UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN DRUG SCIENCES

CICLO XXXVI

Conformational analysis guided by NMR: an analytical strategy for structure elucidation of pharmaceutical compounds

Coordinatore:

Chiar.mo Prof. Marco Mor

Tutori:

Dr. Luca Venturi

Dott.ssa Valentina Mileo

Dottoranda: Emanuela Bua

Anni Accademici 2020/2021 – 2022/2023

UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

Ph.D. COURSE IN DRUG SCIENCES

XXXVI CYCLE

Conformational analysis guided by NMR: an analytical strategy for structure elucidation of pharmaceutical compounds

Coordinator:

Chiar.mo Prof. Marco Mor

Tutors:

Dr. Luca Venturi

Dott.ssa Valentina Mileo

Ph.D. student: Emanuela Bua

Academic Years 2020/2021 – 2022/2023

"Οὐ γὰρ ὡς ἀγγεῖον ὁ νοῦς ἀποπληρώσεως ἀλλ' ὑπεκκαύματος μόνον ὥσπερ ὕλη δεῖται, ὁρμὴν ἐμποιοῦντος εὑρετικὴν καὶ ὄρεξιν ἐπὶ τὴν ἀλήθειαν."

"La mente non ha bisogno, come un vaso, di essere riempita, ma piuttosto, come legna, necessita di una scintilla che l'accenda e vi infonda l'impulso della ricerca e un amore ardente per la verità."

(Plutarco di Cheronea, Moralia – De audiendo, 48C)

Abstract

Nuclear Magnetic Resonance (NMR) spectroscopy is one of the most powerful techniques for structure elucidation and characterization. Moreover, NMR, coupled with computational chemistry, can be useful for the conformational analysis of free ligands in solution to investigate target-ligand interactions within Structure-Based Drug Design (SBDD) approach. Structure-based drug design, using X-ray crystallography, computational molecular modelling, molecular docking and NMR methods, is a powerful and straightforward approach in drug design process.

Conformational analysis by NMR is based upon the idea of ligand conformational selection, in which the free ligand in solution can be fully pre-organized into the bioactive conformation. Thus, the free ligand can adopt multiple three-dimensional conformations due to rotatable bonds, including the one adopted in the binding pocket of the ligand-target complex. This method includes a detailed analysis of multiple NMR constraints, comprising chemical shifts, through-space distances between pairs of atoms derived from rotating-frame Overhauser spectroscopy (ROESY) or quantitative nuclear Overhauser effect (qNOE), torsion angles derived from J-couplings and dynamic average of orientation of internuclear vectors derived from residual dipolar couplings (RDCs).

In addition, quantum mechanical (QM) calculations to predict some NMR parameters, such as isotropic shielding constants, within a reasonable timeframe can be useful. Chemical shift predictions, which are conformation-dependent, provide important information for the conformational analysis as far as some diagnostic protons are detected. Calculated chemical shifts can also be used to confirm the proposed conformational ensemble derived from NMR-based analysis.

The objective of this project is to set up and validate a protocol with high accuracy and precision, which exploits NMR constraints to obtain structural information from free ligands in solution, combining experimental information with computational chemistry simulations. For this purpose, three active macrocyclic compounds, i.e. Lorlatinib, Pacritinib and Simeprevir were tested. Several approaches such as single or multi-parametric fitting, different solvents and various alignment media were investigated.

Table of Contents

1. Introduction

1.1 Structure Elucidation in Drug Discovery

Structure Elucidation is the process of determining chemical structures of compounds which employs various methods. Due to variations in efficiency and practicality, certain techniques are favoured over others, and their respective advantages and limitations are explored. Mass spectrometry (MS) and high-resolution mass spectroscopy (HR-MS), facilitate the determination of the elemental composition of a molecule. This involves obtaining the exact mass, specifically the mass number divided by the charge number of the fragment (m/z), to establish the elementary composition of the compound. Additional MS techniques can identify specific fragments of a molecule. The fragmentation of a molecule results in characteristic patterns in the mass spectrum, which are then utilized to extract structural information about the molecule.

Complementary techniques such as optical spectroscopy (OS) are employed to obtain further details beyond the molecular composition. In the examination of how atoms and bonds interact with light, infrared spectroscopy (IR), Raman spectroscopy, and UV-VIS spectroscopy play pivotal roles. The energy associated with any peak in an absorption spectrum aligns with the frequency of a vibration or an electronic transition within a specific part of the molecule. Each spectral band arises from the absorption of energy at a particular frequency of IR radiation, leading to the excitation of bonds in the molecule to a greater state of vibration, either stretching or bending.

In quality control, infrared spectroscopy (IR) stands out as a reliable technique, offering valuable insights into the structure by identifying functional groups. Moreover, IR spectroscopy provides a distinctive fingerprint for a sample, enabling the identification of molecular families or even entire organic molecules by matching the spectrum with entries in an IR database. Ultraviolet-visible (UV-Vis) spectroscopy serves as a quantitative analytical tool. While offering qualitative information about functional groups and types of bonds present, UV-Vis spectroscopy also allows for the determination of molecule concentrations in a non-destructive way.

Given the crucial role of sample purity in influencing the accuracy of these techniques, high-performance liquid chromatography (HPLC) comes into play for molecular characterization by assessing sample purity. HPLC is a method employed

for the purification and separation of different molecules within a mixture, as well as for quantifying them. From this separation, a parameter known as retention time can be obtained; it furnishes information about the molecule, such as its polarity or apolarity.

Nuclear magnetic resonance (NMR) spectroscopy serves as a complementary approach to the techniques mentioned above. It allows the connection of fragments determined by MS and functional groups identified infrared spectroscopy and UV-Vis. This enables the precise determination of the two-dimensional structure. An additional advantage of NMR lies in its capacity to capture through-space correlations, facilitating the determination of the three-dimensional structure. Consequently, NMR spectroscopy stands as the sole method providing structural information at atomic resolution of molecules in solution.

Lastly, X-ray crystallography provide all the aforementioned information. An X-ray structure offers a comprehensive depiction of the molecular structure, detailing each atom and bond, as well as the conformation in the solid phase and within a specific environment. While X-ray crystallography is a robust technique, its accessibility is often limited, and its application is not always feasible due to the necessity for a single crystal.

1.2 3D-conformation assessment

The compound's 3D structure is a key factor in drug development, as it contributes to its bioactivity through its shape, spatial organization, and functional groups exposed to the solvent. X-ray crystallography and NMR spectroscopy are two complementary methods for obtaining 3D structures. In order to conduct X-ray crystallography, having a crystal of the molecule is crucial. However, this may not always be feasible. Moreover, the X-ray procedure allows for the visualization of the molecule's 3D conformation and configuration in solid state and it is strongly affected by the conditions of crystallization, the specific solvent, and the temperature. This could result in a scenario where the X-Ray structure determined differs from the state observed in the typical aqueous buffer solution used for bioactivity testing.

Furthermore, the crystal structures are influenced by intermolecular packing, potentially constraining the molecular conformation to a specific arrangement. In

contrast, in a solution, the molecule has the flexibility to adopt conformations based on their energy levels and the Boltzmann distribution. One of the major advantages of NMR spectroscopy is its capacity to determine the properties of molecules in a solution. The conventional approach to structure elucidation using NMR relies on determining short proton-proton distances through the acquisition of nuclear Overhauser effect (NOE) or rotating-frame Overhauser effect (ROE) experiments and measuring scalar coupling constants (J). While this method is effective for determining the structure of numerous molecules, the use of NOEs and coupling constants may fail in resolving certain structural details, such as stereochemistry or the molecular conformation in solution. This limitation arises because J couplings and NOEs offer only short-range information; in fact, the upper limit of detectable NOE distances is about 5Å. Hence, common NMR approaches may struggle to determine the configuration and conformation. Consequently, it becomes fundamental to employ less conventional NMR parameters, i.e. Residual Dipolar Couplings (RDCs), capable of providing long-range information. Residual dipolar couplings play a crucial role as they offer angular restraints relative to the static magnetic field, facilitating the correlation of distant parts within a molecule.

1.2.1 NMR-based conformational analysis

Nuclear Magnetic Resonance (NMR) spectroscopy is one of the most powerful techniques for structure elucidation and characterization. Moreover, NMR, coupled with computational chemistry, can be useful for the conformational analysis of free ligands in solution.

