = loadamberparams frcmod.ff99SB
#FRCMOD = loadAmberParams frcmod.pna
FRCMOD = loadAmberParams dimeric.frcmod
source leaprc.ff99bsc0
source leaprc.gaff
source leaprc.pna
#frcmodbsc0 = loadamberparams frcmod.parmbsc0

loadOff ions94.lib
loadOff solvents.lib
loadOff RNA_CI.lib
HOH = TP3
WAT = TP3

# See http://q4md-forcefieldtools.org/Tutorial/leap-mol3.php
APN  = loadmol3 tripos1-mol3.mol2
```

```
NAP = loadmol3 tripos2-mol3.mol2
CAP = loadmol3 tripos3-mol3.mol2
MAPN = loadmol3 tripos4-mol3.mol2

CPN  = loadmol3 tripos5-mol3.mol2
NCP = loadmol3 tripos6-mol3.mol2
CCP = loadmol3 tripos7-mol3.mol2
MCPN = loadmol3 tripos8-mol3.mol2

GPN  = loadmol3 tripos9-mol3.mol2
NGP = loadmol3 tripos10-mol3.mol2
CGP = loadmol3 tripos11-mol3.mol2
MGPN = loadmol3 tripos12-mol3.mol2

IPN  = loadmol3 tripos13-mol3.mol2
NIP = loadmol3 tripos14-mol3.mol2
CIP = loadmol3 tripos15-mol3.mol2
MIPN = loadmol3 tripos16-mol3.mol2

TPN  = loadmol3 tripos17-mol3.mol2
NTP = loadmol3 tripos18-mol3.mol2
CTP = loadmol3 tripos19-mol3.mol2
MTP = loadmol3 tripos20-mol3.mol2

UPN  = loadmol3 tripos21-mol3.mol2
NUP = loadmol3 tripos22-mol3.mol2
CUP = loadmol3 tripos23-mol3.mol2
MUPN = loadmol3 tripos24-mol3.mol2

DIM = loadmol3 dimeric_charged.mol2

charge APN
charge CPN
charge GPN
charge IPN
charge TPN
charge UPN

charge NAP
charge NCP
charge NGP
charge NIP
charge NTP
charge NUP

charge CAP
charge CCP
charge CGP
charge CIP
charge CTP
charge CUP

charge MAPN
charge MCPN
```

```
charge MGPN
charge MIPN
charge MTPN
charge MUPN

charge DIM

DUPLEX = loadpdb duplex.pdb
check DUPLEX
solvatebox DUPLEX TIP3PBOX 12.0
saveAmberParm DUPLEX DUPLEX.top DUPLEX.crd
savepdb DUPLEX DUPLEX.pdb
quit
```

8.3.11. script2.sh

```
logfile 0PNA-from-mol3-files.log

verbosity 2

addAtomTypes {
      { "H"   "H" "sp3" }
      { "HO"  "H" "sp3" }
      { "HS"  "H" "sp3" }
      { "H1"  "H" "sp3" }
      { "H2"  "H" "sp3" }
      { "H3"  "H" "sp3" }
      { "H4"  "H" "sp3" }
      { "H5"  "H" "sp3" }
      { "HW"  "H" "sp3" }
      { "HC"  "H" "sp3" }
      { "HA"  "H" "sp3" }
      { "HP"  "H" "sp3" }
      { "OH"  "O" "sp3" }
      { "OS"  "O" "sp3" }
      { "O"   "O" "sp2" }
      { "O2"  "O" "sp2" }
      { "OW"  "O" "sp3" }
      { "CT"  "C" "sp3" }
      { "CH"  "C" "sp3" }
      { "C2"  "C" "sp3" }
      { "C3"  "C" "sp3" }
      { "C"   "C" "sp2" }
      { "C*"  "C" "sp2" }
      { "CA"  "C" "sp2" }
      { "CB"  "C" "sp2" }
      { "CC"  "C" "sp2" }
      { "CN"  "C" "sp2" }
      { "CM"  "C" "sp2" }
      { "CK"  "C" "sp2" }
      { "CQ"  "C" "sp2" }
      { "CD"  "C" "sp2" }
      { "CE"  "C" "sp2" }
      { "CF"  "C" "sp2" }
      { "CP"  "C" "sp2" }
      { "CI"  "C" "sp2" }
      { "CJ"  "C" "sp2" }
```

```
      { "CW"  "C" "sp2" }
      { "CV"  "C" "sp2" }
      { "CR"  "C" "sp2" }
      { "CA"  "C" "sp2" }
      { "CY"  "C" "sp2" }
      { "C0"  "C" "sp2" }
      { "MG"  "Mg" "sp3" }
      { "N"   "N" "sp2" }
      { "NA"  "N" "sp2" }
      { "N2"  "N" "sp2" }
      { "N*"  "N" "sp2" }
      { "NP"  "N" "sp2" }
      { "NQ"  "N" "sp2" }
      { "NB"  "N" "sp2" }
      { "NC"  "N" "sp2" }
      { "NT"  "N" "sp3" }
      { "N3"  "N" "sp3" }
      { "S"   "S" "sp3" }
      { "SH"  "S" "sp3" }
      { "P"   "P" "sp3" }
      { "LP"  ""  "sp3" }
      { "F"   "F" "sp3" }
      { "CL"  "Cl" "sp3" }
      { "BR"  "Br" "sp3" }
      { "I"   "I"  "sp3" }
      { "FE"  "Fe" "sp3" }
      { "EP"  ""  "sp3" }
}

addPdbResMap {
   { 0 "APN" "NAP" } { 1 "APN" "CAP" }
   { 0 "CPN" "NCP" } { 1 "CPN" "CCP" }
   { 0 "GPN" "NGP" } { 1 "GPN" "CGP" }
   { 0 "IPN" "NIP" } { 1 "IPN" "CIP" }
   { 0 "TPN" "NTP" } { 1 "TPN" "CTP" }
   { 0 "UPN" "NUP" } { 1 "UPN" "CUP" }
}

#parm99 = loadamberparams parm99.dat
#frcmod99SB = loadamberparams frcmod.ff99SB
#FRCMOD = loadAmberParams frcmod.pna
FRCMOD = loadAmberParams dimeric.frcmod
source leaprc.ff99bsc0
source leaprc.gaff
source leaprc.pna
#frcmodbsc0 = loadamberparams frcmod.parmbsc0

loadOff ions94.lib
loadOff solvents.lib
loadOff RNA_CI.lib
HOH = TP3
WAT = TP3

# See http://q4md-forcefieldtools.org/Tutorial/leap-mol3.php
APN  = loadmol3 tripos1-mol3.mol2
NAP = loadmol3 tripos2-mol3.mol2
CAP = loadmol3 tripos3-mol3.mol2
MAPN = loadmol3 tripos4-mol3.mol2

CPN  = loadmol3 tripos5-mol3.mol2
NCP = loadmol3 tripos6-mol3.mol2
CCP = loadmol3 tripos7-mol3.mol2
```

```
MCPN = loadmol3 tripos8-mol3.mol2

GPN  = loadmol3 tripos9-mol3.mol2
NGP = loadmol3 tripos10-mol3.mol2
CGP = loadmol3 tripos11-mol3.mol2
MGPN = loadmol3 tripos12-mol3.mol2

IPN  = loadmol3 tripos13-mol3.mol2
NIP = loadmol3 tripos14-mol3.mol2
CIP = loadmol3 tripos15-mol3.mol2
MIPN = loadmol3 tripos16-mol3.mol2

TPN  = loadmol3 tripos17-mol3.mol2
NTP = loadmol3 tripos18-mol3.mol2
CTP = loadmol3 tripos19-mol3.mol2
MTP = loadmol3 tripos20-mol3.mol2

UPN  = loadmol3 tripos21-mol3.mol2
NUP = loadmol3 tripos22-mol3.mol2
CUP = loadmol3 tripos23-mol3.mol2
MUPN = loadmol3 tripos24-mol3.mol2

DIM = loadmol3 dimeric_charged.mol2

charge APN
charge CPN
charge GPN
charge IPN
charge TPN
charge UPN

charge NAP
charge NCP
charge NGP
charge NIP
charge NTP
charge NUP

charge CAP
charge CCP
charge CGP
charge CIP
charge CTP
charge CUP

charge MAPN
charge MCPN
charge MGPN
charge MIPN
charge MTPN
charge MUPN

charge DIM

DUPLEX = loadpdb duplex.pdb
check DUPLEX
#solvatebox DUPLEX TIP3PBOX 12.0
saveAmberParm DUPLEX DUPLEX.top DUPLEX.crd
savepdb DUPLEX DUPLEX.pdb
quit
```

8.3.12. script3.sh

```
logfile 0PNA-from-mol3-files.log

verbosity 2

addAtomTypes {
      { "H"   "H" "sp3" }
      { "HO"  "H" "sp3" }
      { "HS"  "H" "sp3" }
      { "H1"  "H" "sp3" }
      { "H2"  "H" "sp3" }
      { "H3"  "H" "sp3" }
      { "H4"  "H" "sp3" }
      { "H5"  "H" "sp3" }
      { "HW"  "H" "sp3" }
      { "HC"  "H" "sp3" }
      { "HA"  "H" "sp3" }
      { "HP"  "H" "sp3" }
      { "OH"  "O" "sp3" }
      { "OS"  "O" "sp3" }
      { "O"   "O" "sp2" }
      { "O2"  "O" "sp2" }
      { "OW"  "O" "sp3" }
      { "CT"  "C" "sp3" }
      { "CH"  "C" "sp3" }
      { "C2"  "C" "sp3" }
      { "C3"  "C" "sp3" }
      { "C"   "C" "sp2" }
      { "C*"  "C" "sp2" }
      { "CA"  "C" "sp2" }
      { "CB"  "C" "sp2" }
      { "CC"  "C" "sp2" }
      { "CN"  "C" "sp2" }
      { "CM"  "C" "sp2" }
      { "CK"  "C" "sp2" }
      { "CQ"  "C" "sp2" }
      { "CD"  "C" "sp2" }
      { "CE"  "C" "sp2" }
      { "CF"  "C" "sp2" }
      { "CP"  "C" "sp2" }
      { "CI"  "C" "sp2" }
      { "CJ"  "C" "sp2" }
      { "CW"  "C" "sp2" }
      { "CV"  "C" "sp2" }
      { "CR"  "C" "sp2" }
      { "CA"  "C" "sp2" }
      { "CY"  "C" "sp2" }
      { "C0"  "C" "sp2" }
      { "MG"  "Mg" "sp3" }
      { "N"   "N" "sp2" }
      { "NA"  "N" "sp2" }
      { "N2"  "N" "sp2" }
      { "N*"  "N" "sp2" }
      { "NP"  "N" "sp2" }
      { "NQ"  "N" "sp2" }
      { "NB"  "N" "sp2" }
      { "NC"  "N" "sp2" }
      { "NT"  "N" "sp3" }
      { "N3"  "N" "sp3" }
      { "S"   "S" "sp3" }
```

```
      { "SH"  "S" "sp3" }
      { "P"   "P" "sp3" }
      { "LP"  ""  "sp3" }
      { "F"   "F" "sp3" }
      { "CL"  "Cl" "sp3" }
      { "BR"  "Br" "sp3" }
      { "I"   "I"  "sp3" }
      { "FE"  "Fe" "sp3" }
      { "EP"  ""  "sp3" }
}

addPdbResMap {
   { 0 "APN" "NAP" } { 1 "APN" "CAP" }
   { 0 "CPN" "NCP" } { 1 "CPN" "CCP" }
   { 0 "GPN" "NGP" } { 1 "GPN" "CGP" }
   { 0 "IPN" "NIP" } { 1 "IPN" "CIP" }
   { 0 "TPN" "NTP" } { 1 "TPN" "CTP" }
   { 0 "UPN" "NUP" } { 1 "UPN" "CUP" }
}

#parm99 = loadamberparams parm99.dat
#frcmod99SB = loadamberparams frcmod.ff99SB
#FRCMOD = loadAmberParams frcmod.pna
FRCMOD = loadAmberParams dimeric.frcmod
source leaprc.ff99bsc0
source leaprc.gaff
source leaprc.pna
#frcmodbsc0 = loadamberparams frcmod.parmbsc0

loadOff ions94.lib
loadOff solvents.lib
loadOff RNA_CI.lib
HOH = TP3
WAT = TP3

# See http://q4md-forcefieldtools.org/Tutorial/leap-mol3.php
APN  = loadmol3 tripos1-mol3.mol2
NAP = loadmol3 tripos2-mol3.mol2
CAP = loadmol3 tripos3-mol3.mol2
MAPN = loadmol3 tripos4-mol3.mol2

CPN  = loadmol3 tripos5-mol3.mol2
NCP = loadmol3 tripos6-mol3.mol2
CCP = loadmol3 tripos7-mol3.mol2
MCPN = loadmol3 tripos8-mol3.mol2

GPN  = loadmol3 tripos9-mol3.mol2
NGP = loadmol3 tripos10-mol3.mol2
CGP = loadmol3 tripos11-mol3.mol2
MGPN = loadmol3 tripos12-mol3.mol2

IPN  = loadmol3 tripos13-mol3.mol2
NIP = loadmol3 tripos14-mol3.mol2
CIP = loadmol3 tripos15-mol3.mol2
MIPN = loadmol3 tripos16-mol3.mol2

TPN  = loadmol3 tripos17-mol3.mol2
NTP = loadmol3 tripos18-mol3.mol2
CTP = loadmol3 tripos19-mol3.mol2
MTP = loadmol3 tripos20-mol3.mol2

UPN  = loadmol3 tripos21-mol3.mol2
```

```
NUP = loadmol3 tripos22-mol3.mol2
CUP = loadmol3 tripos23-mol3.mol2
MUPN = loadmol3 tripos24-mol3.mol2

DIM = loadmol3 dimeric_charged.mol2

charge APN
charge CPN
charge GPN
charge IPN
charge TPN
charge UPN

charge NAP
charge NCP
charge NGP
charge NIP
charge NTP
charge NUP

charge CAP
charge CCP
charge CGP
charge CIP
charge CTP
charge CUP

charge MAPN
charge MCPN
charge MGPN
charge MIPN
charge MTPN
charge MUPN

charge DIM

DUPLEX = loadpdb DUPLEX.pdb
check DUPLEX
solvatebox DUPLEX TIP3PBOX 12.0
saveAmberParm DUPLEX DUPLEX.top DUPLEX.crd
savepdb DUPLEX DUPLEX.pdb
quit
```

8.3.13.generate.sh

```
#!/bin/bash

rm nuc.c a.out nab_tmp.pdb tprmtop tleap.out nab_tmp.pdb* leap.log pna.pdb tmp
duplex.pdb

nab nuc.nab
./a.out

nn=`cat nab_tmp.pdb.lpdb| wc -l`

echo $nn

kt=`echo "${nn} - 2" | bc -l`

echo $kt

head -${kt} nab_tmp.pdb.lpdb > pna.pdb

grep 'TER' -B 1000000 pna.pdb > tmp

mv tmp pna.pdb


rm nuc.c a.out nab_tmp.pdb tprmtop tleap.out nab_tmp.pdb* leap.log



nab nuc2.nab
./a.out

grep 'TER' -A 1000000 nuc.pdb > tmp

cat pna.pdb tmp > duplex.pdb

rm nuc.c a.out nab_tmp.pdb tprmtop tleap.out nab_tmp.pdb* leap.log pna.pdb tmp

tleap -f script.sh

sander -O -i min.in -o duplex.out -c DUPLEX.crd -p DUPLEX.top -r duplex.rst

ambpdb -aatm -p DUPLEX.top < duplex.rst > duplex.pdb

grep -v "WAT" duplex.pdb > tmp

mv tmp duplex.pdb

tleap -f script2.sh

rm leap.log 0PNA-from-mol3-files.log duplex.* mdinfo nuc.pdb
```

8.3.14. correction.sh

```
#!/bin/bash
echo "Which file is to minimize?"
read 'file'

sed "s/duplex.pdb/${file}/g" script.sh > script3.sh
tleap -f script3.sh
sander -O -i min.in -o duplex.out -c DUPLEX.crd -p DUPLEX.top -r duplex.rst
ambpdb -aatm -p DUPLEX.top < duplex.rst > duplex.pdb
grep -v "WAT" duplex.pdb > tmp
mv tmp duplex.pdb
tleap -f script2.sh
rm leap.log 0PNA-from-mol3-files.log duplex.* mdinfo nuc.pdb
```

8.4. HPLC purification gradients

Eluents: water +0.1% TFA (A) and acetonitrile +0.1% TFA (B).

8.4.1. Unmodified PNA purification gradient

Time (min)	% A	% B
0	100	0
5	100	0
20	75	25
21	0	100
26	0	100
27	100	0
36	100	0

8.4.2. PNA1 purification gradients

Cycle 1:

Time (min)	% A	% B
0	100	0
5	100	0
50	62.5	37.5
51	0	100
56	0	100
57	100	0
66	100	0

Cycle 2:

Time (min)	% A	% B
0	100	0
5	100	0
41	70	30
42	0	100
47	0	100
48	100	0
57	100	0

8.4.3. PNA2 purification gradients

Cycle 1:

Time (min)	% A	% B
0	100	0
5	100	0
30	70	30
31	0	100
36	0	100
37	100	0
46	100	0

Cycle 2:

Time (min)	% A	% B
0	100	0
5	100	0
41	70	30
42	0	100
47	0	100
48	100	0
57	100	0

8.4.4. PNA3 purification gradients

Cycle 1:

Time (min)	% A	% B
0	100	0
5	100	0
41	70	30
42	0	100
47	0	100
48	100	0
57	100	0

Cycle 2:

Time (min)	% A	% B
0	100	0
5	100	0
31	78.3	21.7
32	0	100
37	0	100
38	100	0
47	100	0

8.5. UPLC gradient

Eluents: water +0.2% HCOOH (A) and acetonitrile +0.2% HCOOH (B)

Gradient:

Time (min)	% A	% B
0.0	100	0
0.9	100	0
6.6	50	50
7.0	50	50
7.2	0	100
8.0	0	100
8.2	100	0
11.0	100	0

8.6. Modified bases

8.6.1. Model 1

DIM