NMR spectroscopy's role in conformational analysis was firmly established by Garbisch[1], drawing from the pioneering work of Eliel. Garbisch's technique demanded the utilization of models with a geometry fixed structure. During the last two decades NMR-based conformational analysis has been mostly applied to biological systems such as proteins[2], nucleotides[3], oligosaccharides[4] and cyclic peptides[5]. Interestingly Luy et al. in 2005 tried to integrate information gained through residual dipolar couplings with data from interproton distances in order to refine the 3D structure of the cyclic undecapeptide cyclosporine A.[6] They found out that experimental RDCs were not fulfilled neither by the 3D conformation obtained through NOEs nor by the

crystal structure. Hence, they obtained an RDC-refined structure which could also explain both interproton distances and the exposure of all side chains to the solvent in contrast to the more compact crystal structure. Similarly, Thiele et al. rejected the conformational hypotheses regarding the cyclic peptide methylgriselimycin obtained through the analysis of internuclear distances because they were not in line with RDCs experimental data.[7] Nevertheless, they succeeded in defining the conformational ensemble of the peptide in solution through the combination of both ROE distances and RDCs constraints.

Other cases in which determining the conformation or relative configuration solely from scalar couplings ($^3J_{HH}$) and Nuclear Overhauser Effect (NOE) data was challenging, leading to ambiguous conclusions are reported in the literature.[5, 8, 9] The combined method offered a comprehensive and effective strategy for obtaining detailed structural information, particularly in cases where traditional NMR parameters alone did not provide sufficient resolution and accuracy.

Conformational analysis by NMR is based upon the idea of ligand conformational selection, in which the free ligand in solution can be fully pre-organized into the bioactive conformation.[10] Thus, the free ligand can adopt multiple three-dimensional conformations due to rotatable bonds, including the one adopted in the binding pocket of the ligand-target complex.[11] Hence, NMR-based conformational analysis is not suitable for early stage of drug discovery when drug candidates' potency is too low, but it can be useful in SAR studies, discrimination of diastereoisomers, elucidation of 3D structure, e.g. whenever X-ray studies of complexes are not feasible or need clarification.

1.3 Macrocyclic compounds in drug discovery

Macrocyclic compounds are molecules presenting a macrocyclic core composed of at least 12 atoms. They belong to the chemical space beyond the rule of five (bRo5)[12]. The interest towards macrocyclic compounds in drug discovery has increased during the last three decades, since the unknown chemical space to explore is getting smaller. Moreover, compounds such as cyclic peptides, oligonucleotides and macrocyclic molecules represent new therapeutic modalities with regards to challenging druggable targets.[13]

1.4 Aim of the study

The research project presented in this thesis aimed at setting up and validating a protocol with high accuracy and precision, which exploits NMR constraints to obtain structural information from the free ligand, combining experimental information with computational chemistry simulations. Free ligands have unique NMR 1H signatures which can be extremely useful; if a compound - when free in solution - has a prevalent conformer in which protons are subject to perturbations, for instance, from an intramolecular hydrogen bond or tautomeric interconversion, these can be detected by 1D NMR spectra.

This project deepens the integration of Residual Dipolar Couplings (RDCs) as structural parameters for NMR-based conformational analysis, offering an alternative method to conventional NMR spectroscopy or X-Ray crystallography in instances where these techniques prove inconclusive.

A specific purpose of this research study is to apply the method described above to synthetic macrocyclic compounds such as Lorlatinib, Pacritinib and Simeprevir. Low MW macrocycles are receiving increasing attention in drug discovery due to enhanced potency at their target, metabolic stability and physicochemical properties.[10] Moreover, the bound conformation is expected to be present in solution ensembles due to the fact that macrocyclic compounds are conformationally constrained;[11] on the other hand, conformational flexibility through bond rotations may result in exposure of functional groups in response to changes in the polarity of the environment.[12]

The acquisition of a standard set of NMR experiments (^{1}H, ^{13}C, COSY, HSQC, HMBC) and the measurement of $^{1}D_{CH,}$ $^{n}D_{CH}$ and $^{n}D_{HH}$ RDCs using reversibly compressible gels, such as pHEMA, PMMA, PAN, or lyotropic liquid crystals allows to measure local and global parameters. The conformational analysis can be routinely performed using a tool on NMR software MestreLab Research MNova. Once the conformers are generated and NMR constraints measured, experimental values are compared with average values from molecular modelling using a general least-squares fitting procedure of a penalty function, available on the above software, useful to determine the overall conformational amplitudes. The number of statistically relevant conformations is filtered by the Akaike Information Criterion (AIC), which aims at

minimizing the information lost when a model with k free parameters is employed to represent reality.

Secondly, computational chemistry allows to perform conformational analysis by using software such as Schrödinger through different methods. Molecular mechanics (MM) methods calculate the steric energy of a molecule, i.e. the one due to the geometry of the molecule, which arises from specific interactions, comprising bonded interactions (stretching, stretching-bending, bending, torsional) and non-bonded interactions (electrostatic, van der Waals). Different searching methods can be employed to explore the complete conformational space. These methods may involve systematic or randomized analyses.

In addition, the progress in high-performance computing allows to make use of quantum mechanical (QM) methods to predict some NMR parameters, such as isotropic shielding constants, within a reasonable timeframe. Chemical shift calculations, which are conformation-dependent, can be performed for all 3D conformations in order to compare predicted spectra to the experimental one and confirm the proposed conformation derived from NMR-based analysis.

2. Theoretical background

2.1 Basic principles of NMR spectroscopy

Nuclear Magnetic Resonance spectroscopy is based on quantum mechanical laws. Nevertheless, NMR processes can be also explained through classical mechanics models and nuclei can be described as charged particles rotating around their axis, i.e, spinners. Nuclei have a spin angular magnetic momentum which can be described by an intrinsic property possessed by any nuclei, i.e. spin quantum number, known as "I". The value of the spin quantum number depends on the atomic number and the mass number of the nuclei. I can correspond to any integer, half-integer, or zero, but only nuclei that have non-zero spin quantum numbers are NMR-active. If I = 0 (e.g., ^{12}C, ^{16}O etc) the nuclei have not spin angular magnetic moments, and hence they cannot be studied through the NMR spectrometer. The spin number determines the number of quantum mechanical states that a spin can have when placed in a homogenous external magnetic field, B_0, according to the formula (2I + 1). In the case of nuclei with I = ½ ($^1_1H, ^{13}_6C, N, ^{19}_9F$) two possible energetic levels can exist; the low-energy state is more populated according to the Boltzmann distribution. These levels, or magnetic orientations, are labelled as α and β, or m = ½ and m = - ½. The difference of energy between them is given by the equation:

$$\Delta E = \frac{h\gamma}{2\pi} B_0 \tag{2.1}$$

where h is Plank's constant, γ is the gyromagnetic ratio which is specific for each nucleus and B_0 is the external magnetic field.

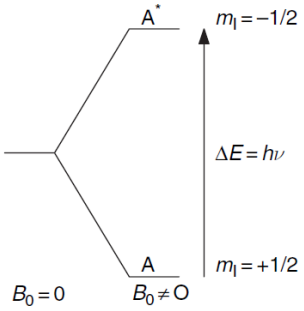


Figure 2.1 Energy levels for spin with I = ½ when placed in a homogenous magnetic field.

A transition from the lower energy state to the higher one can be obtained by using a radiation having a frequency requested from a given magnetic field. This frequency is known as Larmor frequency:

$$\nu = \frac{\gamma}{2\pi} B_0 \tag{2.2}$$

From equation 2.1 and 2.2 the following relationship can be obtained:

$$\Delta E = h\nu \tag{2.3}$$

The higher the magnetic field B_0, the higher the ΔE and the resonance frequency ν. ^1H resonance frequency is used to define the magnetic field strength instead of using Tesla unit.

Before the radiofrequency pulse is applied, all spins precede around the z axis, the vector of the magnetic field. The orientation of these vectors can be parallel (low-energy state) or antiparallel (high-energy state) with respect to the magnetic field. Since the low-energy state is more populated, the sum of these vector is a single vector on the positive z axis and represents the net magnetization vector, M_0 (figure 2.1).