```
@<TRIPOS>MOLECULE
DIM
   73    76     1     0     0
SMALL
resp


@<TRIPOS>ATOM
      1 C4"          4.2100    2.9360   -1.1200 C          1 DIM     0.8982200137
      2 C5"          4.8630    2.9980    0.1810 CM         1 DIM    -0.6087609863
      3 C6"          4.6990    4.1180    0.9390 CM         1 DIM     0.0174510137
      4 C2"          3.2880    5.2280   -0.6700 C          1 DIM     0.8899850137
      5 H3"          2.9960    4.0820   -2.3270 H          1 DIM     0.4506100137
      6 H6"          5.1870    4.1640    1.9070 H4         1 DIM     0.2148800137
      7 C7"          5.7170    1.9110    0.7640 c          1 DIM     0.7989920137
      8 N2L          5.8770    0.8110   -0.0170 n          1 DIM    -0.4741149863
      9 C0L          6.6750   -0.3230    0.4040 c3         1 DIM    -0.2942059863
     10 H0L1         7.5740   -0.4000   -0.2240 h1         1 DIM     0.2033580137
     11 H0L2         7.0180   -0.0970    1.4190 h1         1 DIM     0.2033580137
     12 C7L          5.9410   -1.6530    0.3770 ca         1 DIM     0.2966100137
```

13	C6L	4.5680	-1.7480	0.3090	ca	1 DIM	-0.5013839863
14	C8L	6.7130	-2.8450	0.4500	ca	1 DIM	-0.4221169863
15	C5L	3.9110	-3.0090	0.3140	ca	1 DIM	0.1981480137
16	H6L	3.9660	-0.8460	0.2470	ha	1 DIM	0.2768690137
17	C9L	6.1110	-4.0800	0.4640	ca	1 DIM	-0.3055369863
18	H8L	7.7970	-2.7710	0.4960	ha	1 DIM	0.2752420137
19	C4L	2.4970	-3.1200	0.2410	ca	1 DIM	-0.5118489863
20	C10L	4.6980	-4.2010	0.3970	ca	1 DIM	0.0629710137
21	H9L	6.7140	-4.9830	0.5200	ha	1 DIM	0.2741200137
22	C3L	1.8710	-4.3490	0.2480	ca	1 DIM	0.3130740137
23	H4L	1.9040	-2.2130	0.1750	ha	1 DIM	0.2839450137
24	C11L	4.0330	-5.4550	0.4020	ca	1 DIM	-0.2959469863
25	C12L	2.6620	-5.5270	0.3270	ca	1 DIM	-0.4382819863
26	H11L	4.6270	-6.3640	0.4590	ha	1 DIM	0.2723910137
27	H12L	2.1700	-6.4960	0.3240	ha	1 DIM	0.2794430137
28	C2L	0.3600	-4.4950	0.2080	c3	1 DIM	-0.2976539863
29	H2L1	0.1040	-5.4160	-0.3340	h1	1 DIM	0.1975590137
30	H2L2	-0.0440	-4.6020	1.2200	h1	1 DIM	0.1975590137
31	N1L	-0.3290	-3.3660	-0.3870	n	1 DIM	-0.4289249863
32	C7	-1.3410	-2.7200	0.2460	c	1 DIM	0.7467350137
33	C5	-1.9570	-1.5810	-0.5150	CM	1 DIM	-0.5279789863
34	C4	-1.5400	-1.1440	-1.8410	C	1 DIM	0.8602970137
35	C6	-2.9890	-0.9330	0.0870	CM	1 DIM	0.0158340137
36	H6	-3.3140	-1.2710	1.0650	H4	1 DIM	0.2182980137
37	C2	-3.3290	0.6130	-1.7360	C	1 DIM	0.9211980137
38	H3	-2.0240	0.2710	-3.2570	H	1 DIM	0.4410200137
39	C8'	-4.8160	0.7010	0.1910	CT	1 DIM	-0.8799629863
40	H8'1	-4.6550	0.7180	1.2720	H1	1 DIM	0.3446040137
41	H8'2	-4.9220	1.7270	-0.1660	H1	1 DIM	0.3446040137
42	C7'	-6.0700	-0.1210	-0.1800	C	1 DIM	1.0245680137
43	N4'	-7.2500	0.2550	0.3950	N	1 DIM	-0.3187549863
44	C3'	-7.4190	1.3990	1.2990	CT	1 DIM	-0.3469869863
45	H3'1	-8.3710	1.8710	1.0460	H1	1 DIM	0.2204430137
46	H3'2	-6.6410	2.1430	1.1170	H1	1 DIM	0.2204430137
47	C5'	-8.4580	-0.5010	0.0360	CT	1 DIM	-0.4152869863
48	H5'1	-8.1450	-1.4790	-0.3220	H1	1 DIM	0.2243340137
49	H5'2	-9.0720	-0.6290	0.9320	H1	1 DIM	0.2243340137
50	C'	-9.3190	0.2850	-0.9660	C	1 DIM	0.7478980137
51	C2'	-7.4100	0.9880	2.7870	CT	1 DIM	0.1274780137
52	H2'1	-6.4470	0.5470	3.0550	H1	1 DIM	0.1389980137
53	H2'2	-8.1620	0.2080	2.9550	H1	1 DIM	0.1389980137
54	N1'	-7.6860	2.1070	3.6670	N	1 DIM	-0.7980199863
55	H1'1	-8.6380	2.4490	3.7150	H	1 DIM	0.4787090137
56	CC1	3.8110	6.3810	1.4140	CT	1 DIM	-0.9552239863
57	HC1	4.1850	7.2640	0.8930	H1	1 DIM	0.3107190137
58	HC2	2.7580	6.5430	1.6540	H1	1 DIM	0.3107190137
59	HC3	4.3800	6.2210	2.3290	H1	1 DIM	0.3107190137
60	O2"	2.6180	6.1790	-1.0270	O	1 DIM	-0.7900779863
61	O4"	4.2530	2.0090	-1.9280	O	1 DIM	-0.7882199863
62	N3"	3.4660	4.0830	-1.4290	NA	1 DIM	-0.6127129863
63	O7"	6.2220	2.0480	1.8800	o	1 DIM	-0.7705439863
64	H2L	5.4270	0.8200	-0.9290	hn	1 DIM	0.4077910137
65	H1L	-0.0850	-3.0450	-1.3210	hn	1 DIM	0.3892720137
66	O7	-1.7560	-3.0130	1.3700	o	1 DIM	-0.7597729863
67	O4	-0.6450	-1.6250	-2.5330	O	1 DIM	-0.7739679863
68	N3	-2.2730	-0.0520	-2.3290	NA	1 DIM	-0.5804379863
69	O2	-3.9250	1.5360	-2.2610	O	1 DIM	-0.7950389863
70	N1	-3.6370	0.1410	-0.4490	N*	1 DIM	0.1742120137
71	O7'	-5.9690	-1.0670	-0.9560	O	1 DIM	-0.7750459863
72	O1'	-9.9270	1.2860	-0.5870	O	1 DIM	-0.7948819863
73	N1"	3.9560	5.1960	0.5630	N*	1 DIM	0.3146820137

```
@<TRIPOS>BOND
     1    1    2 1
     2    1   61 1
     3    1   62 1
     4    2    3 1
     5    2    7 1
     6    3    6 1
     7    3   73 1
     8    4   60 1
     9    4   62 1
    10    4   73 1
    11    5   62 1
    12    7    8 1
    13    7   63 1
    14    8    9 1
    15    8   64 1
    16    9   10 1
    17    9   11 1
    18    9   12 1
    19   12   13 1
    20   12   14 1
    21   13   15 1
    22   13   16 1
    23   14   17 1
    24   14   18 1
    25   15   19 1
    26   15   20 1
    27   17   20 1
    28   17   21 1
    29   19   22 1
    30   19   23 1
    31   20   24 1
    32   22   25 1
    33   22   28 1
    34   24   25 1
    35   24   26 1
    36   25   27 1
    37   28   29 1
    38   28   30 1
    39   28   31 1
    40   31   32 1
    41   31   65 1
    42   32   33 1
    43   32   66 1
    44   33   34 1
    45   33   35 1
    46   34   67 1
    47   34   68 1
    48   35   36 1
    49   35   70 1
    50   37   68 1
    51   37   69 1
    52   37   70 1
    53   38   68 1
    54   39   40 1
    55   39   41 1
    56   39   42 1
    57   39   70 1
    58   42   43 1
    59   42   71 1
    60   43   44 1
    61   43   47 1
```

```
    62   44   45 1
    63   44   46 1
    64   44   51 1
    65   47   48 1
    66   47   49 1
    67   47   50 1
    68   50   72 1
    69   51   52 1
    70   51   53 1
    71   51   54 1
    72   54   55 1
    73   56   57 1
    74   56   58 1
    75   56   59 1
    76   56   73 1
@<TRIPOS>SUBSTRUCTURE
     1 DIM        1 TEMP              0 ****  ****    0 ROOT
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

frcmod.pna

```
PNA frcmod file: frcmod.pna
MASS

BOND
CQ-NA        502.0    1.324        by analogy to CQ-NC, parm99.dat
CM-c         349.7    1.487        SOURCE1  from gaff.dat

ANGL
C -CT-N*    50.00     110.13       FF cst by analogy to CT-CT-N*; Eq.val. from
QM geo.opt.
CQ-NA-C     70.0      118.60       by analogy to CA-NC-CQ, parm99.dat
H5-CQ-NA    50.0      115.45       by analogy to H5-CQ-NC, parm99.dat
NC-CQ-NA    70.0      129.10       by analogy to NC-CQ-NC, parm99.dat
CQ-NA-H     50.0      118.00       by analogy to CA-NA-H, parm99.dat
CM-c -o     68.670    123.44       SOURCE3  from gaff.dat
c -CM-C     62.650    120.000    SOURCE3  from gaff.dat
c -CM-CM    64.640    120.140      SOURCE3  from gaff.dat
n -c -CM    68.470    115.140      SOURCE4  from gaff.dat

DIHEDRAL
X -CQ-NA-X    4    6.00        180.0         2.0      by analogy to X -CA-NA-X,
parm99.dat
CM-CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
CM-CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
C -CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
C -CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
n -c -CM-C    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
CM-CM-c -n    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013

NONB
```

8.6.2. Model1 with cutter

NBC

```
@<TRIPOS>MOLECULE
NBC
  106  109    1    0    0
SMALL
NO_CHARGES
****
Energy = 0

@<TRIPOS>ATOM
      1   C4"       4.1950    3.3380   -0.9090 C         1 NBC     0.896168
      2   C5"       4.8370    3.3700    0.3990 CM        1 NBC    -0.610813
      3   C6"       4.6230    4.4490    1.2030 CM        1 NBC     0.015399
      4   C2"       3.1780    5.5690   -0.3690 C         1 NBC     0.887933
      5   H3"       2.9430    4.4850   -2.0750 H         1 NBC     0.448558
      6   H6"       5.1020    4.4730    2.1760 H4        1 NBC     0.212828
      7   C7"       5.7310    2.2950    0.9420 c         1 NBC     0.796940
      8   N2L       5.9400    1.2370    0.1160 n         1 NBC    -0.476167
      9   C0L       6.7810    0.1190    0.4950 c3        1 NBC    -0.296258
     10   H0L1      7.6870    0.1050   -0.1280 h1        1 NBC     0.201306
     11   H0L2      7.1080    0.3150    1.5210 h1        1 NBC     0.201306
     12   C7L       6.1020   -1.2370    0.4050 ca        1 NBC     0.294558
     13   C6L       4.7350   -1.3850    0.3220 ca        1 NBC    -0.503436
     14   C8L       6.9220   -2.3990    0.4340 ca        1 NBC    -0.424169
     15   C5L       4.1300   -2.6700    0.2680 ca        1 NBC     0.196096
     16   H6L       4.0970   -0.5060    0.2940 ha        1 NBC     0.274817
     17   C9L       6.3700   -3.6570    0.3900 ca        1 NBC    -0.307589
     18   H8L       8.0010   -2.2830    0.4920 ha        1 NBC     0.273190
     19   C4L       2.7220   -2.8350    0.1790 ca        1 NBC    -0.513901
     20   C10L      4.9640   -3.8320    0.3070 ca        1 NBC     0.060919
     21   H9L       7.0090   -4.5360    0.4130 ha        1 NBC     0.272068
     22   C3L       2.1460   -4.0880    0.1290 ca        1 NBC     0.311022
     23   H4L       2.0930   -1.9510    0.1470 ha        1 NBC     0.281893
     24   C11L      4.3510   -5.1110    0.2530 ca        1 NBC    -0.297999
     25   C12L      2.9840   -5.2350    0.1640 ca        1 NBC    -0.440334
     26   H11L      4.9810   -5.9970    0.2760 ha        1 NBC     0.270339
     27   H12L      2.5320   -6.2220    0.1160 ha        1 NBC     0.277391
     28   C2L       0.6430   -4.2930    0.0710 c3        1 NBC    -0.299706
     29   H2L1      0.4280   -5.1990   -0.5120 h1        1 NBC     0.195507
     30   H2L2      0.2370   -4.4590    1.0740 h1        1 NBC     0.195507
     31   N1L      -0.0880   -3.1680   -0.4810 n         1 NBC    -0.430977
     32   C7       -1.1290   -2.5910    0.1700 c         1 NBC     0.744683
     33   C5       -1.7860   -1.4470   -0.5460 CM        1 NBC    -0.530031
     34   C4       -1.3790   -0.9370   -1.8490 C         1 NBC     0.858245
     35   C6       -2.8480   -0.8670    0.0740 CM        1 NBC     0.013782
     36   H6       -3.1650   -1.2600    1.0350 H4        1 NBC     0.216246
     37   C2       -3.2390    0.7400   -1.6840 C         1 NBC     0.919146
     38   H3       -1.9110    0.5170   -3.2070 H         1 NBC     0.438968
     39   C8'      -4.7400    0.6850    0.2330 CT        1 NBC    -0.882015
     40   H8'1     -4.5870    0.6630    1.3150 H1        1 NBC     0.342552
     41   H8'2     -4.8860    1.7210   -0.0810 H1        1 NBC     0.342552
     42   C7'      -5.9580   -0.1700   -0.1820 C         1 NBC     1.022516
     43   N4'      -7.1550    0.1330    0.3980 N         1 NBC    -0.320807
     44   C3'      -7.3770    1.2290    1.3490 CT        1 NBC    -0.349039
     45   H3'1     -8.3450    1.6730    1.1090 H1        1 NBC     0.218391
```

```
     46   H3'2      -6.6280    2.0110    1.2050 H1        1 NBC     0.218391
     47   C5'       -8.3290   -0.6550   -0.0020 CT        1 NBC    -0.417339
     48   H5'1      -7.9750   -1.6030   -0.3990 H1        1 NBC     0.222282
     49   H5'2      -8.9430   -0.8460    0.8830 H1        1 NBC     0.222282
     50   C'        -9.2150    0.1380   -0.9770 C         1 NBC     0.745846
     51   C2'       -7.3600    0.7560    2.8180 CT        1 NBC     0.125426
     52   H2'1      -6.3820    0.3430    3.0750 H1        1 NBC     0.136946
     53   H2'2      -8.0810   -0.0610    2.9470 H1        1 NBC     0.136946
     54   N1'       -7.6870    1.8240    3.7430 N         1 NBC    -0.800072
     55   H1'1      -8.6530    2.1250    3.7980 H         1 NBC     0.476657
     56   O2"        2.4720    6.5070   -0.6900 O         1 NBC    -0.792130
     57   O4"        4.2810    2.4490   -1.7550 O         1 NBC    -0.790272
     58   N3"        3.4070    4.4660   -1.1740 NA        1 NBC    -0.614765
     59   O7"        6.2230    2.4040    2.0670 o         1 NBC    -0.772596
     60   H2L        5.4960    1.2670   -0.7980 hn        1 NBC     0.405739
     61   H1L        0.1490   -2.7980   -1.3990 hn        1 NBC     0.387220
     62   O7        -1.5390   -2.9490    1.2780 o         1 NBC    -0.761825
     63   O4        -0.4610   -1.3510   -2.5540 O         1 NBC    -0.776020
     64   N3        -2.1530    0.1450   -2.2960 NA        1 NBC    -0.582490
     65   O2        -3.8690    1.6600   -2.1740 O         1 NBC    -0.797091
     66   N1        -3.5360    0.2020   -0.4200 N*        1 NBC     0.172160
     67   O7'       -5.8130   -1.0770   -0.9970 O         1 NBC    -0.777098
     68   O1'       -9.8660    1.0960   -0.5600 O         1 NBC    -0.796934
     69   N1"        3.8390    5.5110    0.8670 N*        1 NBC     0.312630
     70  C1A         3.6823    6.6196    1.8196 c3        1 NBC     0.020637
     71  H1A1        4.1015    6.3386    2.7631 h1        1 NBC     0.110634
     72  C2A         4.4039    7.8707    1.2934 c3        1 NBC     0.020637
     73  H2A1        4.3974    8.6580    2.0495 h1        1 NBC     0.110634
     74  H2A2        3.8900    8.2542    0.4007 h1        1 NBC     0.110634
     75  N3A         5.7982    7.6056    0.9866 n         1 NBC    -0.721821
     76  H3A1        5.8806    7.3169    0.0327 hn        1 NBC     0.478007
     77  H1A2        2.6418    6.8361    1.9439 h1        1 NBC     0.110634
     78  CCC         6.5876    8.8274    1.1987 c         1 NBC     0.920517
     79  C2C         8.0609    8.7262    1.6474 c3        1 NBC    -0.985691
     80  OCC         6.0621    9.9559    1.0144 o         1 NBC    -0.906523
     81  N0C         8.8047    9.9709    1.3552 n3        1 NBC     0.485873
     82  H2C1        8.0487    8.5664    2.7312 h1        1 NBC     0.348005
     83  H2C2        8.5423    7.8482    1.2003 h1        1 NBC     0.348005
     84  C1C         9.4481    9.9692    0.0048 c3        1 NBC    -0.690862
     85  C3C         9.6909   10.4511    2.4600 c3        1 NBC    -0.690862
     86  C9C         9.7751   11.3888   -0.5012 c3        1 NBC    -0.271161
     87  H1C1       10.3666    9.3684   -0.0006 h1        1 NBC     0.295748
     88  H1C2        8.7444    9.4912   -0.6855 h1        1 NBC     0.295748
     89  C4C         9.0104   11.5214    3.3376 c3        1 NBC    -0.271161
     90  H3C1       10.6040   10.8506    2.0155 h1        1 NBC     0.295748
     91  H3C2       10.0096    9.6228    3.1008 h1        1 NBC     0.295748
     92  N8C         8.6226   12.3497   -0.3482 n3        1 NBC    -0.626037
     93  H9C1       10.6381   11.7995    0.0271 h1        1 NBC     0.238064
     94  H9C2       10.0630   11.3208   -1.5537 h1        1 NBC     0.238064
     95  N5C         8.3544   12.6035    2.5224 n3        1 NBC    -0.626037
     96  H4C1        9.7540   11.9481    4.0198 h1        1 NBC     0.238064
     97  H4C2        8.2241   11.0722    3.9533 h1        1 NBC     0.238064
     98  C7C         8.9790   13.6578    0.2972 c3        1 NBC    -0.360158
     99  H8C1        8.2380   12.5529   -1.2697 hn        1 NBC     0.468702
    100  C6C         9.3200   13.4992    1.7910 c3        1 NBC    -0.360158
    101  H5C1        7.7964   13.1702    3.1595 hn        1 NBC     0.468702
    102  H7C1        8.1136   14.3166    0.1719 h1        1 NBC     0.250612
    103  H7C2        9.8262   14.1296   -0.2125 h1        1 NBC     0.250612
    104  H6C1        9.3279   14.4910    2.2503 h1        1 NBC     0.250612
    105  H6C2       10.3270   13.0951    1.9154 h1        1 NBC     0.250612
    106  Zn          7.3006   11.4904    1.0595 Zn        1 NBC     1.392676
@<TRIPOS>BOND
```

1	1	57	2
2	1	58	1
3	1	2	1
4	2	3	2
5	2	7	1
6	3	6	1
7	3	69	1
8	4	56	2
9	4	58	1
10	4	69	1
11	5	58	1
12	7	8	1
13	7	59	2
14	8	60	1
15	8	9	1
16	9	10	1
17	9	11	1
18	9	12	1
19	12	13	2
20	12	14	1
21	13	15	1
22	13	16	1
23	14	17	2
24	14	18	1
25	15	20	1
26	15	19	1
27	17	20	1
28	17	21	1
29	19	22	2
30	19	23	1
31	20	24	1
32	22	25	1
33	22	28	1
34	24	25	2
35	24	26	1
36	25	27	1
37	28	29	1
38	28	30	1
39	28	31	1
40	31	32	1
41	31	61	1
42	32	33	1
43	32	62	2
44	33	34	1
45	33	35	2
46	34	64	1
47	34	63	2
48	35	66	1
49	35	36	1
50	37	64	1
51	37	65	2
52	37	66	1
53	38	64	1
54	39	66	1
55	39	40	1
56	39	41	1
57	39	42	1
58	42	43	1
59	42	67	2
60	43	44	1
61	43	47	1
62	44	45	1

```
   63    44    51  1
   64    44    46  1
   65    47    48  1
   66    47    50  1
   67    47    49  1
   68    50    68  2
   69    51    53  1
   70    51    54  1
   71    51    52  1
   72    54    55  1
   73    69    70  1
   74    70    71  1
   75    70    72  1
   76    70    77  1
   77    72    73  1
   78    72    74  1
   79    72    75  1
   80    75    76  1
   81    75    78  1
   82    78    79  1
   83    78    80  2
   84    79    81  1
   85    79    82  1
   86    79    83  1
   87    81    84  1
   88    81    85  1
   89    84    86  1
   90    84    87  1
   91    84    88  1
   92    85    89  1
   93    85    90  1
   94    85    91  1
   95    86    92  1
   96    86    93  1
   97    86    94  1
   98    89    95  1
   99    89    96  1
  100    89    97  1
  101    92    98  1
  102    92    99  1
  103    95   100  1
  104    95   101  1
  105    98   100  1
  106    98   102  1
  107    98   103  1
  108   100   104  1
  109   100   105  1

@<TRIPOS>SUBSTRUCTURE
     1 NBC         1 TEMP              0 ****  ****    0 ROOT
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

frcmod.pna

```
PNA frcmod file: frcmod.pna
MASS

BOND
CQ-NA        502.0    1.324        by analogy to CQ-NC, parm99.dat
CM-c         349.7    1.487        SOURCE1  from gaff.dat
n3-Zn         50.4    2.089        ZAFF
c3-N*        337.0    1.475        JCC,7,(1986),230; ADE,CYT,GUA,THY,URA

ANGL
C -CT-N*    50.00     110.13       FF cst by analogy to CT-CT-N*; Eq.val. from
QM geo.opt.
CQ-NA-C     70.0      118.60       by analogy to CA-NC-CQ, parm99.dat
H5-CQ-NA    50.0      115.45       by analogy to H5-CQ-NC, parm99.dat
NC-CQ-NA    70.0      129.10       by analogy to NC-CQ-NC, parm99.dat
CQ-NA-H     50.0      118.00       by analogy to CA-NA-H, parm99.dat
CM-c -o     68.670    123.44       SOURCE3  from gaff.dat
c -CM-C     62.650    120.000    SOURCE3  from gaff.dat
c -CM-CM    64.640    120.140      SOURCE3  from gaff.dat
n -c -CM    68.470    115.140      SOURCE4  from gaff.dat
N*-c3-h1    50.0      109.50       parm99
N*-c3-c3    67.530    112.140    SOURCE4
C-N*-c3     70.0      117.60       parm99
CM-N*-c3    70.0      121.20       parm99

DIHEDRAL
X -CQ-NA-X    4    6.00        180.0         2.0      by analogy to X -CA-NA-X,
parm99.dat
CM-CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
CM-CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
C -CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
C -CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
n -c -CM-C    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
CM-CM-c -n    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
X-c3-N*-X    6    0.00          0.0             2.         JCC,7,(1986),230

NONB
```

8.6.3. Model2

L1D

```
@<TRIPOS>MOLECULE
L1D
   74    77     1     0     0
SMALL
resp


@<TRIPOS>ATOM
      1 C4"          5.5680    3.0780   -1.0530 C        1 L1D      0.875570
      2 C5"          5.7560    2.9470    0.3900 CM       1 L1D     -0.352481
      3 C6"          5.2680    3.9210    1.1940 CM       1 L1D     -0.134303
      4 C2"          4.3940    5.2530   -0.6280 C        1 L1D      0.931158
      5 H3"          4.7430    4.3730   -2.4310 H        1 L1D      0.463365
      6 H6"          5.3740    3.8670    2.2730 H4       1 L1D      0.284295
      7 C7"          6.4960    1.7410    0.9060 c3       1 L1D     -0.511818
      8 N2L          5.7850    0.4990    0.5960 n3       1 L1D     -0.512400
      9 C0L          6.5040   -0.6840    1.0560 c3       1 L1D     -0.328426
     10 H0L1         7.5260   -0.7590    0.6360 h1       1 L1D      0.212106
     11 H0L2         6.6410   -0.5870    2.1450 h1       1 L1D      0.212106
     12 C7L          5.7590   -1.9720    0.7660 ca       1 L1D      0.208201
     13 C6L          4.3890   -1.9980    0.6020 ca       1 L1D     -0.410828
     14 C8L          6.4860   -3.1910    0.6890 ca       1 L1D     -0.389291
     15 C5L          3.6960   -3.2140    0.3580 ca       1 L1D      0.203371
     16 H6L          3.8320   -1.0680    0.6580 ha       1 L1D      0.193653
     17 C9L          5.8460   -4.3860    0.4630 ca       1 L1D     -0.308743
     18 H8L          7.5670   -3.1690    0.8080 ha       1 L1D      0.263920
     19 C4L          2.2860   -3.2530    0.1840 ca       1 L1D     -0.532370
     20 C10L         4.4380   -4.4360    0.2900 ca       1 L1D      0.044448
     21 H9L          6.4150   -5.3110    0.4040 ha       1 L1D      0.272120
     22 C3L          1.6230   -4.4380   -0.0530 ca       1 L1D      0.302742
     23 H4L          1.7280   -2.3230    0.2360 ha       1 L1D      0.303871
     24 C11L         3.7350   -5.6450    0.0450 ca       1 L1D     -0.291310
     25 C12L         2.3710   -5.6450   -0.1260 ca       1 L1D     -0.429242
     26 H11L         4.2950   -6.5750   -0.0150 ha       1 L1D      0.268923
     27 H12L         1.8500   -6.5800   -0.3210 ha       1 L1D      0.272918
     28 C2L          0.1150   -4.5140   -0.2110 c3       1 L1D     -0.297705
     29 H2L1        -0.1300   -5.3010   -0.9380 h1       1 L1D      0.197035
     30 H2L2        -0.3590   -4.8010    0.7330 h1       1 L1D      0.197035
     31 N1L         -0.5020   -3.2640   -0.6130 n        1 L1D     -0.440216
     32 C7          -1.5380   -2.7150    0.0690 c        1 L1D      0.758291
     33 C5          -2.0580   -1.4180   -0.4830 CM       1 L1D     -0.570754
     34 C4          -1.5170   -0.7330   -1.6510 C        1 L1D      0.887496
     35 C6          -3.1180   -0.8630    0.1610 CM       1 L1D      0.040649
     36 H6          -3.5350   -1.3800    1.0190 H4       1 L1D      0.216458
     37 C2          -3.2600    1.0410   -1.3200 C        1 L1D      0.921597
     38 H3          -1.8380    0.9650   -2.7710 H        1 L1D      0.439934
     39 C8'         -4.8990    0.8010    0.4620 CT       1 L1D     -0.847203
     40 H8'1        -4.8420    0.5680    1.5290 H1       1 L1D      0.322992
     41 H8'2        -4.9280    1.8860    0.3390 H1       1 L1D      0.322992
     42 C7'         -6.1440    0.1600   -0.1830 C        1 L1D      0.939704
     43 N4'         -7.3590    0.4730    0.3500 N        1 L1D     -0.062888
     44 C3'         -7.6090    1.5200    1.3450 CT       1 L1D     -0.437516
     45 H3'1        -8.4790    2.0930    1.0080 H1       1 L1D      0.233708
     46 H3'2        -6.7710    2.2200    1.3780 H1       1 L1D      0.233708
```

```
      47 C5'        -8.5430   -0.1400   -0.2750 CT       1 L1D     -0.870227
      48 H5'1       -8.2730   -1.1440   -0.6050 H1       1 L1D      0.324431
      49 H5'2       -9.3410   -0.1990    0.4650 H1       1 L1D      0.324431
      50 C'         -9.0550    0.7300   -1.4350 C        1 L1D      0.825499
      51 C2'        -7.8670    0.9520    2.7560 CT       1 L1D      0.168043
      52 H2'1       -6.9860    0.4110    3.1080 H1       1 L1D      0.125461
      53 H2'2       -8.6810    0.2190    2.7110 H1       1 L1D      0.125461
      54 N1'        -8.2130    1.9860    3.7120 N        1 L1D     -0.824348
      55 H1'1       -9.1470    2.3740    3.6760 H        1 L1D      0.481077
      56 O2"         3.8130    6.2360   -1.0560 O        1 L1D     -0.814395
      57 O4"         5.9400    2.2690   -1.8950 O        1 L1D     -0.773254
      58 N3"         4.8960    4.2470   -1.4370 NA       1 L1D     -0.683505
      59 H2L         5.6820    0.4510   -0.4170 hn       1 L1D      0.400919
      60 H1L        -0.1730   -2.7680   -1.4380 hn       1 L1D      0.391910
      61 O7         -2.0480   -3.2100    1.0760 o        1 L1D     -0.765352
      62 O4         -0.5780   -1.0980   -2.3550 O        1 L1D     -0.784547
      63 N3         -2.1760    0.4680   -1.9550 NA       1 L1D     -0.595513
      64 O2         -3.7880    2.0750   -1.6860 O        1 L1D     -0.791260
      65 N1         -3.6880    0.3280   -0.1880 N*       1 L1D      0.178296
      66 O7'        -6.0150   -0.6140   -1.1310 O        1 L1D     -0.771219
      67 O1'        -9.9490    1.5530   -1.2560 O        1 L1D     -0.748132
      68 N1"         4.6110    5.0420    0.7340 N*       1 L1D      0.257124
      69 CC1         4.1000    6.0710    1.6410 CT       1 L1D     -0.899509
      70 HC1         4.5560    7.0370    1.4120 H1       1 L1D      0.293632
      71 HC2         3.0180    6.1700    1.5310 H1       1 L1D      0.293632
      72 HC3         4.3410    5.7850    2.6650 H1       1 L1D      0.293632
      73 H7"1        7.5220    1.7510    0.4890 h1       1 L1D      0.268316
      74 H7"2        6.6010    1.8190    1.9950 h1       1 L1D      0.268316
@<TRIPOS>BOND
     1    1    2 1
     2    1   57 1
     3    1   58 1
     4    2    3 1
     5    2    7 1
     6    3    6 1
     7    3   68 1
     8    4   56 1
     9    4   58 1
    10    4   68 1
    11    5   58 1
    12    7    8 1
    13    7   73 1
    14    7   74 1
    15    8    9 1
    16    8   59 1
    17    9   10 1
    18    9   11 1
    19    9   12 1
    20   12   13 1
    21   12   14 1
    22   13   15 1
    23   13   16 1
    24   14   17 1
    25   14   18 1
    26   15   19 1
    27   15   20 1
    28   17   20 1
    29   17   21 1
    30   19   22 1
    31   19   23 1
    32   20   24 1
    33   22   25 1
```

```
    34   22   28 1
    35   24   25 1
    36   24   26 1
    37   25   27 1
    38   28   29 1
    39   28   30 1
    40   28   31 1
    41   31   32 1
    42   31   60 1
    43   32   33 1
    44   32   61 1
    45   33   34 1
    46   33   35 1
    47   34   62 1
    48   34   63 1
    49   35   36 1
    50   35   65 1
    51   37   63 1
    52   37   64 1
    53   37   65 1
    54   38   63 1
    55   39   40 1
    56   39   41 1
    57   39   42 1
    58   39   65 1
    59   42   43 1
    60   42   66 1
    61   43   44 1
    62   43   47 1
    63   44   45 1
    64   44   46 1
    65   44   51 1
    66   47   48 1
    67   47   49 1
    68   47   50 1
    69   50   67 1
    70   51   52 1
    71   51   53 1
    72   51   54 1
    73   54   55 1
    74   68   69 1
    75   69   70 1
    76   69   71 1
    77   69   72 1
@<TRIPOS>SUBSTRUCTURE
     1 L1D         1 TEMP              0 ****  ****    0 ROOT
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

frcmod.pna

```
PNA frcmod file: frcmod.pna
MASS

BOND
CQ-NA        502.0    1.324        by analogy to CQ-NC, parm99.dat
CM-c         349.7    1.487        SOURCE1  from gaff.dat
```

```
n3-Zn          50.4    2.089        ZAFF
c3-N*         337.0    1.475        JCC,7,(1986),230; ADE,CYT,GUA,THY,URA
CM-c3         317.0    1.510        by analogy parm99

ANGL
C -CT-N*     50.00     110.13       FF cst by analogy to CT-CT-N*; Eq.val. from
QM geo.opt.
CQ-NA-C      70.0      118.60       by analogy to CA-NC-CQ, parm99.dat
H5-CQ-NA     50.0      115.45       by analogy to H5-CQ-NC, parm99.dat
NC-CQ-NA     70.0      129.10       by analogy to NC-CQ-NC, parm99.dat
CQ-NA-H      50.0      118.00       by analogy to CA-NA-H, parm99.dat
CM-c -o      68.670    123.44       SOURCE3  from gaff.dat
c -CM-C      62.650    120.000    SOURCE3  from gaff.dat
c -CM-CM     64.640    120.140      SOURCE3  from gaff.dat
n -c -CM     68.470    115.140      SOURCE4  from gaff.dat
N*-c3-h1     50.0      109.50       parm99
N*-c3-c3     67.530    112.140    SOURCE4
C-N*-c3      70.0      117.60       parm99
CM-N*-c3     70.0      121.20       parm99
CM-CM-c3     70.0      119.70       parm99
CM-c3-n3     66.18     112.13     SOURCE4          387 1.2309
CM-c3-h1     46.78     110.95     SOURCE3           12 1.1170
C -CM-c3     70.0      119.70       by parm99

DIHEDRAL
X -CQ-NA-X    4    6.00        180.0         2.0      by analogy to X -CA-NA-X,
parm99.dat
CM-CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
CM-CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
C -CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
C -CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
n -c -CM-C    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
CM-CM-c -n    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
X -c3-N*-X    6    0.00          0.0         2.       JCC,7,(1986),230
X -c3-CM-X    6    0.000         0.000       2.       JCC,7,(1986),230

NONB
```

8.6.4. Model3

DCL

```
@<TRIPOS>MOLECULE
DCL
   71    76     1     0     0
SMALL
resp


@<TRIPOS>ATOM
      1 C5          4.1580    0.4530   -1.0520  CM        1 DCL     -0.678993
      2 N1          6.4860   -0.1270   -1.1960  N*        1 DCL      0.353895
      3 C2          6.2210   -1.4180   -1.6720  C         1 DCL      0.884243
      4 C6          5.4740    0.7650   -0.9340  CM        1 DCL     -0.045393
      5 N3          4.8670   -1.7160   -1.7360  NA        1 DCL     -0.692795
      6 O2          7.0980   -2.1970   -1.9860  O         1 DCL     -0.767500
      7 C4          3.7640   -0.8910   -1.4790  C         1 DCL      0.995462
```

8	O4	2.6250	-1.3120	-1.6130	O	1 DCL	-0.790335
9	H3	4.6480	-2.6530	-2.0580	H	1 DCL	0.458389
10	C5"	-8.0960	-1.3320	0.0790	CM	1 DCL	-0.638314
11	N1"	-10.1930	-2.5010	0.0790	N*	1 DCL	0.234916
12	C2"	-9.6020	-3.7370	-0.1940	C	1 DCL	0.896040
13	C6"	-9.4500	-1.3570	0.2050	CM	1 DCL	0.035812
14	N3"	-8.2210	-3.6750	-0.3190	NA	1 DCL	-0.727057
15	O2"	-10.2440	-4.7640	-0.3110	O	1 DCL	-0.780540
16	C4"	-7.3700	-2.5670	-0.2090	C	1 DCL	1.009499
17	O4"	-6.1620	-2.6970	-0.3520	O	1 DCL	-0.796925
18	H3"	-7.7660	-4.5580	-0.5190	H	1 DCL	0.468883
19	C12L	-0.1570	4.4830	0.1420	ca	1 DCL	-0.172488
20	C3L	0.0310	3.0950	-0.1210	ca	1 DCL	-0.092917
21	C4L	-1.0700	2.2570	-0.1290	ca	1 DCL	-0.422894
22	C11L	-1.4120	4.9860	0.3870	ca	1 DCL	-0.440297
23	C10L	-2.5550	4.1440	0.3850	ca	1 DCL	0.191989
24	C5L	-2.3770	2.7480	0.1200	ca	1 DCL	0.193024
25	C6L	-3.5130	1.8990	0.1140	ca	1 DCL	-0.428723
26	C7L	-4.7840	2.3890	0.3580	ca	1 DCL	-0.092585
27	C8L	-4.9500	3.7800	0.6230	ca	1 DCL	-0.172110
28	C9L	-3.8680	4.6250	0.6340	ca	1 DCL	-0.438412
29	CC1	-11.6530	-2.4920	0.2200	CT	1 DCL	-0.966796
30	C8'	7.8960	0.2800	-1.1490	CT	1 DCL	-0.962252
31	C13L	-5.9620	1.5150	0.3500	c2	1 DCL	0.396335
32	C2L	1.3850	2.5900	-0.3750	c2	1 DCL	0.398493
33	C0L	-6.0660	0.1570	0.1020	c2	1 DCL	-0.771309
34	NA1"	-7.3940	-0.1140	0.2280	na	1 DCL	1.002196
35	NA2"	-8.0840	1.0240	0.5400	n2	1 DCL	-0.257207
36	NA3"	-7.2220	1.9910	0.6100	n2	1 DCL	-0.429186
37	NA3	2.4790	3.4180	-0.3820	n2	1 DCL	-0.428986
38	NA2	3.5510	2.7350	-0.6330	n2	1 DCL	-0.261580
39	NA1	3.1730	1.4310	-0.7940	na	1 DCL	1.058193
40	C1L	1.8260	1.3060	-0.6400	c2	1 DCL	-0.797263
41	H6	5.7700	1.7620	-0.6310	H4	1 DCL	0.199919
42	H6"	-9.9940	-0.4460	0.4180	H4	1 DCL	0.179932
43	H12L	0.7120	5.1320	0.1460	ha	1 DCL	0.187421
44	H4L	-0.9500	1.1940	-0.3290	ha	1 DCL	0.280602
45	H11L	-1.5440	6.0470	0.5870	ha	1 DCL	0.292821
46	H6L	-3.3630	0.8410	-0.0880	ha	1 DCL	0.282571
47	H8L	-5.9500	4.1540	0.8140	ha	1 DCL	0.187303
48	H9L	-4.0070	5.6850	0.8360	ha	1 DCL	0.292054
49	HC1	-11.9750	-1.4720	0.4340	H1	1 DCL	0.316690
50	HC2	-11.9530	-3.1550	1.0340	H1	1 DCL	0.316690
51	HC3	-12.1200	-2.8430	-0.7030	H1	1 DCL	0.316690
52	H8'1	7.9230	1.3260	-0.8300	H1	1 DCL	0.344292
53	H8'2	8.3260	0.2120	-2.1500	H1	1 DCL	0.344292
54	H0L	-5.3590	-0.6180	-0.1400	h4	1 DCL	0.371286
55	H1L1	1.3360	0.3530	-0.7440	h4	1 DCL	0.381215
56	C7'	8.8460	-0.5630	-0.2720	C	1 DCL	0.832004
57	O7'	9.9100	-0.9170	-0.7690	O	1 DCL	-0.734164
58	N4'	8.5320	-0.8440	1.0310	N	1 DCL	0.044965
59	C3'	7.4410	-0.2250	1.8000	CT	1 DCL	-0.585281
60	H3'1	7.8600	0.1380	2.7450	H1	1 DCL	0.277921
61	H3'2	7.0870	0.6570	1.2610	H1	1 DCL	0.277921
62	C2'	6.2470	-1.1430	2.1120	CT	1 DCL	0.009452
63	H2'1	5.7970	-1.4990	1.1790	H1	1 DCL	0.150075
64	H2'2	5.5030	-0.5580	2.6600	H1	1 DCL	0.150075
65	N1'	6.5910	-2.2810	2.9360	N	1 DCL	-0.822702
66	C5'	9.5480	-1.6190	1.7530	CT	1 DCL	-0.652017
67	H5'1	9.3000	-1.5790	2.8160	H1	1 DCL	0.285724
68	H5'2	10.5270	-1.1580	1.6050	H1	1 DCL	0.285724
69	C'	9.5520	-3.1080	1.3610	C	1 DCL	0.751758

```
     70 O'          8.5320   -3.7950    1.4750  O         1 DCL     -0.797801
     71 H1'1        7.1310   -3.0250    2.4920  H         1 DCL      0.538828
@<TRIPOS>BOND
     1    1    4 1
     2    1    7 1
     3    1   39 1
     4    2    3 1
     5    2    4 1
     6    2   30 1
     7    3    5 1
     8    3    6 1
     9    4   41 1
    10    5    7 1
    11    5    9 1
    12    7    8 1
    13   10   13 1
    14   10   16 1
    15   10   34 1
    16   11   12 1
    17   11   13 1
    18   11   29 1
    19   12   14 1
    20   12   15 1
    21   13   42 1
    22   14   16 1
    23   14   18 1
    24   16   17 1
    25   19   20 1
    26   19   22 1
    27   19   43 1
    28   20   21 1
    29   20   32 1
    30   21   24 1
    31   21   44 1
    32   22   23 1
    33   22   45 1
    34   23   24 1
    35   23   28 1
    36   24   25 1
    37   25   26 1
    38   25   46 1
    39   26   27 1
    40   26   31 1
    41   27   28 1
    42   27   47 1
    43   28   48 1
    44   29   49 1
    45   29   50 1
    46   29   51 1
    47   30   52 1
    48   30   53 1
    49   30   56 1
    50   31   33 1
    51   31   36 1
    52   32   37 1
    53   32   40 1
    54   33   34 1
    55   33   54 1
    56   34   35 1
    57   35   36 1
    58   37   38 1
    59   38   39 1
```

```
    60   39   40 1
    61   40   55 1
    62   56   57 1
    63   56   58 1
    64   58   59 1
    65   58   66 1
    66   59   60 1
    67   59   61 1
    68   59   62 1
    69   62   63 1
    70   62   64 1
    71   62   65 1
    72   65   71 1
    73   66   67 1
    74   66   68 1
    75   66   69 1
    76   69   70 1
@<TRIPOS>SUBSTRUCTURE
     1 DCL         1 TEMP              0 ****  ****    0 ROOT
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0
```

frcmod.pna

```
PNA frcmod file: frcmod.pna
MASS

BOND
CQ-NA        502.0    1.324       by analogy to CQ-NC, parm99.dat
CM-c         349.7    1.487       SOURCE1  from gaff.dat
n3-Zn         50.4    2.089       ZAFF
c3-N*        337.0    1.475       JCC,7,(1986),230; ADE,CYT,GUA,THY,URA
CM-c3        317.0    1.510       by analogy parm99
c2-CM        357.2    1.4800      SOUECE3       1
CM-na        422.0    1.385       JCC,7,(1986),230; HIS

ANGL
C -CT-N*    50.00     110.13      FF cst by analogy to CT-CT-N*; Eq.val. from
QM geo.opt.
CQ-NA-C     70.0      118.60      by analogy to CA-NC-CQ, parm99.dat
H5-CQ-NA    50.0      115.45      by analogy to H5-CQ-NC, parm99.dat
NC-CQ-NA    70.0      129.10      by analogy to NC-CQ-NC, parm99.dat
CQ-NA-H     50.0      118.00      by analogy to CA-NA-H, parm99.dat
CM-c -o     68.670    123.44      SOURCE3  from gaff.dat
c -CM-C     62.650    120.000   SOURCE3  from gaff.dat
c -CM-CM    64.640    120.140     SOURCE3  from gaff.dat
n -c -CM    68.470    115.140     SOURCE4  from gaff.dat
N*-c3-h1    50.0      109.50      parm99
N*-c3-c3    67.530    112.140   SOURCE4
C-N*-c3     70.0      117.60      parm99
CM-N*-c3    70.0      121.20      parm99
CM-CM-c3    70.0      119.70      parm99
CM-c3-n3    66.18     112.13    SOURCE4          387 1.2309
CM-c3-h1    46.78     110.95    SOURCE3           12 1.1170
C -CM-c3    70.0      119.70      by parm99
c2-c2-CM    66.880    117.000   SOURCE3            1
```

```
c2-c -ca    67.170    116.780    SOURCE3            1
c2-CM-C     67.930    120.700    SOURCE3            1
c2-CM-CM    64.690    120.600    SOURCE3
ha-c2-CM    50.040    120.940    SOURCE3          254 1.3150
ca-c2-n2    71.29     126.01     SOURCE3            1 0.0000
C -CM-na    69.83     121.38       SOURCE3            26    6.9463
CM-CM-na    70.0      121.20
CM-na-c2    67.80     110.37     SOURCE3            6 0.5121
CM-na-n2    65.80     125.00     SOURCE3            1

DIHEDRAL
X -CQ-NA-X    4    6.00        180.0         2.0      by analogy to X -CA-NA-X,
parm99.dat
CM-CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
CM-CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
C -CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
C -CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
n -c -CM-C    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
CM-CM-c -n    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
X -c3-CM-X    6    0.000         0.000       2.       JCC,7,(1986),230
X -c2-CM-X    4   26.600       180.000       2.000    c2=c2 double bond,
intrpol.bsd.on C6H6
X -CM-na-X    4    2.500       180.000       2.000

NONB
```