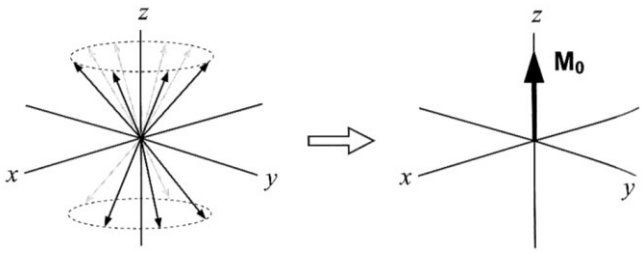


Figure 2.2 – Larmor precession of spins around z-axis. ©Elsevier

When a pulse of a radiofrequency is applied, the net magnetization vector is tilted away from the z axis to the xy plane. The magnetic component in the xy plane can be detected as a signal, the so-called free induction decay (FID), by a receiver. The time-dependent signal is then converted through the Fourier Transform (FT) into a spectrum in the frequency domain.

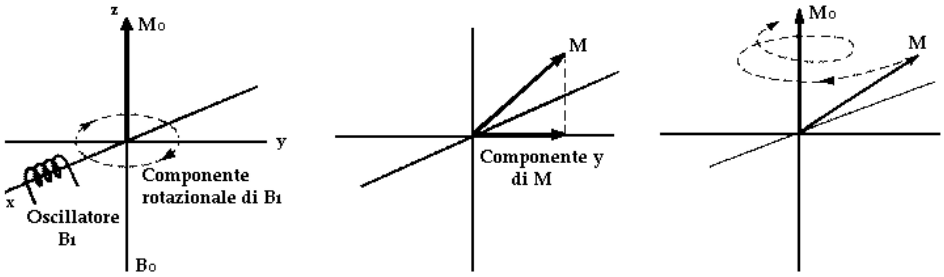


Figure 2.3 – Net magnetization vector M_0 is tilted from z axis towards y axis generating a component in the horizontal plane when the RF pulse is applied. The magnetization M relax back to M_0 in a spiral patter of decreasing amplitude.

2.2. Relaxation processes in NMR

After the RF pulse, M_0 returns to the z axis to restore the equilibrium state through relaxation. There are two relaxation processes involved: spin-lattice and spin-spin relaxation. The spin-lattice relaxation, or longitudinal relaxation (T_1) is the mechanism in which the magnetization vector returns to equilibrium along the axis of the applied magnetic field. During this process the energy is transferred from the nucleus to the molecular lattice as rotational, vibrational and translational energy. It describes the return of the excited spin systems to the Boltzmann distribution.

The spin-spin relaxation, or transverse relaxation (T_2), is the process by which the magnetization vector in the xy plane decays. During spin-spin relaxation there is a mutual exchange of spins between nuclei close to each other with opposite quantum magnetic state. When T_2 relaxation is too fast, peaks become broad, J-coupling may not be resolved or the signal may be difficult or impossible to detect.

2.3 Scalar couplings

Scalar couplings, also known as spin-spin couplings or J couplings, arise from spin-spin interactions mediated through bonds.

Spin-spin coupling is normally observed between nuclei that are one, two and three bonds away. Four-bond coupling can be observed in certain situations, e.g., in conjugated systems like as in aromatic rings, but is not common. The constant J

represents the splitting's size and is independent of the magnetic field; the coupling constant is measured in Hertz. The magnitude of coupling usually decreases when the length or the number of bonds between the two involved nuclei increases. However, three-bond scalar coupling depends on conformation, and hence on dihedral angle. The Karplus equation, published in 1959[14], correlates J coupling of vicinal protons to dihedral angles.

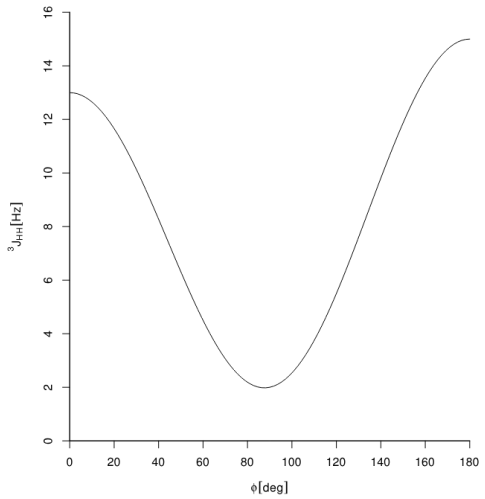


Figure 2.4 – Karplus correlation between J couplings and dihedral angles.

As shown in figure 2.4, the value of J coupling is at its minimum when the dihedral angle is around 90°, while it is at its maximum when the dihedral is 180°.

To predict J coupling, nowadays, Altona equation is usually employed because it considers the effect of substituents according to their electronegativity properties.

2.4 Dipolar couplings

In contrast to scalar couplings, which arise from interactions through bonds, dipolar couplings involve through-space interactions. Conceptually, these spins can be assimilated to two magnets exhibiting inherent rotation at the Larmor frequency. The orientation of each spin in relation to the static magnetic field generates a local magnetic field, either reinforcing or opposing the static magnetic field. Consequently, the nuclei experience a local magnetic field that induces positive or negative shifts in resonance frequencies, contingent upon the orientation of the coupled spins. Since the parallel and antiparallel orientations are approximately equally populated for each

spin, the dipole-dipole interaction, or dipolar coupling, results in a frequency split with twice the dipolar coupling magnitude. Dipolar couplings in solution averaged out to zero due to molecular tumbling. An anisotropic environment is needed to detect them (see chapter 3.2).

Through-space couplings provide the basis for the nuclear Overhauser effect (NOE) which is employed to gain information about interatomic distances.

3. Residual Dipolar Couplings

Residual Dipolar Couplings (RDCs), introduced by an anisotropic medium, are an NMR parameter providing global structural information about orientation of molecules. They represent an important parameter form the determination of the relative spatial arrangement of atoms, i.e. the constitution, configuration and preferred conformation of the molecule.

Residual Dipolar Couplings (RDCs), introduced by an anisotropic medium, are an NMR parameter providing global structural information about orientation of molecules. They represent an important parameter for the determination of the relative spatial arrangement of atoms, i.e. the constitution, configuration and preferred conformation of the molecule.

3.1. Historical background

RDCs have been introduced three decades ago and they have been mostly applied to the analysis of biomolecules, i.e. proteins and nucleic acids.

The pioneering work of Saupe and Englert in 1963 involved measuring molecules in alignment media.[15] They showed the first spectrum of benzene aligned in a liquid crystalline phase.

Since the introduction of the application of RDCs to the analysis of biological macromolecules by Tjandra and Bax in 1997, they have been extensively used in structural biology.

Initially, RDCs were applied to the analysis of proteins; their application was later extended to nucleic acids and studies of various nucleic acid-protein interactions, and to dynamic studies of biological macromolecules in general.

However, RDCs were not as extensively used for the analysis of small organic molecules. The main reason was that the traditional aligning media used for biomolecules were only applicable to water-soluble compounds. In order to be able to partially align small molecules, it was necessary to have access to orienting media compatible with organic solvents.

Courtieu et al.[16] introduced in the mid-1990s the use of chloroform solution of poly-γ-benzyl-L-glutamate (PBLG) to discriminate enantiomers using quadrupolar splittings, dipolar couplings, and chemical shift anisotropies.

The first gel used as alignment medium for the study of a protein was a polyacrylamide gel (PAA) in 2000. More recently, liquid crystalline phases (LLC) soluble in organic solvents were developed. Griesinger and co-workers and Thiele and Berger made significant contributions to the methodology of structure elucidation of small organic molecules by RDCs using PBLG solutions in organic solvents. The development of alignment media compatible with organic compounds opened an extensive range of application, i.e. to small molecules of pharmaceutical interest.

From then, Residual Dipolar Couplings have been extensively used for the structural elucidation of natural products, such as strychnine, sagittamide A[17], archazolide A[18], ludartin[19], cyclosporine A[20]. Interestingly, Luy et al. exploited Residual Dipolar Couplings to distinguish two diastereoisomers of spiroindene assessing the correct relative configuration.[21]

3.2. Theoretical background

Dipole-dipole interactions between spins in a molecule give origin to a parameter known as dipolar coupling. Different from spin-spin scalar couplings, dipolar couplings average to zero when molecules are free to tumble in isotropic solutions.