8.6.5. Model4

ABI

```
@<TRIPOS>MOLECULE
ABI
   71    74     1     0     0
SMALL
resp


@<TRIPOS>ATOM
      1 HC1          9.8980   -5.4250   -0.2580 H1        1 ABI        0.299385
      2 HC2         10.3610   -4.1810    0.9380 H1        1 ABI        0.299385
      3 HC3          9.4220   -5.6160    1.4390 H1        1 ABI        0.299385
      4 H7"          8.1880   -0.1910    0.6860 ha        1 ABI        0.212467
      5 H8'1        -6.9310    0.2930   -1.8650 H1        1 ABI        0.377267
      6 H8'2        -6.9290   -1.4670   -1.6590 H1        1 ABI        0.377267
      7 H7          -3.1490    2.8670   -0.9480 ha        1 ABI        0.217686
      8 H1L         -0.4940    1.4740   -0.3900 ha        1 ABI        0.252844
      9 H0L          5.2460   -0.0360   -0.0950 ha        1 ABI        0.261679
     10 H11L         2.6820    6.7870   -0.1360 ha        1 ABI        0.282344
     11 H12L         0.3160    6.1050   -0.4820 ha        1 ABI        0.237686
     12 H4L          1.4600    1.9660   -0.3280 ha        1 ABI        0.299116
     13 H9L          5.0740    6.1440    0.2150 ha        1 ABI        0.282363
     14 H6L          3.7720    1.3450   -0.0020 ha        1 ABI        0.294156
     15 H8L          6.7930    4.3640    0.4610 ha        1 ABI        0.237926
     16 H6          -5.0880    1.5410   -1.1290 H4        1 ABI        0.284436
     17 H6"          9.2090   -2.3000    0.7740 H4        1 ABI        0.282277
     18 CC1          9.5850   -4.8880    0.6410 CT        1 ABI       -0.905467
```

```
     19 C7"         7.2200   -0.5970    0.3910 c2        1 ABI       -0.097566
     20 C8'        -6.6660   -0.5250   -1.1870 CT        1 ABI       -1.024424
     21 O13L        7.6450    2.1350    0.6840 o         1 ABI       -0.681160
     22 C7         -2.4990    2.0110   -0.7680 c2        1 ABI       -0.121851
     23 C1L        -1.1900    2.2780   -0.5820 c2        1 ABI       -0.581275
     24 O2L        -1.5200    4.6090   -0.8350 o         1 ABI       -0.679453
     25 C0L         6.2280    0.2990    0.2040 c2        1 ABI       -0.601812
     26 C2L        -0.7290    3.6880   -0.6310 c         1 ABI        0.727816
     27 C13L        6.5130    1.7410    0.3980 c         1 ABI        0.729389
     28 C11L        2.4090    5.7350   -0.1780 ca        1 ABI       -0.391046
     29 C12L        1.1010    5.3660   -0.3680 ca        1 ABI       -0.270639
     30 C3L         0.7250    3.9920   -0.4200 ca        1 ABI        0.032919
     31 C4L         1.7040    3.0230   -0.2850 ca        1 ABI       -0.448024
     32 C5L         3.0670    3.3740   -0.0940 ca        1 ABI        0.282537
     33 C10L        3.4300    4.7570   -0.0340 ca        1 ABI        0.184025
     34 C9L         4.7950    5.0940    0.1670 ca        1 ABI       -0.391404
     35 C6L         4.0770    2.3860    0.0400 ca        1 ABI       -0.443389
     36 C7L         5.4030    2.7360    0.2340 ca        1 ABI        0.030775
     37 C8L         5.7500    4.1160    0.3010 ca        1 ABI       -0.270148
     38 H3         -2.7990   -2.5650   -0.3440 H         1 ABI        0.465101
     39 O4         -1.2360   -0.6830   -0.3210 O         1 ABI       -0.749021
     40 C4         -2.4310   -0.5430   -0.5220 C         1 ABI        0.900570
     41 O2         -5.2610   -2.8010   -0.6520 O         1 ABI       -0.764019
     42 N3         -3.2630   -1.6830   -0.5350 NA        1 ABI       -0.701221
     43 C6         -4.5020    0.6440   -0.9710 CM        1 ABI       -0.123054
     44 C2         -4.6230   -1.7570   -0.7240 C         1 ABI        0.907322
     45 N1         -5.2180   -0.5260   -1.0120 N*        1 ABI        0.289300
     46 C5         -3.1540    0.7150   -0.7570 CM        1 ABI       -0.250384
     47 H3"         5.2750   -4.6740   -0.4930 H         1 ABI        0.469041
     48 O4"         4.8430   -2.2620   -0.3930 O         1 ABI       -0.754807
     49 C4"         5.9340   -2.7540   -0.1440 C         1 ABI        0.906914
     50 O2"         7.2730   -6.1190   -0.0880 O         1 ABI       -0.771815
     51 N3"         6.1000   -4.1490   -0.2240 NA        1 ABI       -0.717334
     52 C6"         8.2820   -2.7840    0.4830 CM        1 ABI       -0.116985
     53 C2"         7.2320   -4.9060    0.0110 C         1 ABI        0.909228
     54 N1"         8.3500   -4.1450    0.3800 N*        1 ABI        0.287088
     55 C5"         7.1580   -2.0380    0.2480 CM        1 ABI       -0.281863
     56 C7'        -7.3660   -0.2750    0.1680 C         1 ABI        0.933011
     57 O7'        -6.8650    0.5350    0.9390 O         1 ABI       -0.730295
     58 N4'        -8.5110   -0.9630    0.4500 N         1 ABI       -0.100153
     59 C3'        -9.2040   -1.8240   -0.5190 CT        1 ABI       -0.403446
     60 H3'1      -10.2820   -1.6700   -0.3880 H1        1 ABI        0.210189
     61 H3'2       -8.9840   -1.5090   -1.5410 H1        1 ABI        0.210189
     62 C2'        -8.8620   -3.3190   -0.3610 CT        1 ABI        0.240285
     63 H2'1       -7.7820   -3.4520   -0.4490 H1        1 ABI        0.102330
     64 H2'2       -9.1460   -3.6540    0.6470 H1        1 ABI        0.102330
     65 N1'        -9.5290   -4.1110   -1.3760 N         1 ABI       -0.877097
     66 H1'1      -10.5430   -4.1020   -1.3960 H         1 ABI        0.480588
     67 C5'        -9.2550   -0.6230    1.6760 CT        1 ABI       -0.570402
     68 H5'1       -9.8900    0.2520    1.4990 H1        1 ABI        0.288810
     69 H5'2       -9.9070   -1.4740    1.8990 H1        1 ABI        0.288810
     70 C'         -8.4740   -0.2770    2.9590 C         1 ABI        0.622967
     71 O1'        -8.7820    0.7300    3.5790 O         1 ABI       -0.747267
@<TRIPOS>BOND
     1    1   18 1
     2    2   18 1
     3    3   18 1
     4    4   19 1
     5    5   20 1
     6    6   20 1
     7    7   22 1
     8    8   23 1
```

```
 9    9   25 1
10   10   28 1
11   11   29 1
12   12   31 1
13   13   34 1
14   14   35 1
15   15   37 1
16   16   43 1
17   17   52 1
18   18   54 1
19   19   25 1
20   19   55 1
21   20   45 1
22   20   56 1
23   21   27 1
24   22   23 1
25   22   46 1
26   23   26 1
27   24   26 1
28   25   27 1
29   26   30 1
30   27   36 1
31   28   29 1
32   28   33 1
33   29   30 1
34   30   31 1
35   31   32 1
36   32   33 1
37   32   35 1
38   33   34 1
39   34   37 1
40   35   36 1
41   36   37 1
42   38   42 1
43   39   40 1
44   40   42 1
45   40   46 1
46   41   44 1
47   42   44 1
48   43   45 1
49   43   46 1
50   44   45 1
51   47   51 1
52   48   49 1
53   49   51 1
54   49   55 1
55   50   53 1
56   51   53 1
57   52   54 1
58   52   55 1
59   53   54 1
60   56   57 1
61   56   58 1
62   58   59 1
63   58   67 1
64   59   60 1
65   59   61 1
66   59   62 1
67   62   63 1
68   62   64 1
69   62   65 1
70   65   66 1
```

```
    71   67   68 1
    72   67   69 1
    73   67   70 1
    74   70   71 1
@<TRIPOS>SUBSTRUCTURE
     1 ABI         1 TEMP              0 ****  ****    0 ROOT
@<TRIPOS>HEADTAIL
N1' 1
C' 1
@<TRIPOS>RESIDUECONNECT
1 N1' C' 0 0 0 0
```

frcmod.pna

```
PNA frcmod file: frcmod.pna
MASS

BOND
CQ-NA        502.0    1.324        by analogy to CQ-NC, parm99.dat
CM-c         349.7    1.487        SOURCE1  from gaff.dat
n3-Zn         50.4    2.089        ZAFF
c3-N*        337.0    1.475        JCC,7,(1986),230; ADE,CYT,GUA,THY,URA
CM-c3        317.0    1.510        by analogy parm99
c2-CM        357.2    1.4800       SOUECE3       1

ANGL
C -CT-N*    50.00     110.13       FF cst by analogy to CT-CT-N*; Eq.val. from
QM geo.opt.
CQ-NA-C     70.0      118.60       by analogy to CA-NC-CQ, parm99.dat
H5-CQ-NA    50.0      115.45       by analogy to H5-CQ-NC, parm99.dat
NC-CQ-NA    70.0      129.10       by analogy to NC-CQ-NC, parm99.dat
CQ-NA-H     50.0      118.00       by analogy to CA-NA-H, parm99.dat
CM-c -o     68.670    123.44       SOURCE3  from gaff.dat
c -CM-C     62.650    120.000    SOURCE3  from gaff.dat
c -CM-CM    64.640    120.140      SOURCE3  from gaff.dat
n -c -CM    68.470    115.140      SOURCE4  from gaff.dat
N*-c3-h1    50.0      109.50       parm99
N*-c3-c3    67.530    112.140    SOURCE4
C-N*-c3     70.0      117.60       parm99
CM-N*-c3    70.0      121.20       parm99
CM-CM-c3    70.0      119.70       parm99
CM-c3-n3    66.18     112.13     SOURCE4          387 1.2309
CM-c3-h1    46.78     110.95     SOURCE3           12 1.1170
C -CM-c3    70.0      119.70       by parm99
c2-c2-CM    66.880    117.000    SOURCE3            1
c2-c -ca    67.170    116.780    SOURCE3            1
c2-CM-C     67.930    120.700    SOURCE3            1
c2-CM-CM    64.690    120.600    SOURCE3
ha-c2-CM    50.040    120.940    SOURCE3          254 1.3150

DIHEDRAL
X -CQ-NA-X    4    6.00        180.0         2.0      by analogy to X -CA-NA-X,
parm99.dat
CM-CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
CM-CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
C -CM-c -o    1    2.175       180.0         -2.      Junmei et al, 1999
C -CM-c -o    1    0.30          0.0         3.       Junmei et al, 1999
n -c -CM-C    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
```

```
CM-CM-c -n    4    4.000       180.0         2.       optimized by Junmei Wang,
Jan-2013
X -c3-CM-X    6    0.000         0.000       2.       JCC,7,(1986),230
X -c2-CM-X    4   26.600       180.000       2.000    c2=c2 double bond,
intrpol.bsd.on C6H6

NONB
```

8.7. *Ab initio* study on model4 analogue

8.7.1. Model4 in vacuum

```
HETATM    1  H           0       8.868  -3.569  -0.625                       H
HETATM    2  H           0       9.028  -2.258   0.579                       H
HETATM    3  H           0       8.505  -3.891   1.080                       H
HETATM    4  H           0       5.896   1.039   0.438                       H
HETATM    5  H           0      -5.896   1.038  -0.439                       H
HETATM    6  H           0      -2.988   0.438   0.252                       H
HETATM    7  H           0       2.988   0.438  -0.251                       H
HETATM    8  H           0      -1.247   6.360  -0.082                       H
HETATM    9  H           0      -3.379   5.087  -0.218                       H
HETATM   10  H           0      -1.209   1.383  -0.069                       H
HETATM   11  H           0       1.247   6.360   0.083                       H
HETATM   12  H           0       1.209   1.383   0.068                       H
HETATM   13  H           0       3.379   5.087   0.219                       H
HETATM   14  H           0      -7.426  -0.738  -0.477                       H
HETATM   15  H           0       7.426  -0.737   0.474                       H
HETATM   16  C           0       8.453  -3.139   0.290                       C
HETATM   17  C           0       5.056   0.399   0.167                       C
HETATM   18  O           0       4.782   3.150   0.459                       O
HETATM   19  C           0      -5.056   0.399  -0.168                       C
HETATM   20  C           0      -3.862   1.011  -0.024                       C
HETATM   21  O           0      -4.782   3.150  -0.460                       O
HETATM   22  C           0       3.862   1.011   0.025                       C
HETATM   23  C           0      -3.776   2.477  -0.228                       C
HETATM   24  C           0       3.776   2.477   0.227                       C
HETATM   25  C           0      -1.245   5.272  -0.078                       C
HETATM   26  C           0      -2.425   4.576  -0.150                       C
HETATM   27  C           0      -2.440   3.152  -0.136                       C
HETATM   28  C           0      -1.238   2.468  -0.063                       C
HETATM   29  C           0      -0.000   3.160   0.000                       C
HETATM   30  C           0      -0.000   4.592   0.000                       C
HETATM   31  C           0       1.245   5.272   0.079                       C
HETATM   32  C           0       1.238   2.468   0.063                       C
HETATM   33  C           0       2.440   3.151   0.136                       C
HETATM   34  C           0       2.425   4.576   0.150                       C
HETATM   35  H           0      -4.204  -4.034   0.735                       H
HETATM   36  O           0      -3.168  -1.816   0.585                       O
HETATM   37  C           0      -4.354  -2.012   0.371                       C
HETATM   38  O           0      -6.517  -4.918   0.402                       O
HETATM   39  N           0      -4.872  -3.316   0.477                       N
HETATM   40  C           0      -6.647  -1.441  -0.200                       C
HETATM   41  C           0      -6.167  -3.758   0.282                       C
HETATM   42  N           0      -7.060  -2.738  -0.073                       N
HETATM   43  C           0      -5.363  -1.009  -0.004                       C
HETATM   44  H           0       4.204  -4.035  -0.733                       H
```

```
HETATM   45  O           0        3.168  -1.816  -0.583                       O
HETATM   46  C           0        4.354  -2.012  -0.370                       C
HETATM   47  O           0        6.517  -4.918  -0.401                       O
HETATM   48  N           0        4.872  -3.316  -0.476                       N
HETATM   49  C           0        6.647  -1.441   0.198                       C
HETATM   50  C           0        6.167  -3.758  -0.282                       C
HETATM   51  N           0        7.061  -2.737   0.072                       N
HETATM   52  C           0        5.363  -1.009   0.003                       C
HETATM   53  C           0       -8.453  -3.139  -0.292                       C
HETATM   54  H           0       -8.504  -3.895  -1.079                       H
HETATM   55  H           0       -8.870  -3.565   0.624                       H
HETATM   56  H           0       -9.027  -2.259  -0.585                       H
END
```

8.7.2. Linear analogue of Model4 in vacuum

```
HETATM    1  H           0       11.543  -1.904   0.886                       H
HETATM    2  H           0       10.438  -2.993  -0.000                       H
HETATM    3  H           0       11.543  -1.903  -0.886                       H
HETATM    4  H           0        6.014  -1.938  -0.000                       H
HETATM    5  H           0       -6.014  -1.938  -0.000                       H
HETATM    6  H           0       -5.079   0.962   0.000                       H
HETATM    7  H           0        5.079   0.962   0.000                       H
HETATM    8  H           0       -1.253   3.227  -0.000                       H
HETATM    9  H           0       -3.363   2.001  -0.000                       H
HETATM   10  H           0       -1.256  -1.745   0.000                       H
HETATM   11  H           0        1.253   3.227  -0.000                       H
HETATM   12  H           0        1.256  -1.745  -0.000                       H
HETATM   13  H           0        3.363   2.001  -0.000                       H
HETATM   14  H           0       -8.319  -2.364   0.000                       H
HETATM   15  H           0        8.319  -2.364  -0.000                       H
HETATM   16  C           0       10.913  -2.011  -0.000                       C
HETATM   17  C           0        6.145  -0.856  -0.000                       C
HETATM   18  O           0        3.630  -2.034  -0.000                       O
HETATM   19  C           0       -6.145  -0.856   0.000                       C
HETATM   20  C           0       -5.016  -0.116   0.000                       C
HETATM   21  O           0       -3.630  -2.034   0.000                       O
HETATM   22  C           0        5.016  -0.116  -0.000                       C
HETATM   23  C           0       -3.703  -0.805  -0.000                       C
HETATM   24  C           0        3.703  -0.805  -0.000                       C
HETATM   25  C           0       -1.246   2.140  -0.000                       C
HETATM   26  C           0       -2.433   1.444  -0.000                       C
HETATM   27  C           0       -2.446   0.019  -0.000                       C
HETATM   28  C           0       -1.238  -0.660  -0.000                       C
HETATM   29  C           0        0.000   0.027  -0.000                       C
HETATM   30  C           0        0.000   1.461  -0.000                       C
HETATM   31  C           0        1.246   2.140  -0.000                       C
HETATM   32  C           0        1.238  -0.660  -0.000                       C
HETATM   33  C           0        2.446   0.019  -0.000                       C
HETATM   34  C           0        2.433   1.444  -0.000                       C
HETATM   35  H           0       -9.554   2.239  -0.000                       H
HETATM   36  O           0       -7.114   1.992   0.000                       O
HETATM   37  C           0       -7.879   1.039   0.000                       C
HETATM   38  O           0      -11.481   0.649   0.000                       O
HETATM   39  N           0       -9.268   1.266  -0.000                       N
HETATM   40  C           0       -8.536  -1.300   0.000                       C
HETATM   41  C           0      -10.301   0.347   0.000                       C
HETATM   42  N           0       -9.866  -0.985   0.000                       N
```

```
HETATM   43  C           0       -7.518  -0.387   0.000                       C
HETATM   44  H           0        9.554   2.239   0.000                       H
HETATM   45  O           0        7.114   1.992   0.000                       O
HETATM   46  C           0        7.879   1.039   0.000                       C
HETATM   47  O           0       11.481   0.649   0.000                       O
HETATM   48  N           0        9.268   1.266   0.000                       N
HETATM   49  C           0        8.536  -1.300  -0.000                       C
HETATM   50  C           0       10.301   0.347   0.000                       C
HETATM   51  N           0        9.866  -0.985   0.000                       N
HETATM   52  C           0        7.518  -0.387   0.000                       C
HETATM   53  C           0      -10.913  -2.011   0.000                       C
HETATM   54  H           0      -11.543  -1.903   0.886                       H
HETATM   55  H           0      -11.543  -1.904  -0.886                       H
HETATM   56  H           0      -10.438  -2.993   0.000                       H
END
```

8.7.3. PES scan

```
56
 generated by VMD
 H        11.542610       -1.903560        0.886190
 H        10.438060       -2.992710       -0.000150
 H        11.542580       -1.903400       -0.886340
 H         6.013820       -1.937900       -0.000060
 H        -6.013820       -1.937900       -0.000010
 H        -5.078510        0.962270        0.000010
 H         5.078510        0.962270        0.000010
 H        -1.252530        3.227170       -0.000130
 H        -3.363070        2.000880       -0.000090
 H        -1.255840       -1.745470        0.000000
 H         1.252530        3.227170       -0.000150
 H         1.255840       -1.745470       -0.000040
 H         3.363070        2.000880       -0.000140
 H        -8.318880       -2.363730        0.000040
 H         8.318880       -2.363730       -0.000060
 C        10.912850       -2.010700       -0.000070
 C         6.145290       -0.855720       -0.000020
 O         3.630270       -2.034090       -0.000120
 C        -6.145290       -0.855720        0.000000
 C        -5.016160       -0.116180        0.000010
 O        -3.630270       -2.034090        0.000030
 C         5.016160       -0.116180       -0.000030
 C        -3.702580       -0.804700       -0.000000
 C         3.702580       -0.804700       -0.000080
 C        -1.246410        2.139950       -0.000110
 C        -2.432810        1.443970       -0.000080
 C        -2.445500        0.018690       -0.000040
 C        -1.238140       -0.659870       -0.000030
 C         0.000000        0.027320       -0.000060
 C         0.000000        1.461350       -0.000100
 C         1.246400        2.139950       -0.000120
 C         1.238140       -0.659870       -0.000060
 C         2.445500        0.018690       -0.000080
 C         2.432810        1.443970       -0.000110
 H        -9.554090        2.239140       -0.000010
 O        -7.113690        1.991540        0.000010
 C        -7.878930        1.039450        0.000010
 O       -11.480560        0.649090        0.000070
 N        -9.267580        1.266150       -0.000010
```

```
 C        -8.536430       -1.300440        0.000030
 C       -10.300730        0.347200        0.000040
 N        -9.866310       -0.984990        0.000050
 C        -7.517530       -0.386510        0.000020
 H         9.554090        2.239140        0.000230
 O         7.113690        1.991540        0.000110
 C         7.878930        1.039450        0.000100
 O        11.480570        0.649090        0.000110
 N         9.267580        1.266150        0.000190
 C         8.536430       -1.300440       -0.000000
 C        10.300730        0.347200        0.000110
 N         9.866310       -0.984990        0.000030
 C         7.517530       -0.386510        0.000030
 C       -10.912850       -2.010700        0.000120
 H       -11.542560       -1.903450        0.886410
 H       -11.542630       -1.903510       -0.886120
 H       -10.438070       -2.992710        0.000130
56
 generated by VMD
 H        11.562535       -1.909432        0.664069
 H        10.432876       -2.984544       -0.207638
 H        11.509879       -1.878929       -1.107419
 H         6.010105       -1.935086       -0.059218
 H        -6.024895       -1.926991        0.241249
 H        -5.057913        0.945289       -0.085641
 H         5.072131        0.963326        0.014764
 H        -1.255451        3.212644        0.363910
 H        -3.366883        1.981674        0.354621
 H        -1.256990       -1.754205        0.108290
 H         1.248176        3.216818        0.290904
 H         1.253608       -1.749342        0.039001
 H         3.358725        1.995501        0.177644
 H        -8.330005       -2.343455        0.168376
 H         8.314178       -2.358075       -0.134559
 C        10.906809       -2.002123       -0.204832
 C         6.140617       -0.852883       -0.045281
 O         3.627877       -2.034862       -0.006637
 C        -6.145649       -0.852951        0.098599
 C        -5.010260       -0.123967        0.062561
 O        -3.639709       -2.038301        0.313308
 C         5.011219       -0.115129        0.000376
 C        -3.705075       -0.813072        0.209360
 C         3.699073       -0.805878        0.027473
 C        -1.249471        2.127311        0.299916
 C        -2.435273        1.430418        0.288966
 C        -2.446761        0.007323        0.212885
 C        -1.240173       -0.669558        0.155321
 C        -0.002672        0.019090        0.157361
 C        -0.003395        1.451347        0.230753
 C         1.242467        2.131051        0.234634
 C         1.235356       -0.665074        0.092579
 C         2.442058        0.014635        0.096896
 C         2.428826        1.438029        0.169561
 H        -9.521233        2.229643       -0.448962
 O        -7.086403        1.976682       -0.291007
 C        -7.859368        1.033721       -0.218484
 O       -11.459863        0.657146       -0.355952
 N        -9.244064        1.262786       -0.317931
 C        -8.537544       -1.287340        0.028643
 C       -10.284307        0.353815       -0.260014
 N        -9.862734       -0.970083       -0.077903
 C        -7.511582       -0.383141       -0.029551
```

```
 H         9.545877        2.245529       -0.091930
 O         7.106631        1.995129       -0.025033
 C         7.872188        1.044104       -0.063893
 O        11.472579        0.657899       -0.175022
 N         9.260121        1.272349       -0.100339
 C         8.530962       -1.294686       -0.122882
 C        10.293433        0.354733       -0.145994
 N         9.860065       -0.977781       -0.156198
 C         7.511890       -0.382059       -0.077718
 C       -10.917870       -1.984330       -0.004574
 H       -11.581774       -1.774574        0.837415
 H       -11.510248       -1.974858       -0.922358
 H       -10.452671       -2.962264        0.125701
56
 generated by VMD
 H        11.564872       -1.920660        0.385697
 H        10.402583       -2.974864       -0.468556
 H        11.445558       -1.847413       -1.381267
 H         5.990149       -1.931088       -0.116298
 H        -6.050785       -1.891751        0.490593
 H        -5.010286        0.915433       -0.094130
 H         5.055054        0.964498        0.049497
 H        -1.255373        3.197792        0.689924
 H        -3.371496        1.962560        0.692584
 H        -1.266009       -1.761322        0.292404
 H         1.243844        3.204519        0.541478
 H         1.240685       -1.751546        0.139963
 H         3.349988        1.988548        0.329539
 H        -8.349317       -2.299431        0.289139
 H         8.288795       -2.351156       -0.293934
 C        10.876833       -1.992633       -0.459878
 C         6.121225       -0.849277       -0.086233
 O         3.612668       -2.034578        0.023433
 C        -6.144361       -0.839724        0.220182
 C        -4.996713       -0.130379        0.179522
 O        -3.667434       -2.020805        0.709213
 C         4.994316       -0.113629        0.017574
 C        -3.714430       -0.811943        0.481676
 C         3.684495       -0.806489        0.080930
 C        -1.252713        2.115348        0.588067
 C        -2.438814        1.418963        0.583156
 C        -2.451646       -0.000706        0.463372
 C        -1.248821       -0.677778        0.361960
 C        -0.011322        0.011597        0.347142
 C        -0.009717        1.441224        0.463099
 C         1.236169        2.120979        0.451988
 C         1.224100       -0.669345        0.226056
 C         2.430545        0.010743        0.214316
 C         2.419975        1.431175        0.330745
 H        -9.426795        2.188998       -0.896720
 O        -7.013661        1.938455       -0.532922
 C        -7.803335        1.013867       -0.418381
 O       -11.389146        0.647411       -0.783168
 N        -9.173927        1.240081       -0.642913
 C        -8.531181       -1.262877        0.023640
 C       -10.228373        0.349095       -0.566972
 N        -9.840441       -0.950154       -0.212997
 C        -7.490683       -0.377012       -0.055296
 H         9.523285        2.250913       -0.200602
 O         7.088428        1.998159       -0.043331
 C         7.851433        1.048498       -0.132663
 O        11.444517        0.666203       -0.394033
```

```
 N          9.236987        1.278037       -0.218911
 C          8.506360       -1.288179       -0.268051
 C         10.267275        0.362048       -0.325209
 N          9.833222       -0.970097       -0.347451
 C          7.490275       -0.377157       -0.162887
 C        -10.912106       -1.944771       -0.114826
 H        -11.643213       -1.634405        0.635234
 H        -11.422809       -2.041469       -1.075769
 H        -10.473832       -2.902302        0.169365
56
 generated by VMD
 H         11.536103       -1.951524        0.194675
 H         10.351790       -2.984182       -0.655672
 H         11.381438       -1.844047       -1.567787
 H          5.953333       -1.928209       -0.184694
 H         -6.071867       -1.839372        0.695814
 H         -4.962651        0.898644       -0.073289
 H          5.035370        0.967245        0.060700
 H         -1.241833        3.205326        0.960559
 H         -3.367445        1.973998        0.983763
 H         -1.285300       -1.743169        0.434261
 H          1.250109        3.206712        0.730540
 H          1.213710       -1.736175        0.195019
 H          3.343792        1.989889        0.422506
 H         -8.354798       -2.259652        0.377030
 H          8.245248       -2.354755       -0.421042
 C         10.830770       -2.004375       -0.637990
 C          6.090112       -0.847750       -0.136708
 O          3.579662       -2.024940        0.004297
 C         -6.138692       -0.815633        0.326518
 C         -4.982996       -0.118901        0.293275
 O         -3.701827       -1.957246        1.079144
 C          4.969326       -0.109815        0.008844
 C         -3.728030       -0.777338        0.731118
 C          3.658545       -0.798975        0.090983
 C         -1.248179        2.126624        0.824769
 C         -2.437004        1.434960        0.833679
 C         -2.458525        0.020082        0.671954
 C         -1.263823       -0.661013        0.523400
 C         -0.024194        0.024863        0.486581
 C         -0.013041        1.450801        0.641965
 C          1.235094        2.126065        0.611720
 C          1.204074       -0.656602        0.310330
 C          2.412669        0.019219        0.281346
 C          2.411844        1.435923        0.436167
 H         -9.320211        2.115601       -1.247705
 O         -6.936155        1.889675       -0.713545
 C         -7.738975        0.980567       -0.571522
 O        -11.297923        0.593894       -1.128816
 N         -9.091253        1.190813       -0.899609
 C         -8.511213       -1.248820        0.014104
 C        -10.155292        0.312102       -0.815718
 N         -9.800488       -0.953741       -0.329870
 C         -7.460900       -0.375400       -0.073320
 H          9.503080        2.239525       -0.273344
 O          7.071379        1.994368       -0.064438
 C          7.827607        1.043244       -0.188681
 O         11.412062        0.650367       -0.538570
 N          9.211893        1.268362       -0.302775
 C          8.468356       -1.293360       -0.381170
 C         10.235229        0.350073       -0.448718
 N          9.794563       -0.979577       -0.485061
```

```
 C         7.459176       -0.380061       -0.236514
 C       -10.883851       -1.934106       -0.217219
 H       -11.662189       -1.556270        0.449665
 H       -11.328336       -2.116379       -1.198517
 H       -10.473323       -2.862757        0.181275
56
 generated by VMD
 H        11.486932       -2.014568       -0.036404
 H        10.270168       -3.007380       -0.888401
 H        11.287019       -1.847268       -1.789607
 H         5.895320       -1.924897       -0.274947
 H        -6.095401       -1.757506        0.910244
 H        -4.910124        0.888119       -0.053808
 H         5.015460        0.968687        0.098261
 H        -1.215368        3.238671        1.210322
 H        -3.354996        2.021345        1.257201
 H        -1.315630       -1.698290        0.580273
 H         1.266924        3.224875        0.902521
 H         1.172345       -1.703393        0.255056
 H         3.339906        1.999689        0.503006
 H        -8.356378       -2.204021        0.493412
 H         8.176081       -2.365529       -0.582727
 C        10.759435       -2.033225       -0.851223
 C         6.045018       -0.848314       -0.191641
 O         3.526947       -2.005301       -0.025544
 C        -6.130517       -0.775453        0.437515
 C        -4.968140       -0.089569        0.406865
 O        -3.746491       -1.840198        1.454724
 C         4.936291       -0.104997        0.008498
 C        -3.746944       -0.705676        0.980395
 C         3.620524       -0.783608        0.099090
 C        -1.236025        2.163860        1.047748
 C        -2.430206        1.481846        1.073219
 C        -2.467996        0.072275        0.877286
 C        -1.285591       -0.617634        0.685669
 C        -0.041424        0.059752        0.624418
 C        -0.013320        1.481535        0.812323
 C         1.239262        2.147341        0.759941
 C         1.175117       -0.626767        0.395723
 C         2.388230        0.039904        0.346179
 C         2.404384        1.452511        0.533264
 H        -9.183393        1.984819       -1.624500
 O        -6.840531        1.812807       -0.915667
 C        -7.657220        0.923250       -0.735717
 O       -11.176404        0.485428       -1.477993
 N        -8.983974        1.099751       -1.170923
 C        -8.481833       -1.235012        0.020846
 C       -10.057936        0.235576       -1.066151
 N        -9.744176       -0.975082       -0.433387
 C        -7.421585       -0.375937       -0.086951
 H         9.486742        2.207825       -0.304429
 O         7.058625        1.979601       -0.043181
 C         7.801131        1.025995       -0.219988
 O        11.371368        0.610354       -0.673582
 N         9.184501        1.241257       -0.360916
 C         8.411632       -1.308498       -0.510957
 C        10.193962        0.318686       -0.564756
 N         9.738243       -1.004398       -0.636713
 C         7.416211       -0.390982       -0.308939
 C       -10.840567       -1.936727       -0.289525
 H       -11.646867       -1.500320        0.304724
 H       -11.238932       -2.199242       -1.272236
```

```
 H       -10.457870       -2.830128        0.205574
56
 generated by VMD
 H        11.390501       -2.184673       -0.281498
 H        10.146338       -3.042139       -1.234798
 H        11.197776       -1.820410       -2.005399
 H         5.802499       -1.903040       -0.504267
 H        -6.103592       -1.622231        1.164922
 H        -4.870404        0.892306       -0.060462
 H         5.012090        0.953818        0.203640
 H        -1.167469        3.347423        1.303524
 H        -3.322340        2.156849        1.440835
 H        -1.356102       -1.592214        0.708302
 H         1.295997        3.300630        0.878368
 H         1.111353       -1.629353        0.265957
 H         3.334259        2.046675        0.393996
 H        -8.339573       -2.137788        0.693801
 H         8.070762       -2.375841       -0.858442
 C        10.663789       -2.092646       -1.092014
 C         5.984888       -0.847965       -0.299443
 O         3.435703       -1.947406       -0.211667
 C        -6.117173       -0.707199        0.571764
 C        -4.954926       -0.022562        0.512556
 O        -3.782695       -1.609866        1.844178
 C         4.899311       -0.098332       -0.014236
 C        -3.762442       -0.554106        1.215745
 C         3.563045       -0.741790        0.005416
 C        -1.209563        2.272448        1.146155
 C        -2.410396        1.606312        1.225304
 C        -2.474640        0.197442        1.035405
 C        -1.311658       -0.510855        0.802896
 C        -0.062241        0.150374        0.682667
 C        -0.007532        1.572474        0.861273
 C         1.249638        2.222118        0.748552
 C         1.134241       -0.552671        0.404608
 C         2.353036        0.097600        0.302936
 C         2.395442        1.510980        0.478711
 H        -9.067243        1.746274       -1.971140
 O        -6.763861        1.685682       -1.127233
 C        -7.586271        0.818184       -0.879562
 O       -11.061088        0.258586       -1.739632
 N        -8.889215        0.926659       -1.400730
 C        -8.443050       -1.236123        0.098516
 C        -9.964817        0.072803       -1.243057
 N        -9.681224       -1.045878       -0.447261
 C        -7.380765       -0.387848       -0.062496
 H         9.516218        2.097912       -0.079349
 O         7.082272        1.914791        0.158408
 C         7.796327        0.964551       -0.122820
 O        11.353365        0.495754       -0.625234
 N         9.185633        1.152816       -0.241551
 C         8.337616       -1.340726       -0.670524
 C        10.167628        0.228851       -0.547299
 N         9.672996       -1.064092       -0.763864
 C         7.369547       -0.422023       -0.366605
 C       -10.780919       -1.991989       -0.239910
 H       -11.625899       -1.485290        0.231838
 H       -11.112674       -2.397657       -1.198591
 H       -10.428570       -2.800394        0.401997
56
 generated by VMD
 H        11.294396       -2.312974       -0.464550
```

```
 H        10.016449       -3.097900       -1.435300
 H        11.086329       -1.868820       -2.167843
 H         5.709028       -1.882935       -0.620764
 H        -6.117898       -1.477838        1.352397
 H        -4.822178        0.889290       -0.086695
 H         4.998381        0.951705        0.243987
 H        -1.109265        3.445127        1.503131
 H        -3.283509        2.290409        1.665275
 H        -1.406063       -1.472108        0.776978
 H         1.341011        3.363992        1.012397
 H         1.043256       -1.541456        0.264313
 H         3.342282        2.085307        0.442028
 H        -8.327491       -2.054105        0.842857
 H         7.961647       -2.396274       -1.015128
 C        10.558773       -2.168073       -1.259123
 C         5.920486       -0.844907       -0.363470
 O         3.345902       -1.884895       -0.301030
 C        -6.102425       -0.636637        0.658444
 C        -4.938271        0.043222        0.579324
 O        -3.834336       -1.369167        2.147175
 C         4.857038       -0.083992       -0.029151
 C        -3.783504       -0.398165        1.397283
 C         3.505125       -0.694368       -0.028528
 C        -1.176179        2.375908        1.317398
 C        -2.387126        1.729941        1.412022
 C        -2.483485        0.328983        1.184765
 C        -1.342679       -0.394369        0.900073
 C        -0.083823        0.247124        0.763169
 C         0.003867        1.661876        0.980888
 C         1.270568        2.290577        0.854126
 C         1.090834       -0.470001        0.433993
 C         2.319704        0.159261        0.321636
 C         2.395267        1.565837        0.536417
 H        -8.925651        1.483613       -2.294988
 O        -6.664986        1.535352       -1.340850
 C        -7.498899        0.699392       -1.031699
 O       -10.928598        0.022440       -1.984355
 N        -8.775552        0.738390       -1.623453
 C        -8.402444       -1.229769        0.140721
 C        -9.857672       -0.096616       -1.417166
 N        -9.613202       -1.111333       -0.481050
 C        -7.333484       -0.400552       -0.068730
 H         9.528343        1.996296       -0.034326
 O         7.092509        1.863405        0.218204
 C         7.779529        0.910228       -0.115476
 O        11.318910        0.376480       -0.674762
 N         9.172157        1.068862       -0.238676
 C         8.256570       -1.378748       -0.779269
 C        10.127433        0.136130       -0.597883
 N         9.597934       -1.132031       -0.871059
 C         7.315136       -0.451614       -0.422147
 C       -10.721077       -2.033821       -0.217813
 H       -11.591396       -1.479410        0.140315
 H       -10.998979       -2.558464       -1.135074
 H       -10.403356       -2.753215        0.537979
56
 generated by VMD
 H        11.165831       -2.431311       -0.702036
 H         9.845806       -3.149990       -1.667907
 H        10.926145       -1.919215       -2.381953
 H         5.589408       -1.866551       -0.702822
 H        -6.157228       -1.303771        1.538817
```

```
 H        -4.739594        0.857157       -0.098518
 H         4.967806        0.949565        0.282908
 H        -1.053106        3.551016        1.723284
 H        -3.246070        2.438441        1.922876
 H        -1.471159       -1.343366        0.905177
 H         1.381536        3.428571        1.169329
 H         0.960906       -1.453907        0.337382
 H         3.340070        2.120411        0.519763
 H        -8.333075       -1.938181        0.951711
 H         7.818684       -2.415827       -1.175608
 C        10.414391       -2.240101       -1.471527
 C         5.832058       -0.844053       -0.412858
 O         3.236941       -1.829607       -0.311581
 C        -6.084781       -0.559148        0.745233
 C        -4.910867        0.103594        0.660328
 O        -3.931637       -1.077353        2.483057
 C         4.795652       -0.071927       -0.023916
 C        -3.820076       -0.224826        1.607816
 C         3.430117       -0.651364       -0.009142
 C        -1.146777        2.486656        1.521662
 C        -2.368340        1.864125        1.637275
 C        -2.500062        0.470612        1.386523
 C        -1.384646       -0.270389        1.055349
 C        -0.114663        0.347045        0.901389
 C         0.008941        1.754867        1.142237
 C         1.284793        2.359988        0.993093
 C         1.035301       -0.387360        0.527299
 C         2.273686        0.218776        0.392860
 C         2.385273        1.618929        0.631565
 H        -8.672719        1.171554       -2.644886
 O        -6.482396        1.337955       -1.550346
 C        -7.345367        0.551188       -1.195763
 O       -10.708944       -0.232677       -2.291115
 N        -8.577703        0.517613       -1.875336
 C        -8.349588       -1.210460        0.146577
 C        -9.680620       -0.281785       -1.640890
 N        -9.512004       -1.171170       -0.569979
 C        -7.259736       -0.416602       -0.090783
 H         9.515795        1.894762       -0.057488
 O         7.083093        1.811095        0.243002
 C         7.738871        0.855490       -0.142536
 O        11.250834        0.259427       -0.802801
 N         9.132199        0.985067       -0.289981
 C         8.143890       -1.415791       -0.907072
 C        10.055646        0.045007       -0.708477
 N         9.488622       -1.197686       -1.020481
 C         7.234132       -0.481589       -0.491271
 C       -10.646230       -2.049194       -0.269462
 H       -11.526740       -1.452222       -0.020747
 H       -10.881965       -2.666727       -1.139235
 H       -10.379911       -2.685855        0.575224
56
 generated by VMD
 H        11.023061       -2.561355       -0.884480
 H         9.681513       -3.184041       -1.886461
 H        10.802689       -1.953413       -2.534871
 H         5.480412       -1.845752       -0.769373
 H        -6.210670       -1.094899        1.687839
 H        -4.635711        0.775512       -0.152025
 H         4.942783        0.949222        0.320442
 H        -0.996422        3.642311        1.928193
 H        -3.207832        2.569810        2.145123
```

```
 H        -1.538101       -1.212049        0.955053
 H         1.427837        3.480799        1.340610
 H         0.880772       -1.361269        0.360479
 H         3.346012        2.149312        0.620075
 H        -8.371660       -1.750889        1.076934
 H         7.685894       -2.429393       -1.304471
 C        10.279379       -2.304883       -1.642759
 C         5.752223       -0.839610       -0.449617
 O         3.137382       -1.768230       -0.334973
 C        -6.070387       -0.476732        0.800311
 C        -4.874544        0.142949        0.693990
 O        -4.048254       -0.763557        2.736086
 C         4.741195       -0.057095       -0.017189
 C        -3.856850       -0.058953        1.750928
 C         3.362571       -0.604936        0.001408
 C        -1.116339        2.585645        1.701528
 C        -2.348783        1.986128        1.824362
 C        -2.514723        0.603479        1.538047
 C        -1.426395       -0.148328        1.149945
 C        -0.143492        0.444586        0.998366
 C         0.016508        1.840309        1.281996
 C         1.303821        2.420828        1.132182
 C         0.982895       -0.303306        0.582717
 C         2.233148        0.278260        0.448664
 C         2.381358        1.666830        0.730654
 H        -8.409360        0.790470       -2.955692
 O        -6.275637        1.060430       -1.772647
 C        -7.183821        0.360101       -1.355726
 O       -10.509768       -0.489601       -2.515473
 N        -8.378846        0.257959       -2.092904
 C        -8.319777       -1.152310        0.172997
 C        -9.519988       -0.467456       -1.806712
 N        -9.440226       -1.189781       -0.607214
 C        -7.191729       -0.432120       -0.115902
 H         9.507262        1.797321       -0.068125
 O         7.079202        1.760006        0.276932
 C         7.704382        0.803794       -0.155064
 O        11.188343        0.148789       -0.902514
 N         9.097456        0.905874       -0.325252
 C         8.040151       -1.447557       -1.007056
 C         9.990090       -0.040317       -0.793782
 N         9.387389       -1.257027       -1.139283
 C         7.161097       -0.507722       -0.541030
 C       -10.619031       -1.982096       -0.246071
 H       -11.492658       -1.332558       -0.155850
 H       -10.820998       -2.727414       -1.018971
 H       -10.426885       -2.478763        0.705853
56
 generated by VMD
 H        10.916286       -2.557835       -1.076633
 H         9.514832       -3.256622       -1.936655
 H        10.567876       -2.058557       -2.741313
 H         5.369055       -1.870067       -0.682879
 H        -6.316017       -0.867461        1.787568
 H        -4.467225        0.627331       -0.139490
 H         4.875640        0.974012        0.293632
 H        -0.959013        3.664366        2.268998
 H        -3.189817        2.630301        2.471476
 H        -1.622449       -1.123970        1.069099
 H         1.461393        3.467319        1.675689
 H         0.794214       -1.310914        0.483382
 H         3.342914        2.126401        0.884443
```

```
 H        -8.453477       -1.494144        1.067036
 H         7.545849       -2.473284       -1.306595
 C        10.115880       -2.357900       -1.792737
 C         5.652491       -0.848989       -0.426969
 O         3.046219       -1.767402       -0.167663
 C        -6.056972       -0.397165        0.838210
 C        -4.819690        0.140361        0.761150
 O        -4.244262       -0.468207        2.991417
 C         4.661891       -0.049205        0.020875
 C        -3.918958        0.059649        1.933770
 C         3.288196       -0.597988        0.134409
 C        -1.103037        2.619388        2.005014
 C        -2.347266        2.042748        2.116441
 C        -2.546046        0.676414        1.777726
 C        -1.482970       -0.077630        1.330397
 C        -0.184595        0.488802        1.204996
 C         0.010217        1.866281        1.548279
 C         1.310253        2.422090        1.416843
 C         0.921362       -0.265807        0.749345
 C         2.183797        0.292427        0.628290
 C         2.366838        1.662833        0.971934
 H        -7.959655        0.414520       -3.273190
 O        -5.927271        0.744704       -1.937160
 C        -6.914057        0.164222       -1.514993
 O       -10.170961       -0.674356       -2.863143
 N        -8.042121        0.015947       -2.344079
 C        -8.280852       -1.041268        0.095868
 C        -9.249469       -0.598595       -2.070942
 N        -9.325508       -1.138456       -0.778546
 C        -7.086539       -0.431376       -0.180432
 H         9.408637        1.813427       -0.377049
 O         7.003338        1.786307        0.101043
 C         7.609139        0.811928       -0.317807
 O        11.050620        0.131612       -1.222740
 N         8.990481        0.909214       -0.567583
 C         7.910375       -1.477201       -1.076507
 C         9.861367       -0.055854       -1.038462
 N         9.247662       -1.289998       -1.289944
 C         7.052619       -0.518064       -0.610111
 C       -10.582173       -1.805120       -0.426237
 H       -11.413116       -1.099038       -0.494216
 H       -10.773171       -2.630615       -1.115675