Residual dipolar couplings (RDCs) are measured as a result of partial alignment of the molecule, in which case the dipolar coupling is not averaged to zero by isotropic tumbling. Partial alignment of molecules is obtained by means of an inert ordered medium.

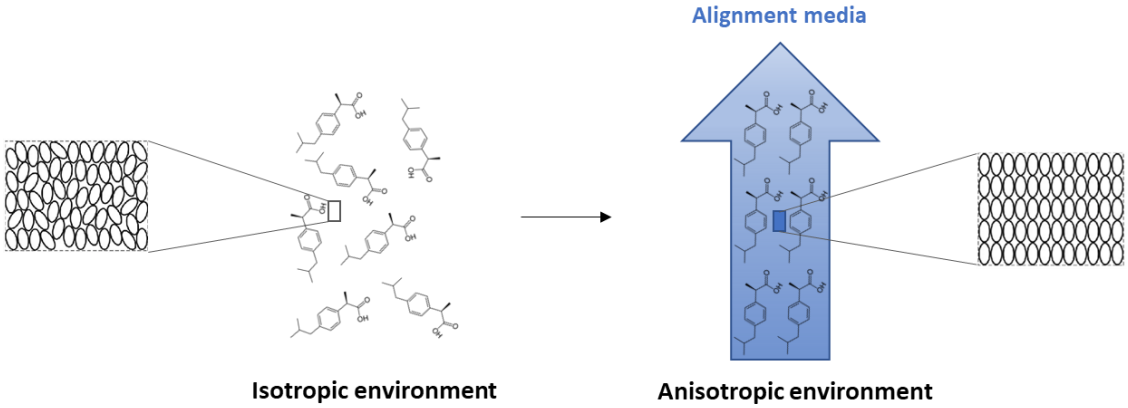


Figure 3.1 – Orientation of a molecule in isotropic vs anisotropic environment.

Dipolar couplings between the nuclei i and j (D_{ij}) contain important structural information since their values depend on the internuclear distance (bond length r) and on the angle θ between the internuclear vector and the axis of the external magnetic field, as shown in Figure 3.1, according to the following equation:

$$D_{ij} = \frac{k}{R^3}\left(cos^2\vartheta - \frac{1}{3}\right) \tag{3.1}$$

where

$$k = -\frac{3}{8\pi}\mu_0\gamma_i\gamma_j\hbar \tag{3.2}$$

in which γ_i and γ_j are the gyromagnetic ratios of nuclear spins i and j, μ_0 is the magnetic constant and ħ the Plank constant.

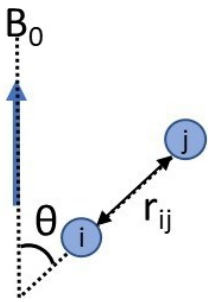


Figure 3.2 – Dipole-dipole interaction between a spin pair. The size of the resultant coupling depends on the distance r between the two nuclei i and j as well as the angle of their internuclear vector relative to the external magnetic field (B_0).

The RDCs still contain the same structural information as the original dipolar couplings and provide global orientation information between remote internuclear vectors. The great power of RDCs is that they provide information about the relative orientations between the internuclear vectors (H-H, H-C, H-N, C-C, etc.) even if they are not covalently bonded.

RDCs report on the orientations of individual bond vectors with regard to the axes of the alignment tensor of the molecule. Thus, the RDCs provide the long-range or global structural information.

3.2.1 Alignment tensor

To derive structural information from Residual Dipolar Couplings, i.e. from the equation (3.1), a complete understanding of the function governing the molecular orientation is essential.

Kramer et al. in 2004 gave an extensive derivation of the alignment tensor matrix by fixing the Cartesian frame of reference (x, y, z) in the molecule itself; they demonstrated how residual dipolar couplings can be calculated by expressing the $\cos^2\theta$ term in equation (3.1) by means of the probability tensor *P*.[22] Since the molecule's axes are fixed, the probability tensor *P* stands for the probability distribution of the orientation of the magnetic field B_0 along x, y, z axes. *P* is a 3x3 matrix, but its corresponding traceless matrix, i.e. the alignment tensor *A*, is usually employed when dealing with anisotropic measurements. Through the relationship between *A* and *P*, Kramer et al. established that equation (3.1) was equal to:

$$D_{ij} = \frac{k}{R^3} \overrightarrow{r^T} A \vec{r} \tag{3.3}$$

in which $\vec{r}$ is the vector between nuclei i and j and *A* is the alignment tensor.

The alignment tensor *A* is a symmetric and traceless matrix (i.e. the sum of the element on the diagonal is zero). Hence, only five terms are unknown to solve the matrix, which means that at least five independent RDCs, i.e. non-parallel RDCs, are necessary to calculate the alignment tensor. Parallel and antiparallel CH vectors would give the same RDC value, and hence redundant orientational information because the angle between parallel internuclear vectors and the magnetic field B_0 is the same.[23]

$$A = \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix}$$

(3.4)

Consequently, since residual dipolar couplings D_{ij} can be experimentally derived, equation (3.3) can be solved and the alignment tensor calculated. In particular, singular value decomposition (SVD) calculations, proposed by Losonczi et al., allows to determine the alignment tensor by solving a system of linear equations.[24] The best-fitting common alignment tensor for all structure can be then determined using the single tensor approximation.[25] Once the alignment tensor is derived, back-calculated RDCs can be easily generated and compared with the experimental RDCs.

Various software such as PALES[26], REDCAT[27], RDC@hotFCHT[28,29] and StereoFitter[30] allows to derive the alignment tensor of an ensemble of conformers from experimental values using the SVD approach. In the calculation of RDCs, the C-H bond distances from the input structures are used. This approach ensures a robust and comprehensive analysis of structural features based on experimental and calculated RDCs.

3.3 Alignment media

3.3.1 Polymeric stretched gels

Polymeric gels are formed by cross-linked polymers which create a sort of grid composed of chains with crosslinking points when they swell in a solvent.[31] These polymeric gels present some cavities in which the molecule can diffuse and be confined into them. When a gel is mechanically stretched or compressed, the molecular grid becomes oriented, resulting in the formation of an anisotropic matrix. Consequently, solute molecules entering such a stretched gel experience partial orientation due to interactions with the polymeric network. The degree of alignment is proportional to the density of the cross-linking.

The concept of stretched polymer gels was initially introduced by Deloche and Samulski in 1981[32] but the first molecules diffusing into stretched polymer gels was studied in 2000 using polyacrylamide (PAA) gels in aqueous solution.[33] In more

recent years, gelatin has emerged as an inert, highly cost-effective, and broadly applicable hydrogel.[34]

As regards polymeric gels compatible with apolar solvents, polystyrene (PS) was used in 2004 by Luy and his coworkers to study strychnine[35] and then menthol and norcamphor.[36] More recently, poly(methyl methacrylate) (PMMA)[19] gels and polyurethane (PU)[37] gels have also been reported as alignment media for apolar organic solvents.

Some polymeric gels compatible with polar organic solvents have been identified. Poly(vinyl acetate) (PVAc)[38] serves as a reliable alignment medium for numerous polar organic solvents such as methanol, acetone, and acetonitrile, dimethyl sulfoxide and dimethylformamide. Other options are poly(2-hydroxyethyl methacrylate) (p-HEMA)[39], also compatible with water, and poly(acrylonitrile)[40] (PAN).

3.3.2 Liquid crystalline phases

Liquid crystalline (LC) phases were identified as the first alignment media for partially orienting solute molecules in NMR assessments. One of the first spectra of a solute exhibiting alignment in a liquid crystal was published in 1963 by Saupe and Englert who exploited the thermotropic LC phase of a p-azoxyanisole derivative to dissolve and orient benzene.[15]

Over the years numerous liquid crystalline phases were developed, both thermotropic and lyotropic liquid crystals. In both scenarios, the alignment is induced by the spontaneous orientation of the liquid crystal when placed in a magnetic field. This induces an anisotropic grid that promotes the alignment of the solute.

Numerous alignment media for aqueous solutions based on liquid crystalline phases have been identified and used in biomolecular NMR spectroscopy. A diverse range of structures, commonly referred to as bicelles[41,42], have been customized to meet specific requirements, particularly in protein NMR. Additionally, for nucleic acids, the preferred alignment medium consists of negatively charged filamentous Pf1 phages[3,43], which are also utilized for negatively charged or neutral proteins[44] and sugars.