 H       -10.501863       -2.185132        0.593077
56
 generated by VMD
 H        10.916286       -2.557835       -1.076633
 H         9.514832       -3.256622       -1.936655
 H        10.567876       -2.058557       -2.741313
 H         5.369055       -1.870067       -0.682879
 H        -6.316017       -0.867461        1.787568
 H        -4.467225        0.627331       -0.139490
 H         4.875640        0.974012        0.293632
 H        -0.959013        3.664366        2.268998
 H        -3.189817        2.630301        2.471476
 H        -1.622449       -1.123970        1.069099
 H         1.461393        3.467319        1.675689
 H         0.794214       -1.310914        0.483382
 H         3.342914        2.126401        0.884443
 H        -8.453477       -1.494144        1.067036
 H         7.545849       -2.473284       -1.306595
 C        10.115880       -2.357900       -1.792737
 C         5.652491       -0.848989       -0.426969
```

```
 O         3.046219       -1.767402       -0.167663
 C        -6.056972       -0.397165        0.838210
 C        -4.819690        0.140361        0.761150
 O        -4.244262       -0.468207        2.991417
 C         4.661891       -0.049205        0.020875
 C        -3.918958        0.059649        1.933770
 C         3.288196       -0.597988        0.134409
 C        -1.103037        2.619388        2.005014
 C        -2.347266        2.042748        2.116441
 C        -2.546046        0.676414        1.777726
 C        -1.482970       -0.077630        1.330397
 C        -0.184595        0.488802        1.204996
 C         0.010217        1.866281        1.548279
 C         1.310253        2.422090        1.416843
 C         0.921362       -0.265807        0.749345
 C         2.183797        0.292427        0.628290
 C         2.366838        1.662833        0.971934
 H        -7.959655        0.414520       -3.273190
 O        -5.927271        0.744704       -1.937160
 C        -6.914057        0.164222       -1.514993
 O       -10.170961       -0.674356       -2.863143
 N        -8.042121        0.015947       -2.344079
 C        -8.280852       -1.041268        0.095868
 C        -9.249469       -0.598595       -2.070942
 N        -9.325508       -1.138456       -0.778546
 C        -7.086539       -0.431376       -0.180432
 H         9.408637        1.813427       -0.377049
 O         7.003338        1.786307        0.101043
 C         7.609139        0.811928       -0.317807
 O        11.050620        0.131612       -1.222740
 N         8.990481        0.909214       -0.567583
 C         7.910375       -1.477201       -1.076507
 C         9.861367       -0.055854       -1.038462
 N         9.247662       -1.289998       -1.289944
 C         7.052619       -0.518064       -0.610111
 C       -10.582173       -1.805120       -0.426237
 H       -11.413116       -1.099038       -0.494216
 H       -10.773171       -2.630615       -1.115675
 H       -10.501863       -2.185132        0.593077
56
 generated by VMD
 H        10.752528       -2.739484       -1.196435
 H         9.338100       -3.301784       -2.131718
 H        10.461488       -2.096634       -2.822493
 H         5.240221       -1.827651       -0.820634
 H        -6.383453       -0.517070        1.898384
 H        -4.340913        0.457890       -0.161593
 H         4.857290        0.955618        0.360532
 H        -0.908585        3.841133        2.244310
 H        -3.149327        2.845248        2.535184
 H        -1.696278       -0.942417        1.102136
 H         1.492742        3.589264        1.599215
 H         0.705294       -1.187775        0.476545
 H         3.329118        2.202445        0.778516
 H        -8.521059       -1.143536        1.184859
 H         7.397321       -2.481496       -1.460838
 C         9.976971       -2.446484       -1.907926
 C         5.564505       -0.841817       -0.486851
 O         2.921608       -1.671625       -0.289685
 C        -6.043695       -0.265065        0.893074
 C        -4.772509        0.181570        0.792166
 O        -4.372021        0.029767        3.135514
```

```
 C         4.603294       -0.033019        0.006509
 C        -3.950106        0.306900        2.017661
 C         3.205663       -0.527110        0.066057
 C        -1.077695        2.794319        2.003474
 C        -2.327422        2.241536        2.160224
 C        -2.559454        0.874057        1.847339
 C        -1.526916        0.099548        1.364053
 C        -0.219641        0.636939        1.206597
 C         0.009768        2.014123        1.529559
 C         1.316240        2.543516        1.359307
 C         0.859852       -0.143581        0.731466
 C         2.130116        0.387962        0.577596
 C         2.347726        1.758160        0.900886
 H        -7.664742       -0.228383       -3.414468
 O        -5.687189        0.257152       -2.043900
 C        -6.733885       -0.161669       -1.577116
 O        -9.965933       -1.076789       -2.935526
 N        -7.824465       -0.415624       -2.430425
 C        -8.266031       -0.916901        0.154421
 C        -9.085028       -0.887462       -2.116594
 N        -9.267878       -1.134288       -0.748059
 C        -7.018300       -0.450567       -0.162604
 H         9.433185        1.637859       -0.197208
 O         7.021918        1.684519        0.248787
 C         7.590518        0.716867       -0.233026
 O        11.011938       -0.050390       -1.144544
 N         8.978357        0.769925       -0.459073
 C         7.801942       -1.522202       -1.154298
 C         9.812853       -0.196687       -0.989049
 N         9.149073       -1.380109       -1.338326
 C         6.980488       -0.562008       -0.628508
 C       -10.584659       -1.639113       -0.349228
 H       -11.361403       -0.929645       -0.643340
 H       -10.784966       -2.594142       -0.840660
 H       -10.594240       -1.770944        0.733484
56
 generated by VMD
 H        10.613051       -2.877159       -1.231827
 H         9.174441       -3.419854       -2.141395
 H        10.313986       -2.247424       -2.861544
 H         5.126849       -1.839643       -0.795385
 H        -6.445934       -0.185131        1.907274
 H        -4.224625        0.330505       -0.132916
 H         4.821476        0.965425        0.356290
 H        -0.853774        4.000700        2.282224
 H        -3.118757        3.059215        2.597611
 H        -1.778444       -0.758154        1.139659
 H         1.538091        3.680699        1.633922
 H         0.617217       -1.074691        0.520530
 H         3.332910        2.243485        0.809205
 H        -8.560463       -0.847380        1.157919
 H         7.260418       -2.548508       -1.456264
 C         9.834448       -2.575938       -1.936499
 C         5.477330       -0.857551       -0.477480
 O         2.818641       -1.623833       -0.241191
 C        -6.027637       -0.130996        0.901433
 C        -4.732162        0.243222        0.819413
 O        -4.511780        0.517111        3.175341
 C         4.540818       -0.021493        0.018025
 C        -3.993690        0.563855        2.063458
 C         3.133237       -0.482805        0.099678
 C        -1.051821        2.957463        2.048060
```

```
 C        -2.313854        2.438035        2.215278
 C        -2.584220        1.076438        1.905822
 C        -1.576967        0.276317        1.408882
 C        -0.255079        0.775672        1.250987
 C         0.012476        2.146731        1.571563
 C         1.332071        2.639972        1.395753
 C         0.801120       -0.034745        0.773353
 C         2.084740        0.461919        0.613907
 C         2.340158        1.826129        0.934610
 H        -7.356271       -0.801264       -3.454147
 O        -5.451303       -0.156456       -2.047217
 C        -6.546709       -0.444059       -1.593065
 O        -9.724007       -1.462778       -3.005810
 N        -7.590114       -0.803622       -2.467125
 C        -8.226337       -0.818423        0.125603
 C        -8.890772       -1.163994       -2.169707
 N        -9.175793       -1.155962       -0.796574
 C        -6.938352       -0.464201       -0.174848
 H         9.403436        1.538139       -0.265571
 O         7.000107        1.643108        0.211840
 C         7.540590        0.657974       -0.266928
 O        10.931706       -0.194994       -1.214554
 N         8.926108        0.677700       -0.511626
 C         7.690239       -1.595196       -1.165761
 C         9.731895       -0.312937       -1.041645
 N         9.037602       -1.485196       -1.368731
 C         6.897207       -0.611151       -0.640290
 C       -10.539903       -1.530966       -0.414004
 H       -11.258539       -0.853328       -0.880519
 H       -10.756886       -2.548201       -0.748396
 H       -10.626113       -1.472154        0.671718
56
 generated by VMD
 H        10.511412       -2.913450       -1.224081
 H         9.034096       -3.575904       -1.979510
 H        10.106266       -2.477491       -2.893676
 H         5.049644       -1.920117       -0.539612
 H        -6.513734        0.093183        1.815941
 H        -4.109000        0.248298       -0.068133
 H         4.743126        1.005438        0.256861
 H        -0.794520        4.078800        2.502628
 H        -3.104043        3.212679        2.729518
 H        -1.865345       -0.605832        1.192587
 H         1.604419        3.678993        1.928375
 H         0.535655       -1.002720        0.647164
 H         3.366661        2.196391        1.114801
 H        -8.586746       -0.589627        0.972078
 H         7.150635       -2.664337       -1.263614
 C         9.687502       -2.705218       -1.910736
 C         5.393546       -0.899215       -0.372564
 O         2.761193       -1.661563        0.072694
 C        -6.009179       -0.008447        0.854377
 C        -4.696604        0.309745        0.838985
 O        -4.672675        0.914755        3.142206
 C         4.469263       -0.023830        0.076010
 C        -4.054770        0.782799        2.088519
 C         3.081112       -0.494670        0.303130
 C        -1.025319        3.050536        2.234676
 C        -2.309916        2.575984        2.351060
 C        -2.623246        1.235166        1.993095
 C        -1.632276        0.411303        1.499619
 C        -0.287931        0.861684        1.398104
```

```
 C         0.021599        2.211932        1.766501
 C         1.364197        2.656765        1.646289
 C         0.750596        0.023565        0.929564
 C         2.055147        0.476103        0.814740
 C         2.352725        1.818955        1.185056
 H        -6.985822       -1.250469       -3.468797
 O        -5.181315       -0.478179       -1.995637
 C        -6.324751       -0.657994       -1.609132
 O        -9.412950       -1.755040       -3.154588
 N        -7.304322       -1.102660       -2.517173
 C        -8.163201       -0.721198       -0.018572
 C        -8.642090       -1.369171       -2.294556
 N        -9.044493       -1.153537       -0.968718
 C        -6.838446       -0.459915       -0.244680
 H         9.263429        1.571799       -0.714366
 O         6.894168        1.692677       -0.092672
 C         7.426818        0.666447       -0.486259
 O        10.769809       -0.236656       -1.551189
 N         8.792809        0.679974       -0.823586
 C         7.574312       -1.675731       -1.118291
 C         9.586984       -0.353196       -1.285445
 N         8.902200       -1.567755       -1.424159
 C         6.792610       -0.649361       -0.661283
 C       -10.453109       -1.420703       -0.665433
 H       -11.094987       -0.792489       -1.287195
 H       -10.689206       -2.467259       -0.872025
 H       -10.630159       -1.202669        0.388542
56
 generated by VMD
 H        10.390401       -3.099792       -1.096316
 H         8.899596       -3.747789       -1.837393
 H        10.004927       -2.709481       -2.781852
 H         4.955503       -1.939636       -0.471566
 H        -6.548151        0.417755        1.712130
 H        -4.059952        0.214801       -0.052852
 H         4.723488        1.017548        0.227645
 H        -0.726048        4.298095        2.392630
 H        -3.068841        3.509984        2.599836
 H        -1.935135       -0.367651        1.135399
 H         1.665308        3.809306        1.856074
 H         0.455826       -0.852753        0.628535
 H         3.386980        2.255547        1.089087
 H        -8.596328       -0.363439        0.888391
 H         7.038100       -2.762999       -1.160824
 C         9.575634       -2.892869       -1.794200
 C         5.325962       -0.923371       -0.336070
 O         2.670105       -1.598808        0.112663
 C        -6.001265        0.144657        0.808966
 C        -4.683262        0.438605        0.803171
 O        -4.756886        1.439711        2.961585
 C         4.423535       -0.009716        0.079512
 C        -4.089251        1.113191        1.982067
 C         3.022233       -0.435798        0.313647
 C        -0.988506        3.274702        2.135432
 C        -2.289281        2.845469        2.240258
 C        -2.644467        1.511705        1.892893
 C        -1.673875        0.646875        1.427327
 C        -0.314040        1.050128        1.339023
 C         0.035897        2.394533        1.692843
 C         1.394079        2.791865        1.585109
 C         0.701679        0.169982        0.898257
 C         2.022226        0.576637        0.795002
```

```
 C         2.359686        1.914139        1.149565
 H        -6.792602       -1.830155       -3.272778
 O        -5.044603       -0.839306       -1.864014
 C        -6.210612       -0.927220       -1.514293
 O        -9.245249       -2.227442       -3.007184
 N        -7.154448       -1.509992       -2.380920
 C        -8.127221       -0.675648       -0.039272
 C        -8.508540       -1.709087       -2.187747
 N        -8.970929       -1.255053       -0.944469
 C        -6.786645       -0.480528       -0.236280
 H         9.261307        1.431548       -0.743289
 O         6.894690        1.635108       -0.132915
 C         7.400763        0.583019       -0.491793
 O        10.721486       -0.442944       -1.514096
 N         8.767360        0.549457       -0.825160
 C         7.487640       -1.781809       -1.046324
 C         9.535188       -0.519204       -1.249324
 N         8.818832       -1.719262       -1.350238
 C         6.732193       -0.720441       -0.626500
 C       -10.400108       -1.436932       -0.676568
 H       -10.991771       -0.910327       -1.428935
 H       -10.656716       -2.498137       -0.715871
 H       -10.622484       -1.038828        0.314300
56
 generated by VMD
 H        10.296026       -3.273384       -0.889184
 H         8.798899       -3.917633       -1.620831
 H         9.939881       -2.947151       -2.594663
 H         4.880121       -1.956515       -0.401321
 H        -6.572243        0.705391        1.562857
 H        -4.035999        0.210902       -0.067866
 H         4.714215        1.029347        0.186705
 H        -0.659845        4.503889        2.235273
 H        -3.033696        3.793423        2.395719
 H        -1.990195       -0.137496        1.010675
 H         1.725120        3.931404        1.752638
 H         0.388647       -0.704174        0.556248
 H         3.410552        2.311474        1.043175
 H        -8.598448       -0.198253        0.810416
 H         6.951803       -2.859340       -1.019469
 C         9.494727       -3.077773       -1.605611
 C         5.275024       -0.945838       -0.296608
 O         2.592418       -1.533440        0.118338
 C        -5.998651        0.287279        0.734999
 C        -4.679260        0.573504        0.722953
 O        -4.817677        1.923142        2.676742
 C         4.390401        0.005106        0.070880
 C        -4.116438        1.430154        1.794160
 C         2.974831       -0.376220        0.295834
 C        -0.950897        3.487422        1.981306
 C        -2.266600        3.102476        2.060877
 C        -2.659953        1.778154        1.715833
 C        -1.706329        0.874628        1.287950
 C        -0.332762        1.232904        1.223968
 C         0.053595        2.569305        1.569875
 C         1.425957        2.920274        1.487548
 C         0.662462        0.313603        0.817738
 C         1.997738        0.675297        0.738787
 C         2.371141        2.005473        1.083852
 H        -6.669982       -2.330173       -2.991471
 O        -4.961944       -1.125464       -1.706112
 C        -6.139632       -1.152218       -1.386105
```

```
 O        -9.133073       -2.672714       -2.748151
 N        -7.058536       -1.867654       -2.176654
 C        -8.101434       -0.659027       -0.037339
 C        -8.419846       -2.029411       -1.999474
 N        -8.919198       -1.377959       -0.863072
 C        -6.753965       -0.500557       -0.218852
 H         9.274719        1.292138       -0.733987
 O         6.906031        1.577919       -0.165613
 C         7.390074        0.500502       -0.476537
 O        10.697656       -0.647327       -1.407582
 N         8.759728        0.420009       -0.788471
 C         7.424560       -1.885888       -0.936711
 C         9.506206       -0.683568       -1.157930
 N         8.761292       -1.868761       -1.221461
 C         6.690351       -0.789993       -0.571248
 C       -10.357996       -1.508991       -0.618157
 H       -10.920822       -1.095219       -1.458249
 H       -10.624442       -2.562792       -0.509437
 H       -10.606647       -0.967998        0.295790
56
 generated by VMD
 H        10.149799       -3.546568       -0.584114
 H         8.679520       -4.088768       -1.442006
 H         9.941606       -3.161438       -2.301713
 H         4.795664       -1.950843       -0.426370
 H        -6.591646        1.037261        1.327094
 H        -4.032973        0.189536       -0.107451
 H         4.747154        1.027736        0.215610
 H        -0.610822        4.802728        1.721735
 H        -3.005436        4.158574        1.882017
 H        -2.026703        0.115421        0.808668
 H         1.766366        4.145804        1.306583
 H         0.336614       -0.531235        0.420884
 H         3.421781        2.442812        0.739415
 H        -8.606953       -0.012486        0.753010
 H         6.855743       -2.949624       -0.927189
 C         9.416600       -3.289736       -1.352418
 C         5.233921       -0.964213       -0.275182
 O         2.510059       -1.426911       -0.000946
 C        -6.003984        0.445884        0.624333
 C        -4.683338        0.720864        0.574089
 O        -4.855684        2.473644        2.168992
 C         4.377268        0.025709        0.053653
 C        -4.135275        1.793569        1.438416
 C         2.932740       -0.284631        0.182126
 C        -0.922562        3.778677        1.530423
 C        -2.247784        3.430594        1.611191
 C        -2.670045        2.095787        1.346949
 C        -1.729193        1.140944        1.009539
 C        -0.346990        1.463697        0.942162
 C         0.066155        2.811843        1.199253
 C         1.447098        3.124686        1.112887
 C         0.631371        0.495226        0.617704
 C         1.976607        0.818518        0.537587
 C         2.376087        2.161357        0.792329
 H        -6.613137       -2.928886       -2.446456
 O        -4.925334       -1.472172       -1.422384
 C        -6.109458       -1.427176       -1.128553
 O        -9.082218       -3.208331       -2.185410
 N        -7.015662       -2.298523       -1.761308
 C        -8.095221       -0.648855        0.038051
 C        -8.381140       -2.416667       -1.581308
```

```
 N        -8.899853       -1.531586       -0.626215
 C        -6.743763       -0.534367       -0.146168
 H         9.362141        1.070087       -0.406338
 O         6.978301        1.469238        0.010348
 C         7.426809        0.373282       -0.289081
 O        10.726433       -0.929210       -1.024393
 N         8.808275        0.227012       -0.512427
 C         7.370554       -2.003441       -0.794316
 C         9.520673       -0.907488       -0.854617
 N         8.722165       -2.050684       -0.992059
 C         6.670355       -0.877484       -0.454540
 C       -10.343966       -1.604705       -0.388755
 H       -10.887495       -1.383484       -1.310390
 H       -10.618310       -2.609250       -0.058727
 H       -10.607999       -0.877382        0.380054
56
 generated by VMD
 H        10.111968       -3.628240       -0.272097
 H         8.633402       -4.227057       -1.076419
 H         9.896517       -3.379958       -2.013877
 H         4.762278       -1.980761       -0.265545
 H        -6.614395        1.276445        1.027942
 H        -4.023384        0.171107       -0.144528
 H         4.739316        1.038252        0.149670
 H        -0.569166        4.953079        1.453708
 H        -2.984910        4.370742        1.507397
 H        -2.052428        0.277038        0.595891
 H         1.808710        4.228594        1.165044
 H         0.301315       -0.432742        0.325209
 H         3.444285        2.470891        0.716347
 H        -8.616521        0.135373        0.597829
 H         6.815819       -3.033589       -0.675833
 C         9.376118       -3.428531       -1.054547
 C         5.209593       -0.989981       -0.184070
 O         2.474694       -1.399418        0.075279
 C        -6.008958        0.566138        0.464392
 C        -4.685150        0.823291        0.408209
 O        -4.897739        2.836276        1.646237
 C         4.360792        0.030137        0.062559
 C        -4.154224        2.030489        1.085542
 C         2.911664       -0.254492        0.198022
 C        -0.897209        3.932897        1.269045
 C        -2.232670        3.619109        1.292956
 C        -2.678326        2.288933        1.039701
 C        -1.745166        1.302608        0.777271
 C        -0.353963        1.591698        0.762961
 C         0.081590        2.935098        1.006666
 C         1.472404        3.211639        0.978563
 C         0.612917        0.591276        0.508413
 C         1.968110        0.879868        0.481073
 C         2.389908        2.218199        0.722675
 H        -6.542746       -3.311509       -1.958505
 O        -4.879625       -1.690302       -1.168697
 C        -6.071608       -1.594344       -0.921830
 O        -9.018863       -3.544039       -1.720634
 N        -6.962436       -2.567455       -1.411643
 C        -8.086889       -0.619847        0.025841
 C        -8.332877       -2.653599       -1.251810
 N        -8.875547       -1.609913       -0.489990
 C        -6.730289       -0.538839       -0.136712
 H         9.359680        0.988692       -0.426909
 O         6.976998        1.440202       -0.059580
```

```
 C         7.416380        0.321175       -0.275593
 O        10.707659       -1.063328       -0.888069
 N         8.797934        0.145585       -0.476089
 C         7.339347       -2.085342       -0.607524
 C         9.501050       -1.017641       -0.729507
 N         8.691933       -2.160099       -0.789340
 C         6.648238       -0.930799       -0.356728
 C       -10.325789       -1.642301       -0.283092
 H       -10.842222       -1.606914       -1.245176
 H       -10.610725       -2.564856        0.228027
 H       -10.610899       -0.780709        0.321997
56
 generated by VMD
 H        10.099829       -3.666275        0.151464
 H         8.622169       -4.330136       -0.601951
 H         9.893999       -3.578910       -1.606933
 H         4.752614       -1.997328       -0.076616
 H        -6.632129        1.491816        0.604904
 H        -4.018954        0.106771       -0.136525
 H         4.744330        1.046676        0.074080
 H        -0.548362        5.086164        1.002353
 H        -2.972638        4.538639        0.962541
 H        -2.055765        0.386886        0.341528
 H         1.830219        4.316001        0.867807
 H         0.289671       -0.364304        0.212715
 H         3.457166        2.514020        0.590948
 H        -8.637403        0.305312        0.352624
 H         6.805151       -3.094577       -0.354626
 C         9.368790       -3.536392       -0.649818
 C         5.206132       -1.006020       -0.071745
 O         2.461350       -1.373664        0.147599
 C        -6.017697        0.663033        0.252694
 C        -4.686525        0.881737        0.212203
 O        -4.916661        3.117027        0.963160
 C         4.360548        0.036744        0.068171
 C        -4.160493        2.202593        0.632311
 C         2.906550       -0.226374        0.194814
 C        -0.882637        4.061610        0.857138
 C        -2.221993        3.766711        0.832176
 C        -2.677947        2.429815        0.635752
 C        -1.747330        1.418493        0.477612
 C        -0.352720        1.689964        0.507343
 C         0.092279        3.038697        0.696887
 C         1.486916        3.295050        0.719816
 C         0.608891        0.664237        0.352702
 C         1.968230        0.933430        0.371443
 C         2.399501        2.277148        0.559857
 H        -6.530095       -3.661261       -1.241229
 O        -4.867966       -1.920425       -0.767598
 C        -6.067253       -1.756877       -0.605554
 O        -9.018591       -3.805595       -1.072264
 N        -6.956911       -2.806376       -0.900908
 C        -8.099965       -0.567258       -0.004881
 C        -8.334606       -2.838129       -0.790295
 N        -8.887073       -1.640165       -0.317779
 C        -6.736008       -0.541755       -0.115431
 H         9.375717        0.918275       -0.400667
 O         6.990090        1.414429       -0.119015
 C         7.424686        0.278357       -0.230415
 O        10.715485       -1.174002       -0.661019
 N         8.808337        0.077599       -0.388109
 C         7.334786       -2.147266       -0.357125
```

```
 C         9.506704       -1.107402       -0.528666
 N         8.689941       -2.245795       -0.505100
 C         6.648280       -0.971091       -0.219667
 C       -10.345126       -1.607738       -0.176339
 H       -10.818666       -1.809440       -1.139834
 H       -10.670581       -2.370623        0.534960
 H       -10.638651       -0.619758        0.180705
56
 generated by VMD
 H        10.078094       -3.632574        0.757989
 H         8.619473       -4.377295        0.043886
 H         9.924972       -3.752690       -1.003781
 H         4.750357       -1.988921        0.173791
 H        -6.639862        1.606674       -0.002257
 H        -4.021380        0.047128        0.035001
 H         4.754152        1.051394       -0.040164
 H        -0.546145        5.178929        0.252440
 H        -2.969912        4.632639        0.182218
 H        -2.050222        0.435771        0.087874
 H         1.832723        4.395448        0.288835
 H         0.290072       -0.326995        0.125675
 H         3.458662        2.571073        0.271950
 H        -8.651558        0.407133       -0.097806
 H         6.803217       -3.116264        0.090908
 C         9.371827       -3.596488       -0.074615
 C         5.209200       -1.005084        0.073978
 O         2.457473       -1.339983        0.246020
 C        -6.024120        0.707083       -0.007361
 C        -4.688649        0.896853        0.037318
 O        -4.917444        3.252590        0.095031
 C         4.365507        0.048510        0.062444
 C        -4.160075        2.280899        0.087913
 C         2.906948       -0.195672        0.174108
 C        -0.880141        4.144642        0.216413
 C        -2.219118        3.850084        0.177890
 C        -2.676217        2.499766        0.129954
 C        -1.745052        1.476403        0.123361
 C        -0.350760        1.748456        0.162459
 C         0.094724        3.109268        0.209474
 C         1.489164        3.364732        0.247953
 C         0.610055        0.710401        0.156138
 C         1.969494        0.978093        0.186865
 C         2.401256        2.333906        0.234368
 H        -6.540838       -3.865622       -0.143799
 O        -4.872833       -2.069984       -0.044032
 C        -6.075413       -1.859741       -0.076084
 O        -9.036722       -3.941381       -0.226799
 N        -6.968258       -2.945848       -0.132608
 C        -8.112989       -0.535151       -0.105378
 C        -8.350012       -2.936744       -0.177195
 N        -8.903070       -1.649406       -0.159804
 C        -6.745131       -0.550023       -0.062442
 H         9.398046        0.856095       -0.357049
 O         7.008125        1.388149       -0.207640
 C         7.439085        0.245895       -0.169632
 O        10.732187       -1.255606       -0.327216
 N         8.825534        0.024371       -0.260768
 C         7.338420       -2.177341       -0.008767
 C         9.520540       -1.170693       -0.238873
 N         8.696675       -2.296169       -0.103682
 C         6.655255       -0.991435       -0.033193
 C       -10.365466       -1.569001       -0.205643
```

```
 H       -10.738201       -2.031725       -1.122516
 H       -10.796278       -2.097685        0.647751
 H       -10.657677       -0.518532       -0.176098
```