Some chiral polypeptides, such as poly-γ-benzyl-L/D-glutamate (PBLG/PBDG), poly-ε-carbobenzyloxy-L/D-lysine (PCBLL/PCBDL), and poly-γ-ethyl-L/D-glutamate (PELG/PEDG) are used both to distinguish between two diastereotopic protons in different molecules (such as in strychnine) and for the assignment of the relative configuration.[45] Another interesting application of chiral lyotropic liquid crystals is the discrimination of enantiomers using a mixture of two chiral polypeptides, poly-γ-benzyl-L-glutamate (PBLG) and poly-ε-carbobenzyloxy-L-lysine (PCBLL).[46]

On the one hand, various liquid crystals-based alignment media are available for non-polar organic solvents (e.g. chloroform, dichloromethane, tetrahydrofuran) such as the aforementioned PBLG, PELG, PCBLL, but on the other hand liquid crystalline phases for organic polar solvents are limited; one example is PBLG dissolved in dimethylformamide (DMF).[47]

4. Materials and methods

4.1 Materials

Lorlatinib was purchased from Sigma-Aldrich (Merck). Pacritinib and Simeprevir were bought from MedChemExpress. PMMA and pHEMA gels were purchased from MestreLab Research and PAN gels were kindly provided by Professor Armando Navarro-Vázquez. Lyotropic liquid crystals tested such as PBPMLG[48] were graciously provided by Professor Christina M. Thiele. PTFE compression device was purchased from New Era. NMR deuterated solvents (0.75 mL ampoules) were purchased from Sigma-Aldrich (Merck).

4.2 Sample preparation

4.2.1 Polymeric gels

Aligning medium is placed into high-precision NMR tube 24/48 hours before the acquisition and the solution containing the analyte of interest is added into it.

After waiting for 24/48 hours for diffusion and swelling processes, polymeric gel is compressed into the 5-mm NMR tube by means of PFTE compression device. The swollen gel normally has a diameter of about 3.6 mm; because the inner diameter of a 5-mm tube is typically 4.0 - 4.2 mm, a layer of isotropic solution forms between the gel and the wall of the tube: as a result, signals from the solution inside and outside the gel are simultaneously observed in NMR spectrum.

When the position of the piston is locked at a given degree of compression, as shown in figure 4.1, the nut and bolt seal the tube tightly and long-duration experiments can be performed. The compression/relaxation process is reversible.

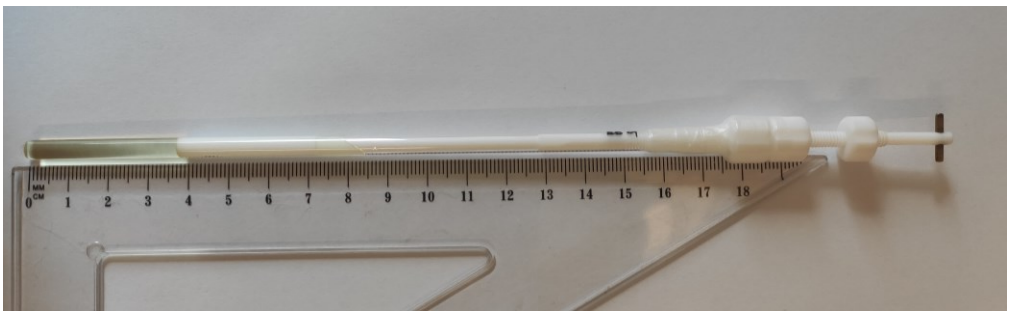


Figure 4.1 – Compression gel device purchased from New Era. It is composed of a plunger, two adapters (similar to a bolt and a screw) and a nut. Shown compression height is 38 mm

whereas swollen height was 45 mm. The plunger can be loosened up or lowered to change the alignment degree.

One of the most important advantages of polymeric stretched gel is that the compression is reversible; both your analyte and the gel can be recovered. The gel can be washed in acetone and MeOH and then put in a vial with the solvent used for the analysis and stored in the fridge.

There is always a thin layer of solvent remaining between the gel and the inner side of the NMR tube. For this reason, an interesting feature of using polymeric gels is that in one spectrum it is possible to measure both isotropic and anisotropic couplings. Nevertheless, this aspect can be regarded as an advantage or a disadvantage because it can lead to overcrowded spectra and overlapping signals.

4.2.2 Lyotropic liquid crystals

The fiber-like polymer or powder forming LLC phases is placed into the NMR tube and the analyte and the solvent are added. The amount of each component is calculated in order to reach the critical concentration to form the liquid crystalline phase. In fact, only above this concentration the system shows properties between solid and liquid states.

The NMR tube is centrifugated up and down at about 1000 rpm for 10 times to homogenize liquid crystals whenever a series of spectra are acquired.

Figure 4.2 – Polymeric fibers composed of PBPMLG into an NMR tube before solvent addition.

The homogeneity of the LLC phase can be visually checked by means of the polarizer filter and experimentally assessed trough ^{2}H spectra (see section 4.2.3). Both polymer and analyte can be recovered from LLC phase through the addition of solvents which keep the analyte in solution and promote the precipitation of the polymer.

4.2.3 Deuterium quadrupolar couplings

A deuterium (1D ^{2}H) spectrum is commonly employed to assess the quadrupolar coupling ($\Delta\nu_Q$) of the solvent, providing information about the level of compression/stretching of polymeric gels; it also allows to verify whether the crystalline liquid phase is formed and equilibrated. The observation of quadrupolar coupling arises from the interaction between the quadrupole moment of a nucleus with a magnetic spin greater than ½ and the electric field gradient (EFG) at the nucleus. In NMR, $\Delta\nu_Q$ is an anisotropic parameter that is typically averaged to zero in solution. Nevertheless, within an anisotropic medium, $\Delta\nu_Q$ causes the solvent signal to split, facilitating the experimental determination of the quadrupolar coupling. This determined value is subsequently employed to monitor the alignment strength of the orienting medium.

The magnitude of the splitting is primarily influenced by the electric field gradient (EFG) on deuterium nuclei within the molecules and the extent of their alignment. For instance, chloroform (CDCl3), characterized by a strong EFG, results in a larger ΔνQ compared to DMSO-d6. Consequently, ΔνQ values from distinct samples, containing different solvents or media, cannot be directly compared.

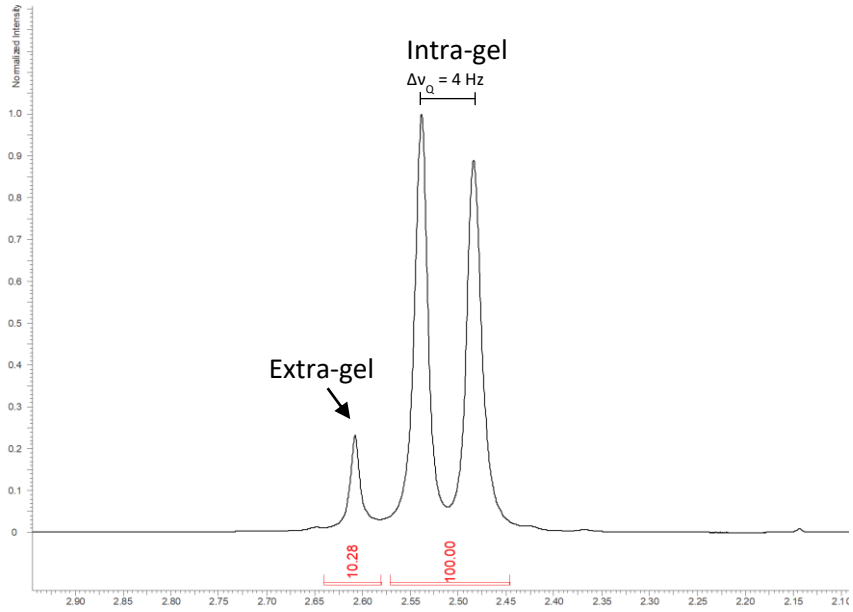


Figure 4.3 – Deuterium splitting of the solvent signal of DMSO-d_6 in poly(2-hydroxyethyl methacrylate) gel, indicated as intra-gel. A peak from residual solvent which has not diffused into the polymer is also present (extra-gel).

The deuterium spectrum is essential for assessing the anisotropy of the medium and ensuring sample equilibration before measuring anisotropic parameters such as residual dipolar couplings (RDCs) or residual chemical shift anisotropies (RCSAs).

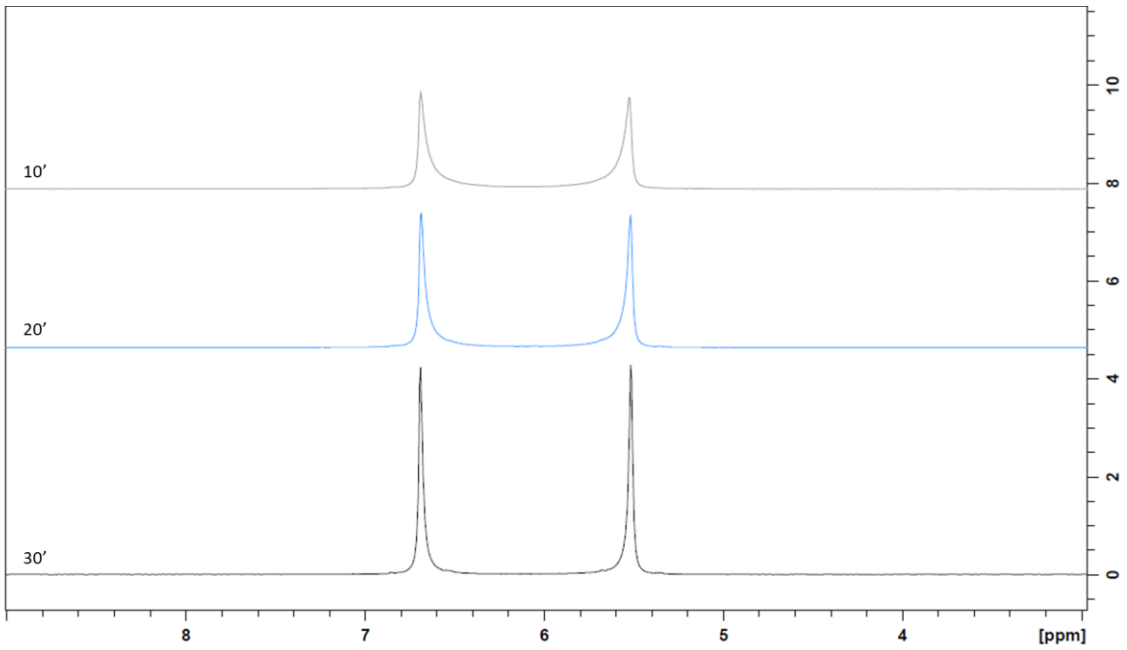


Figure 4.4 – ^{2}H quadrupolar splitting of $CDCl_3$ in PBPMLG liquid crystalline phase. After 10 minutes of equilibration into the NMR magnet, the typical Pake pattern[49,50] of non-equilibrated liquid crystals was evident.

Modifying the ΔνQ value through different media or by means of a stretching/compression is useful to find a balance between spectral resolution and alignment strength. On the one hand, if the sample is excessively compressed, all recorded spectra can result in broad peaks making the extraction of coupling constants challenging or not possible. On the other hand, insufficient alignment provides good resolution but yields small RDC values comparable to experimental error. Consequently, a modulation of alignment is essential.

While the 1D ^{2}H spectrum of the deuterated solvent provides a general characterization of the strength and equilibration of the alignment, it does not yield information about the homogeneity of the sample. Hence, 2H imaging experiment is employed to assess the uniformity of liquid crystals quality and gel. Inhomogeneity can be caused by a non-uniform magnetic field resulting from shimming issues or inhomogeneity within the sample itself such as incomplete equilibration of liquid crystalline phase. This experiment enables the identification of broken gels or

inhomogeneity within gels or liquid crystals by examining the spatial distribution of the 2H quadrupolar splitting along the z-axis of anisotropic samples.

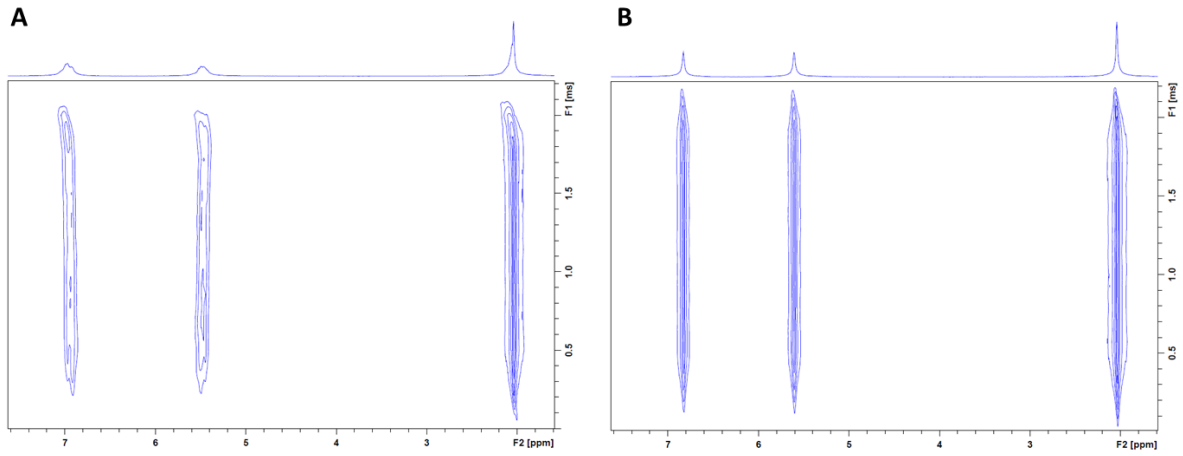


Figure 4.5 – 2H-image spectra of deuterated solvents in liquid crystals to assess the homogeneity of the sample. Inhomogeneous (A) vs homogenous (B) sample. Suboptimal shims usually lead to tilt in both the upper and the lower part of the signals. Sample shown in panel B is well equilibrated and well shimmed.

4.3 NMR equipment

Most of the NMR experiments were performed at the 1H proton resonance frequency of 600 MHz, at a temperature of 298 K, on a Bruker AVANCE III HD 600 instrument equipped with a triple resonance TCI INVERSE H-C/N-D-0.5-Z ATMA 5 mm cryoprobe located at Research Centre of Chiesi Farmaceutici in Parma (IT) at Analytics and Early Formulations department.

Some NMR spectra were recorded using a Bruker AVANCE III HD spectrometer equipped with a QCI cryo probe (1H/19F-31P/13C/15N/2H) with z-gradient and on 600 MHz instrument (Bruker Avance III) is equipped with a triple resonance broadband inverse probe (TBI; $^{1}H/^{31}P$, BB/^{2}H) located at Technische Universität Darmstadt at Organic Chemistry department.

4.4 NMR experiments

4.4.1 RDCs measurements

To measure Residual Dipolar Couplings it is necessary to acquire spectra in both isotropic and anisotropic environments. This need arises from the fact that the couplings measured in an anisotropic state give the sum of scalar couplings (namely J couplings, achievable also in isotropic solution) and dipolar couplings (namely D couplings), as in the following equation:

$$T = J + D \tag{4.1}$$

where T represents total couplings extracted from anisotropic spectra. The difference between total and J splittings will provide the value of residual dipolar couplings.

There are different NMR pulse sequences to measure homonuclear and heteronuclear couplings. During the research project presented in this thesis HD-J-HSQC (Homodecoupled-Heteronuclear Single Quantum Coherence), F1-coupled HSQC, $^{1}J_{CH}$-resolved HSQC, HECADE, PSYCHEDELIC were used. Regarding $^{1}D_{CH}$ it is possible to decide whether to extract couplings using the direct (F2) or the indirect dimension (F1). Advantages and drawbacks of some of these experiments are discussed below.