8.8. *Ab initio* study on TACN and TACNA

8.8.1. 6-31G* optimized structure of Zn^{2+} in vacuum

```
HETATM    1 Zn            0       0.000   0.000   0.000                          Zn
END
```

8.8.2. 6-31G* optimized structure of TACN in vacuum

```
HETATM    1  C            0       0.000   2.105  -0.040                           C
HETATM    2  N            0       1.172   1.348  -0.472                           N
HETATM    3  C            0       1.755   0.402   0.483                           C
HETATM    4  C            0       1.823  -1.052  -0.040                           C
HETATM    5  H            0       1.193   0.426   1.425                           H
HETATM    6  H            0       2.782   0.708   0.737                           H
HETATM    7  N            0       0.582  -1.689  -0.472                           N
HETATM    8  H            0       2.281  -1.671   0.743                           H
HETATM    9  H            0       2.512  -1.085  -0.895                           H
HETATM   10  C            0      -0.529  -1.721   0.483                           C
HETATM   11  C            0      -1.823  -1.052  -0.040                           C
HETATM   12  H            0      -0.227  -1.246   1.425                           H
HETATM   13  H            0      -0.778  -2.763   0.737                           H
HETATM   14  N            0      -1.754   0.341  -0.472                           N
HETATM   15  H            0      -2.588  -1.140   0.743                           H
HETATM   16  H            0      -2.196  -1.633  -0.895                           H
HETATM   17  C            0      -1.225   1.319   0.483                           C
HETATM   18  H            0      -2.004   2.055   0.737                           H
HETATM   19  H            0      -0.965   0.820   1.425                           H
HETATM   20  H            0      -0.316   2.718  -0.895                           H
HETATM   21  H            0       0.307   2.811   0.743                           H
HETATM   22  H            0      -1.264   0.410  -1.360                           H
HETATM   23  H            0       0.277  -1.300  -1.360                           H
HETATM   24  H            0       0.988   0.890  -1.360                           H
END
```

8.8.3. 6-31G* optimized structure of $[Zn(TACNA)]^{2+}$ complex in vacuum

```
HETATM    1  C            0       1.413   1.649  -0.519                           C
HETATM    2  N            0       0.000   1.700   0.048                           N
HETATM    3  C            0      -1.107   1.460  -0.972                           C
HETATM    4  C            0      -2.134   0.399  -0.519                           C
HETATM    5  H            0      -0.637   1.172  -1.915                           H
HETATM    6  H            0      -1.639   2.398  -1.159                           H
HETATM    7  N            0      -1.472  -0.850   0.048                           N
HETATM    8  H            0      -2.770   0.135  -1.371                           H
```

```
HETATM    9  H           0       -2.784    0.796    0.268                       H
HETATM   10  C           0       -0.711   -1.689   -0.972                       C
HETATM   11  C           0        0.722   -2.048   -0.519                       C
HETATM   12  H           0       -0.696   -1.138   -1.915                       H
HETATM   13  H           0       -1.257   -2.618   -1.159                       H
HETATM   14  N           0        1.472   -0.850    0.048                       N
HETATM   15  H           0        1.268   -2.467   -1.371                       H
HETATM   16  H           0        0.702   -2.809    0.268                       H
HETATM   17  C           0        1.818    0.229   -0.972                       C
HETATM   18  H           0        2.896    0.220   -1.159                       H
HETATM   19  H           0        1.334   -0.034   -1.915                       H
HETATM   20  H           0        2.081    2.013    0.268                       H
HETATM   21  H           0        1.502    2.332   -1.371                       H
HETATM   22  H           0        2.323   -1.183    0.507                       H
HETATM   23  H           0       -2.186   -1.421    0.507                       H
HETATM   24  H           0       -0.137    2.603    0.507                       H
HETATM   25 Zn           0        0.000    0.000    1.228                      Zn
END
```

8.8.4. 6-31G* optimized structure of TACNA in vacuum

```
HETATM    1  C           0       -2.097    2.057   -0.290                       C
HETATM    2  N           0       -0.938    1.705    0.548                       N
HETATM    3  C           0       -1.238    0.939    1.772                       C
HETATM    4  C           0       -0.396   -0.351    1.883                       C
HETATM    5  H           0       -2.304    0.687    1.794                       H
HETATM    6  H           0       -1.057    1.541    2.674                       H
HETATM    7  N           0       -0.404   -1.139    0.659                       N
HETATM    8  H           0       -0.741   -0.921    2.768                       H
HETATM    9  H           0        0.643   -0.068    2.081                       H
HETATM   10  C           0       -1.665   -1.781    0.294                       C
HETATM   11  C           0       -2.051   -1.489   -1.180                       C
HETATM   12  H           0       -2.449   -1.403    0.959                       H
HETATM   13  H           0       -1.636   -2.872    0.457                       H
HETATM   14  N           0       -1.960   -0.088   -1.584                       N
HETATM   15  H           0       -3.067   -1.860   -1.366                       H
HETATM   16  H           0       -1.388   -2.058   -1.844                       H
HETATM   17  C           0       -2.837    0.839   -0.864                       C
HETATM   18  H           0       -3.617    1.227   -1.537                       H
HETATM   19  H           0       -3.379    0.331   -0.053                       H
HETATM   20  H           0       -1.710    2.661   -1.120                       H
HETATM   21  H           0       -2.838    2.680    0.243                       H
HETATM   22  H           0       -0.998    0.197   -1.423                       H
HETATM   23  H           0       -0.448    2.561    0.791                       H
HETATM   24  C           0        0.806   -1.897    0.399                       C
HETATM   25  C           0        1.989   -1.073   -0.144                       C
HETATM   26  H           0        0.594   -2.653   -0.365                       H
HETATM   27  H           0        1.181   -2.449    1.277                       H
HETATM   28  O           0        3.098   -1.588   -0.257                       O
HETATM   29  N           0        1.699    0.204   -0.513                       N
HETATM   30  C           0        2.729    1.097   -1.026                       C
HETATM   31  H           0        0.777    0.578   -0.283                       H
HETATM   32  C           0        3.463    1.873    0.072                       C
HETATM   33  H           0        2.255    1.788   -1.733                       H
HETATM   34  H           0        3.441    0.482   -1.583                       H
HETATM   35  H           0        2.768    2.489    0.656                       H
HETATM   36  H           0        4.221    2.536   -0.362                       H
HETATM   37  H           0        3.964    1.178    0.754                       H
END
```

8.8.5. 6-31G* optimized structure of $[Zn(TACNA)]^{2+}$ complex in vacuum

```
HETATM    1  C           0        3.139  -1.442   0.045                       C
HETATM    2  N           0        1.828  -1.481   0.777                       N
HETATM    3  C           0        1.762  -0.601   2.000                       C
HETATM    4  C           0        0.502   0.288   2.037                       C
HETATM    5  H           0        2.669   0.007   2.035                       H
HETATM    6  H           0        1.773  -1.221   2.901                       H
HETATM    7  N           0        0.250   0.989   0.740                       N
HETATM    8  H           0        0.594   1.003   2.866                       H
HETATM    9  H           0       -0.376  -0.335   2.241                       H
HETATM   10  C           0        1.283   2.008   0.378                       C
HETATM   11  C           0        1.795   1.839  -1.067                       C
HETATM   12  H           0        2.110   1.921   1.086                       H
HETATM   13  H           0        0.894   3.025   0.497                       H
HETATM   14  N           0        2.157   0.413  -1.387                       N
HETATM   15  H           0        2.657   2.499  -1.220                       H
HETATM   16  H           0        1.025   2.135  -1.787                       H
HETATM   17  C           0        3.376  -0.093  -0.659                       C
HETATM   18  H           0        4.205  -0.210  -1.362                       H
HETATM   19  H           0        3.687   0.668   0.061                       H
HETATM   20  H           0        3.113  -2.256  -0.688                       H
HETATM   21  H           0        3.974  -1.640   0.727                       H
HETATM   22  H           0        2.324   0.357  -2.392                       H
HETATM   23  H           0        1.656  -2.448   1.055                       H
HETATM   24  C           0       -1.151   1.438   0.584                       C
HETATM   25  C           0       -2.020   0.352  -0.087                       C
HETATM   26  H           0       -1.170   2.312  -0.077                       H
HETATM   27  H           0       -1.587   1.749   1.541                       H
HETATM   28  O           0       -1.472  -0.576  -0.769                       O
HETATM   29  N           0       -3.326   0.449   0.024                       N
HETATM   30  C           0       -4.319  -0.443  -0.632                       C
HETATM   31  H           0       -3.704   1.216   0.573                       H
HETATM   32  C           0       -5.553  -0.614   0.247                       C
HETATM   33  H           0       -3.818  -1.393  -0.822                       H
HETATM   34  H           0       -4.578   0.003  -1.599                       H
HETATM   35  H           0       -5.306  -1.083   1.205                       H
HETATM   36  H           0       -6.271  -1.258  -0.269                       H
HETATM   37  H           0       -6.053   0.342   0.437                       H
HETATM   38 Zn           0        0.490  -0.659  -0.634                      Zn
END
```

8.8.6. 6-31G* optimized structure of $[Zn(H_2O)_6]^{2+}$ complex in vacuum

```
HETATM    1 Zn           0        0.000   0.000   0.000                      Zn
HETATM    2  O           0        0.000   1.740   1.231                       O
HETATM    3  H           0        0.554   2.530   1.112                       H
HETATM    4  H           0       -0.552   1.893   2.017                       H
HETATM    5  O           0        1.507  -0.870   1.231                       O
HETATM    6  H           0        1.916  -0.468   2.017                       H
HETATM    7  H           0        1.915  -1.745   1.112                       H
HETATM    8  O           0       -1.507  -0.870   1.231                       O
HETATM    9  H           0       -1.364  -1.425   2.017                       H
HETATM   10  H           0       -2.468  -0.786   1.112                       H
HETATM   11  O           0        1.507   0.870  -1.231                       O
```

```
HETATM   12  H           0        2.468    0.786   -1.112                          H
HETATM   13  H           0        1.364    1.425   -2.017                          H
HETATM   14  O           0       -1.507    0.870   -1.231                          O
HETATM   15  H           0       -1.916    0.468   -2.017                          H
HETATM   16  H           0       -1.915    1.745   -1.112                          H
HETATM   17  O           0        0.000   -1.740   -1.231                          O
HETATM   18  H           0        0.552   -1.893   -2.017                          H
HETATM   19  H           0       -0.554   -2.530   -1.112                          H
END
```

8.8.7. 6-31G* optimized structure of water in vacuum

```
HETATM    1  O           0        0.000    0.000    0.120                          O
HETATM    2  H           0        0.000    0.761   -0.479                          H
HETATM    3  H           0        0.000   -0.761   -0.479                          H
END
```

8.8.8. 6-31G* optimized structure of $[Zn(TACN)(H_2O)]^{2+}$ complex in vacuum

```
HETATM    1  C           0       -0.823    0.085    2.156                          C
HETATM    2  N           0       -0.299   -1.044    1.312                          N
HETATM    3  C           0       -1.343   -1.732    0.469                          C
HETATM    4  C           0       -0.918   -1.878   -1.003                          C
HETATM    5  H           0       -2.275   -1.168    0.549                          H
HETATM    6  H           0       -1.549   -2.728    0.872                          H
HETATM    7  N           0       -0.330   -0.610   -1.555                          N
HETATM    8  H           0       -1.784   -2.196   -1.595                          H
HETATM    9  H           0       -0.150   -2.653   -1.105                          H
HETATM   10  C           0       -1.321    0.512   -1.733                          C
HETATM   11  C           0       -0.835    1.841   -1.126                          C
HETATM   12  H           0       -2.270    0.206   -1.285                          H
HETATM   13  H           0       -1.518    0.668   -2.797                          H
HETATM   14  N           0       -0.252    1.663    0.250                          N
HETATM   15  H           0       -1.669    2.551   -1.106                          H
HETATM   16  H           0       -0.046    2.278   -1.748                          H
HETATM   17  C           0       -1.260    1.295    1.311                          C
HETATM   18  H           0       -1.420    2.145    1.980                          H
HETATM   19  H           0       -2.217    1.102    0.822                          H
HETATM   20  H           0       -0.015    0.367    2.840                          H
HETATM   21  H           0       -1.670   -0.247    2.768                          H
HETATM   22  H           0        0.191    2.544    0.512                          H
HETATM   23  H           0        0.086   -0.829   -2.461                          H
HETATM   24  H           0        0.122   -1.729    1.941                          H
HETATM   25 Zn           0        0.994   -0.024   -0.013                         Zn
HETATM   26  O           0        3.037   -0.039   -0.027                          O
HETATM   27  H           0        3.623    0.739    0.021                          H
HETATM   28  H           0        3.613   -0.825   -0.086                          H
END
```

8.8.9. 6-31G* optimized structure of $[Zn(TACNA)(H_2O)]^{2+}$ complex in vacuum

```
HETATM    1  C           0       -3.236    0.759    0.848                          C
HETATM    2  N           0       -1.931    0.407    1.491                          N
HETATM    3  C           0       -1.807   -1.032    1.900                          C
```

```
HETATM    4  C           0       -0.494  -1.673   1.418                          C
HETATM    5  H           0       -2.671  -1.580   1.516                          H
HETATM    6  H           0       -1.852  -1.117   2.990                          H
HETATM    7  N           0       -0.208  -1.397  -0.018                          N
HETATM    8  H           0       -0.527  -2.753   1.618                          H
HETATM    9  H           0        0.341  -1.262   1.997                          H
HETATM   10  C           0       -1.172  -2.028  -0.965                          C
HETATM   11  C           0       -1.695  -1.028  -2.010                          C
HETATM   12  H           0       -2.001  -2.444  -0.389                          H
HETATM   13  H           0       -0.717  -2.878  -1.485                          H
HETATM   14  N           0       -2.147   0.256  -1.385                          N
HETATM   15  H           0       -2.508  -1.496  -2.579                          H
HETATM   16  H           0       -0.904  -0.770  -2.724                          H
HETATM   17  C           0       -3.389   0.128  -0.548                          C
HETATM   18  H           0       -4.233   0.605  -1.054                          H
HETATM   19  H           0       -3.643  -0.931  -0.463                          H
HETATM   20  H           0       -3.267   1.852   0.777                          H
HETATM   21  H           0       -4.083   0.451   1.473                          H
HETATM   22  H           0       -2.329   0.914  -2.142                          H
HETATM   23  H           0       -1.827   0.994   2.319                          H
HETATM   24  C           0        1.207  -1.635  -0.358                          C
HETATM   25  C           0        2.060  -0.403  -0.018                          C
HETATM   26  H           0        1.288  -1.788  -1.441                          H
HETATM   27  H           0        1.601  -2.538   0.126                          H
HETATM   28  O           0        1.514   0.739   0.022                          O
HETATM   29  N           0        3.354  -0.580   0.168                          N
HETATM   30  C           0        4.350   0.492   0.404                          C
HETATM   31  H           0        3.713  -1.530   0.140                          H
HETATM   32  C           0        5.677   0.158  -0.272                          C
HETATM   33  H           0        4.475   0.604   1.486                          H
HETATM   34  H           0        3.921   1.417   0.016                          H
HETATM   35  H           0        6.105  -0.775   0.115                          H
HETATM   36  H           0        6.397   0.955  -0.064                          H
HETATM   37  H           0        5.568   0.074  -1.357                          H
HETATM   38 Zn           0       -0.520   0.825  -0.076                         Zn
HETATM   39  O           0       -0.221   2.902  -0.283                          O
HETATM   40  H           0        0.701   3.209  -0.237                          H
HETATM   41  H           0       -0.781   3.677  -0.452                          H
END
```

8.8.10. 6-31G* optimized structure of Zn^{2+} in PCM water

```
HETATM    1 Zn           0        0.000   0.000   0.000                         Zn
END
```

8.8.11. 6-31G* optimized structure of TACN in PCM water

```
HETATM    1  C           0        0.000   2.106  -0.040                          C
HETATM    2  N           0        1.178   1.353  -0.474                          N
HETATM    3  C           0        1.751   0.399   0.484                          C
HETATM    4  C           0        1.824  -1.053  -0.040                          C
HETATM    5  H           0        1.185   0.423   1.423                          H
HETATM    6  H           0        2.778   0.698   0.747                          H
HETATM    7  N           0        0.583  -1.697  -0.474                          N
HETATM    8  H           0        2.285  -1.667   0.744                          H
HETATM    9  H           0        2.509  -1.081  -0.898                          H
```

```
HETATM   10  C           0      -0.530  -1.716   0.484                       C
HETATM   11  C           0      -1.824  -1.053  -0.040                       C
HETATM   12  H           0      -0.226  -1.238   1.423                       H
HETATM   13  H           0      -0.785  -2.755   0.747                       H
HETATM   14  N           0      -1.761   0.343  -0.474                       N
HETATM   15  H           0      -2.586  -1.146   0.744                       H
HETATM   16  H           0      -2.191  -1.633  -0.898                       H
HETATM   17  C           0      -1.221   1.317   0.484                       C
HETATM   18  H           0      -1.994   2.057   0.747                       H
HETATM   19  H           0      -0.959   0.815   1.423                       H
HETATM   20  H           0      -0.318   2.714  -0.898                       H
HETATM   21  H           0       0.301   2.813   0.744                       H
HETATM   22  H           0      -1.262   0.409  -1.359                       H
HETATM   23  H           0       0.276  -1.297  -1.359                       H
HETATM   24  H           0       0.985   0.888  -1.359                       H
END
```

8.8.12. 6-31G* optimized structure of $[Zn(TACN)]^{2+}$ complex in PCM water

```
HETATM    1  C           0      -0.490  -2.079   0.535                       C
HETATM    2  N           0      -1.333  -0.982  -0.010                       N
HETATM    3  C           0      -1.802   0.019   0.991                       C
HETATM    4  C           0      -1.557   1.467   0.530                       C
HETATM    5  H           0      -1.292  -0.172   1.938                       H
HETATM    6  H           0      -2.871  -0.110   1.181                       H
HETATM    7  N           0      -0.186   1.647  -0.019                       N
HETATM    8  H           0      -1.742   2.144   1.373                       H
HETATM    9  H           0      -2.254   1.727  -0.272                       H
HETATM   10  C           0       0.911   1.554   0.989                       C
HETATM   11  C           0       2.045   0.617   0.537                       C
HETATM   12  H           0       0.488   1.209   1.935                       H
HETATM   13  H           0       1.333   2.545   1.181                       H
HETATM   14  N           0       1.518  -0.662  -0.014                       N
HETATM   15  H           0       2.720   0.439   1.383                       H
HETATM   16  H           0       2.622   1.089  -0.264                       H
HETATM   17  C           0       0.891  -1.573   0.987                       C
HETATM   18  H           0       1.537  -2.438   1.163                       H
HETATM   19  H           0       0.812  -1.045   1.940                       H
HETATM   20  H           0      -0.373  -2.810  -0.271                       H
HETATM   21  H           0      -0.981  -2.579   1.378                       H
HETATM   22  H           0       2.278  -1.147  -0.486                       H
HETATM   23  H           0      -0.140   2.547  -0.491                       H
HETATM   24  H           0      -2.136  -1.394  -0.480                       H
HETATM   25 Zn           0       0.002  -0.013  -1.414                      Zn
END
```

8.8.13. 6-31G* optimized structure of TACNA in PCM water

```
HETATM    1  C           0      -2.016   2.078  -0.314                       C
HETATM    2  N           0      -0.873   1.700   0.536                       N
HETATM    3  C           0      -1.213   0.965   1.770                       C
HETATM    4  C           0      -0.417  -0.351   1.905                       C
HETATM    5  H           0      -2.287   0.749   1.786                       H
HETATM    6  H           0      -1.017   1.576   2.661                       H
HETATM    7  N           0      -0.429  -1.143   0.681                       N
HETATM    8  H           0      -0.799  -0.908   2.781                       H
```

```
HETATM    9  H           0       0.627  -0.102   2.124                       H
HETATM   10  C           0      -1.698  -1.761   0.304                       C
HETATM   11  C           0      -2.059  -1.476  -1.176                       C
HETATM   12  H           0      -2.479  -1.359   0.957                       H
HETATM   13  H           0      -1.694  -2.851   0.474                       H
HETATM   14  N           0      -1.923  -0.079  -1.598                       N
HETATM   15  H           0      -3.085  -1.819  -1.366                       H
HETATM   16  H           0      -1.406  -2.074  -1.825                       H
HETATM   17  C           0      -2.785   0.876  -0.884                       C
HETATM   18  H           0      -3.552   1.278  -1.564                       H
HETATM   19  H           0      -3.342   0.384  -0.075                       H
HETATM   20  H           0      -1.608   2.669  -1.143                       H
HETATM   21  H           0      -2.743   2.721   0.213                       H
HETATM   22  H           0      -0.959   0.179  -1.394                       H
HETATM   23  H           0      -0.367   2.549   0.776                       H
HETATM   24  C           0       0.772  -1.907   0.416                       C
HETATM   25  C           0       1.948  -1.096  -0.157                       C
HETATM   26  H           0       0.549  -2.679  -0.330                       H
HETATM   27  H           0       1.159  -2.443   1.300                       H
HETATM   28  O           0       3.047  -1.640  -0.321                       O
HETATM   29  N           0       1.681   0.188  -0.482                       N
HETATM   30  C           0       2.701   1.087  -1.008                       C
HETATM   31  H           0       0.765   0.577  -0.228                       H
HETATM   32  C           0       3.455   1.850   0.087                       C
HETATM   33  H           0       2.209   1.790  -1.688                       H
HETATM   34  H           0       3.400   0.486  -1.596                       H
HETATM   35  H           0       2.766   2.454   0.689                       H
HETATM   36  H           0       4.198   2.522  -0.358                       H
HETATM   37  H           0       3.975   1.154   0.753                       H
END
```

8.8.14. 6-31G* optimized structure of $[Zn(TACNA)]^{2+}$ complex in PCM water

```
HETATM    1  C           0       3.231  -1.288   0.323                       C
HETATM    2  N           0       1.919  -1.257   1.023                       N
HETATM    3  C           0       1.787  -0.205   2.074                       C
HETATM    4  C           0       0.484   0.604   1.947                       C
HETATM    5  H           0       2.655   0.455   2.019                       H
HETATM    6  H           0       1.810  -0.663   3.067                       H
HETATM    7  N           0       0.228   1.058   0.555                       N
HETATM    8  H           0       0.518   1.453   2.644                       H
HETATM    9  H           0      -0.361  -0.028   2.238                       H
HETATM   10  C           0       1.224   2.039   0.044                       C
HETATM   11  C           0       1.757   1.657  -1.347                       C
HETATM   12  H           0       2.047   2.102   0.758                       H
HETATM   13  H           0       0.794   3.045  -0.006                       H
HETATM   14  N           0       2.201   0.237  -1.398                       N
HETATM   15  H           0       2.576   2.338  -1.611                       H
HETATM   16  H           0       0.972   1.769  -2.100                       H
HETATM   17  C           0       3.421  -0.067  -0.593                       C
HETATM   18  H           0       4.270  -0.257  -1.255                       H
HETATM   19  H           0       3.680   0.815  -0.003                       H
HETATM   20  H           0       3.246  -2.206  -0.272                       H
HETATM   21  H           0       4.065  -1.334   1.035                       H
HETATM   22  H           0       2.368  -0.011  -2.371                       H
HETATM   23  H           0       1.754  -2.172   1.438                       H
HETATM   24  C           0      -1.166   1.471   0.343                       C
HETATM   25  C           0      -2.047   0.297  -0.103                       C
HETATM   26  H           0      -1.200   2.204  -0.470                       H
```

```
HETATM   27  H           0      -1.593   1.955   1.230                       H
HETATM   28  O           0      -1.542  -0.754  -0.579                       O
HETATM   29  N           0      -3.360   0.461  -0.014                       N
HETATM   30  C           0      -4.351  -0.501  -0.520                       C
HETATM   31  H           0      -3.708   1.341   0.350                       H
HETATM   32  C           0      -5.617  -0.484   0.331                       C
HETATM   33  H           0      -3.880  -1.485  -0.508                       H
HETATM   34  H           0      -4.581  -0.253  -1.563                       H
HETATM   35  H           0      -5.401  -0.766   1.367                       H
HETATM   36  H           0      -6.339  -1.197  -0.078                       H
HETATM   37  H           0      -6.087   0.506   0.330                       H
HETATM   38 Zn           0       0.516  -0.874  -0.592                      Zn
END
```

8.8.15. 6-31G* optimized structure of $[Zn(H_2O)_6]^{2+}$ complex in PCM water

```
HETATM    1 Zn           0       0.000   0.000   0.000                      Zn
HETATM    2  O           0       0.671   1.718   1.077                       O
HETATM    3  H           0       1.613   1.901   0.911                       H
HETATM    4  H           0       0.546   1.788   2.039                       H
HETATM    5  O           0       1.152  -1.432   1.085                       O
HETATM    6  H           0       1.274  -1.353   2.046                       H
HETATM    7  H           0       0.846  -2.342   0.924                       H
HETATM    8  O           0      -1.822  -0.295   1.079                       O
HETATM    9  H           0      -1.814  -0.443   2.040                       H
HETATM   10  H           0      -2.462   0.420   0.920                       H
HETATM   11  O           0       1.822   0.295  -1.079                       O
HETATM   12  H           0       2.462  -0.420  -0.920                       H
HETATM   13  H           0       1.814   0.443  -2.040                       H
HETATM   14  O           0      -1.152   1.432  -1.085                       O
HETATM   15  H           0      -1.274   1.353  -2.046                       H
HETATM   16  H           0      -0.846   2.342  -0.924                       H
HETATM   17  O           0      -0.671  -1.718  -1.077                       O
HETATM   18  H           0      -0.546  -1.788  -2.039                       H
HETATM   19  H           0      -1.613  -1.901  -0.911                       H
END
```

8.8.16. 6-31G* optimized structure of water in PCM water

```
HETATM    1  O           0       0.000   0.000   0.121                       O
HETATM    2  H           0       0.000   0.759  -0.483                       H
HETATM    3  H           0       0.000  -0.759  -0.483                       H
END
```

8.8.17. 6-31G* optimized structure of $[Zn(TACN)(H_20)]^{2+}$ complex in PCM water

```
HETATM    1  C           0       0.028   0.031   0.028                       C
HETATM    2  N           0       0.060   0.063   1.514                       N
HETATM    3  C           0       1.425   0.020   2.113                       C
HETATM    4  C           0       1.632   1.109   3.180                       C
HETATM    5  H           0       2.162   0.128   1.314                       H
HETATM    6  H           0       1.606  -0.956   2.573                       H
HETATM    7  N           0       1.133   2.434   2.724                       N
HETATM    8  H           0       2.695   1.153   3.446                       H
```

```
HETATM    9  H           0       1.069   0.857   4.083                       H
HETATM   10  C           0       1.926   3.065   1.630                       C
HETATM   11  C           0       1.040   3.561   0.475                       C
HETATM   12  H           0       2.656   2.339   1.265                       H
HETATM   13  H           0       2.497   3.913   2.019                       H
HETATM   14  N           0       0.025   2.547   0.082                       N
HETATM   15  H           0       1.680   3.838  -0.371                       H
HETATM   16  H           0       0.487   4.452   0.786                       H
HETATM   17  C           0       0.579   1.331  -0.581                       C
HETATM   18  H           0       0.347   1.343  -1.649                       H
HETATM   19  H           0       1.668   1.355  -0.500                       H
HETATM   20  H           0      -1.019  -0.101  -0.262                       H
HETATM   21  H           0       0.599  -0.818  -0.367                       H
HETATM   22  H           0      -0.653   2.990  -0.534                       H
HETATM   23  H           0       1.103   3.061   3.526                       H
HETATM   24  H           0      -0.491  -0.716   1.867                       H
HETATM   25 Zn           0      -0.872   1.996   2.006                      Zn
HETATM   26  O           0      -2.864   2.247   2.683                       O
HETATM   27  H           0      -3.485   2.437   1.957                       H
HETATM   28  H           0      -3.212   1.452   3.124                       H
END
```

8.8.18. 6-31G* optimized structure of $[Zn(TACNA)(H_20)]^{2+}$ complex in PCM water

```
HETATM    1  C           0       3.212   0.594  -1.110                       C
HETATM    2  N           0       1.891   0.108  -1.591                       N
HETATM    3  C           0       1.772  -1.378  -1.683                       C
HETATM    4  C           0       0.484  -1.908  -1.033                       C
HETATM    5  H           0       2.651  -1.829  -1.218                       H
HETATM    6  H           0       1.783  -1.691  -2.730                       H
HETATM    7  N           0       0.241  -1.327   0.313                       N
HETATM    8  H           0       0.526  -3.005  -0.990                       H
HETATM    9  H           0      -0.371  -1.638  -1.661                       H
HETATM   10  C           0       1.248  -1.735   1.331                       C
HETATM   11  C           0       1.807  -0.534   2.109                       C
HETATM   12  H           0       2.057  -2.265   0.824                       H
HETATM   13  H           0       0.819  -2.449   2.041                       H
HETATM   14  N           0       2.230   0.568   1.202                       N
HETATM   15  H           0       2.638  -0.876   2.739                       H
HETATM   16  H           0       1.038  -0.121   2.768                       H
HETATM   17  C           0       3.434   0.267   0.375                       C
HETATM   18  H           0       4.293   0.840   0.735                       H
HETATM   19  H           0       3.691  -0.788   0.500                       H
HETATM   20  H           0       3.219   1.678  -1.258                       H
HETATM   21  H           0       4.034   0.168  -1.699                       H
HETATM   22  H           0       2.416   1.389   1.774                       H
HETATM   23  H           0       1.714   0.516  -2.506                       H
HETATM   24  C           0      -1.147  -1.532   0.751                       C
HETATM   25  C           0      -2.068  -0.503   0.093                       C
HETATM   26  H           0      -1.205  -1.358   1.830                       H
HETATM   27  H           0      -1.494  -2.556   0.562                       H
HETATM   28  O           0      -1.629   0.643  -0.185                       O
HETATM   29  N           0      -3.326  -0.862  -0.134                       N
HETATM   30  C           0      -4.354   0.016  -0.709                       C
HETATM   31  H           0      -3.605  -1.802   0.123                       H
HETATM   32  C           0      -5.357   0.508   0.335                       C
HETATM   33  H           0      -4.860  -0.551  -1.495                       H
HETATM   34  H           0      -3.832   0.852  -1.177                       H
HETATM   35  H           0      -5.874  -0.329   0.814                       H
```

```
HETATM   36  H           0      -6.109   1.139  -0.151                       H
HETATM   37  H           0      -4.858   1.099   1.110                       H
HETATM   38 Zn           0       0.488   0.890  -0.119                      Zn
HETATM   39  O           0      -0.114   2.910   0.194                       O
HETATM   40  H           0      -1.075   2.890   0.345                       H
HETATM   41  H           0       0.269   3.394   0.945                       H
END
```

9. Acknowledgments

I really gratefully acknowledge Prof. Michele Parrinello for hosting and supervising the computationally activities described in this thesis work and for making available all the advanced computing facilities and expertise. Of course, special thanks go also to Vincenzo Verdolino, who taught me all I know about Molecular Dynamics, to Ferruccio Palazzesi for his help in simulations setup and analysis and to all the unique people of the group in Lugano.

My special thanks go to my supervisor, Prof. Roberto Corradini, for introducing me in exciting world of artificial DNAs engineering and for his advices during all the thesis work. I express my gratitude also to all the kind colleagues of lab 50, in particular to Dr. Alex Manicardi who helped me to increase my laboratory expertise and who more and more over the years became a good friend. Thanks also to Saša Korom for the help in experimental data revision. In this acknowledgements I want to include also my bachelors' students, Luca Pigoli, Tania Peracchia, Laura Ceriani and Ambra Rossi. Thanks also to Prof. Matteo Tegoni for the measures on TACN.

My best thanks go to the treasures of my life, Francesca and little Maria Cristina. Francesca supported me during all the PhD activity, encouraging me and last year she became my beloved wife. Maria Cristina entered in our lives as a wonderful gift and her graceful caresses sustained me during all PhD writing. My grateful tanks go to my family for supporting me during all my studies and in everyday life; a big hug goes in particular to my sister Elisa.

Last but not least, thanks to all my friends that shared with me joys and sorrows during all my life.

Massimiliano Donato Verona


I'VE REACHED MY GOAL!
AND THERE'S THE NEXT ONE!
©www.cartoonmotivators.com