4.4.1.1 F1-coupled HSQC

F1-coupled HSQC pulse sequence using non uniform sampling (NUS) published by Thiele et al.[51] are employed to extract $^{1}D_{CH}$. One of the drawbacks of this pulse sequence is that individual values of diastereotopic methylene groups cannot be measured[52]. $^{1}D_{CH}$ coming from methylene and methyl groups are measurable as sum of two ($^{1}D_{CHa} + {}^{1}D_{CHb}$) and three contributions ($^{1}D_{CHa} + {}^{1}D_{CHb} + {}^{1}D_{CHc}$) respectively.

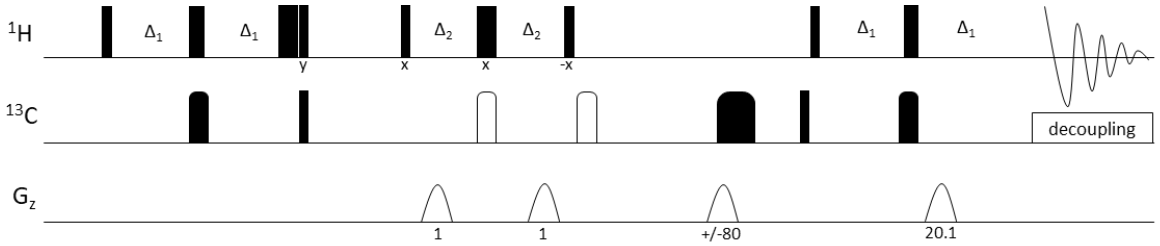
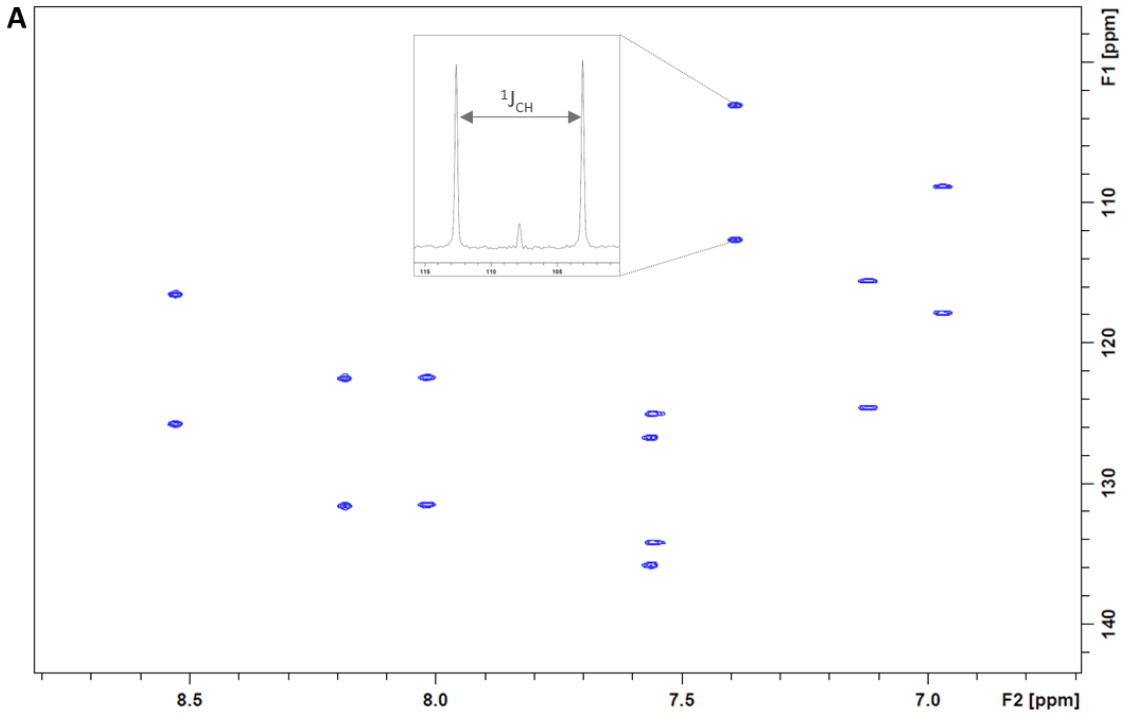


Figure 4.6 – F1-coupled-HSQC pulse sequence (adapted from reference[51]). Thin filled

rectangles are hard 90° pulses, thick filled rectangles stand for 180° pulses. Narrow shaped filled pulses on 13C are adiabatic inversion chirp pulses of 60 kHz sweep width (500 µs duration), wide shaped filled pulses are adiabatic refocusing pulses; open shaped pulses on 13C are BIP (broad band inversion) pulses of length 160 µs. 1 ms gradients were used for echo/antiecho protocol, GARP4 was used for decoupling. Inter-pulse delays Δ1 for INEPT (D1 = 1/(4$^{1}J_{CH}$) and D2 for G-BIRD (Δ2 = 1/(2*$^{1}J_{CH}$) were optimized for 1J = 145 Hz; G-BIRD was used to remove long-range coupling in t1. The scaling factor κ for J-evolution was set to 8, 10 or 16. Gradients were set to 80:20.1:11.*

In general, F1-coupled-HSQC spectra presented throughout this thesis were acquired using 1024 t_1 increments and 2048 complex points in t_2. Zero-filling to 4096 in F1 and squared cosine window function (QSINE, SSB: 2) in both dimensions, direct (F2) and indirect (F1), were applied. Resolution resulted after zero-filling was about 7.7 (F2) and 4.8 (F1).

As regards the measurement of $^{1}J_{CH}$ or $^{1}T_{CH}$ from F1-coupled-HSQC, F1 columns are extracted and superimposed until the maximum overlap of the peak is obtained. Experimental errors associated with the measurements are calculated superposing the two involved peaks on the left side, middle point and right side. This procedure is used for both accurately extracting couplings and discarding measurements with large experimental errors.



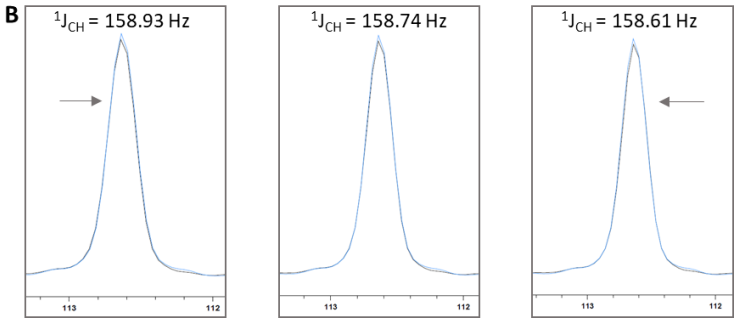


Figure 4.7 – Measurement of $^1J_{CH}$ from F1-coupled-HSQC experiment. (A) Excerpt of aromatic region of isotropic solution of Pacritinib in DMSO-d_6. (B) Extraction of scalar coupling value and experimental error of splitting boxed in panel A. F1 columns of the corresponding peaks of the coupling are extracted from 2D spectrum and left side, middle and right side of peaks are overlaid.

4.4.1.2 HD-J-HSQC

The BIRD-homodecoupled J-resolved HSQC (HD-J-HSQC) reported from Parella et al.[53] is used to measure $^1J_{CH}$ and $^1T_{CH}$.

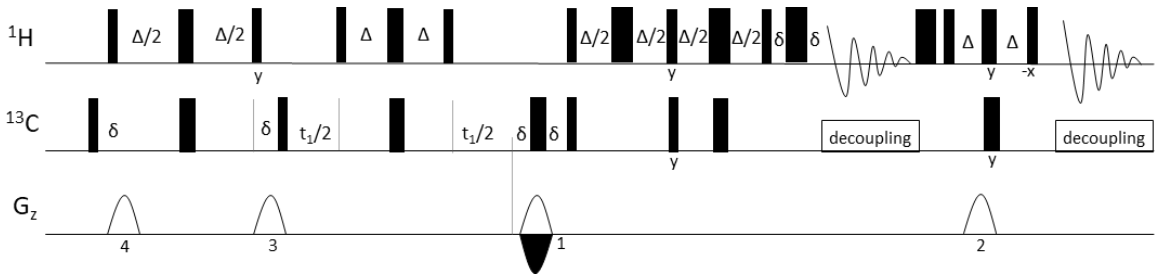


Figure 4.8 – Homodecoupled J-resolved HSQC pulse sequence (adapted from reference [53]). HD-J-HSQC spectra were recorded with a P90 pulse length of 8 μs and 13 μs for 1H and 13C respectively. Broadband ^{13}C inversion and refocusing was accomplished by means of 0.5 ms Chirp pulses (sweep frequency of 60 KHz) and a composite four-Chirp pulse of 2 ms duration. The interval between pulses in the INEPT and BIRD modules was set to 3.5 ms ($\Delta=1/2*^1J_{CH}$) optimized for a $^1J_{CH}$=145 Hz, with a recycle delay (d1) of 1 s. The gradient ratio G1:G2:G3:G4 has been set to 80:20.1:17:33. All experiments were acquired and processed using an echo/antiecho protocol with G1 gradient inversion at each FID new acquisition. The homonuclear BIRD decoupling was used through modules of 13.5 ms (τ) with a total number of loops (n) equal to 4 or 8.

In general, HD-J-HSQC spectra were acquired with 4 scans for each t1 increment using an acquisition matrix of 2048 (F2) *512 (F1). Each dataset was processed by

applying a zero-filling of 2048 in F1 and a function of the type π/2-shifted cos^2 (QSINE, SSB: 2) in both dimensions. After zero-filling the digital resolution of the spectra resulted into 2.6 Hz/pt in indirect dimension and 0.31 Hz/pt in direct dimension.

4.4.1.3 CLIP-HSQC

CLIP-HSQC sequence allows to extract heteronuclear couplings along F2 dimension. On one hand, using F2 dimension. As previously stated (see chapter 4.2.1), whenever a polymeric gel is used as alignment medium, both isotropic and anisotropic signals will be visible in all spectra. Hence, one of the drawbacks of this pulse sequence is that signals in F2 can easily overlap. Moreover, there is no decoupling during acquisition; so, if CLIP-HSQC sequence is used for extracting residual dipolar couplings in solute molecules diffused into compressed polymeric gels it can result in overcrowded rows (which are used to measure couplings in F2 dimension), leading to inaccurate or even impossible measurements.

Acquisition of HSQC spectra with coupling evolution in the direct dimension is the faster way, allows for almost unlimited digital resolution and also allows the assignment of coupling constants in methylene pairs.

The measurement of couplings from CLIP-HSQC can be conducted by extracting rows from the 2D spectra, aligning them, and superimposing one spectrum onto the other to achieve maximum overlap of the multiplet's components. This alignment procedure is employed to estimate the experimental error of the measurements. The strategy of superimposing and aligning the spectra allows for a more accurate determination of coupling values by minimizing any systematic errors that may arise from variations in experimental conditions or spectral features.

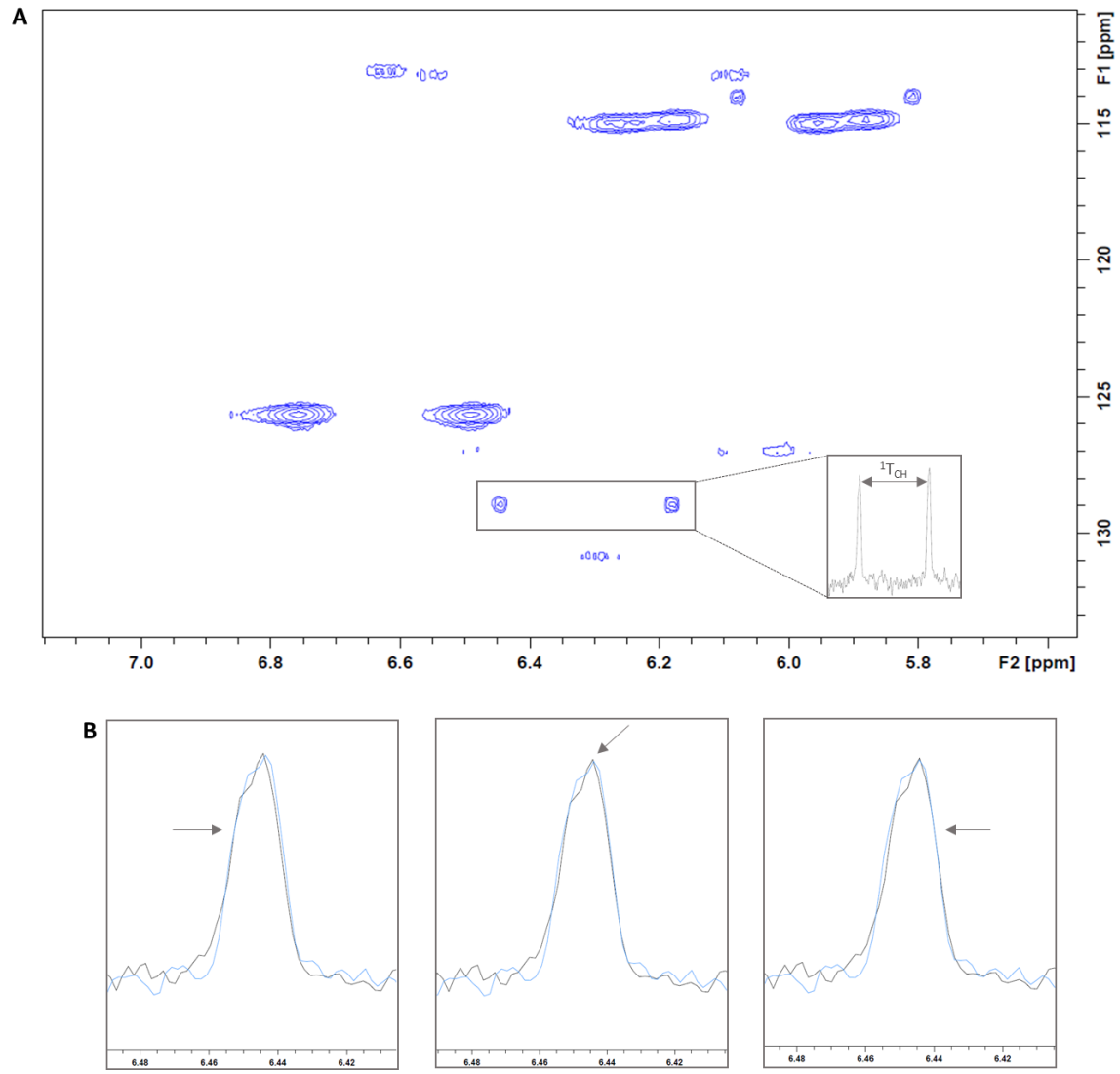


Figure 4.9 – Measurement of $^1J_{CH}$ from CLIP-HSQC experiment. Aromatic region of CLIP-HSQC of Lorlatinib in LLC phase (A). Superposition of the rows of the total coupling boxed in spectrum A to show how to accurately measure the coupling (B). Three values can be measured since peaks are superimposed at the left side, middle point and right side; from these three measurements experimental error is calculated.

4.4.1.4 Long-range Residual Dipolar Couplings' measurements

HSQC-HECADE (developed by Kozminski and Nanz[54]) and PSYCHEDELIC (published by Sinnaeve et al.[55]) sequences are used to measure heteronuclear long-range $^nD_{CH}$ and homonuclear couplings $^nD_{HH}$ respectively.

HSQC-HECADE presents E.COSY-type (exclusive correlation spectroscopy) cross peaks. In HECADE both short-range and long-range couplings are detectable.

Splitting in F1 represents the coupling between an heteronucleus and the hydrogen directly bonded to it (i.e. $^{1}J_{XH}$ or $^{1}T_{XH}$), whereas splitting in F2 represents long-range $^{n}J_{XH}$ or $^{n}T_{XH}$. Spectra are analyzed extracting F2 rows (as shown for CLIP-HSQC in Figure 4.7). Since the sign of couplings needs to be determined when dealing with long-range couplings, HECADE experiment is advantageous because it allows to extract both the magnitude and the sign of couplings. The sign can be determined through the analysis of the tilt E.COSY cross-peaks with respect to the positive or negative slope.[56]

PSYCHEDELIC sequence allows to measure ^{1}H-^{1}H couplings by selectively invert one spin to detect all homonuclear couplings whit respect to the selected proton appearing as doublets in F1 dimension. Measurements of scalar and total couplings from PSYCHEDELIC spectra are performed by extracting F1 columns as demonstrated in Figure 4.5 for F1-coupled-HSQC spectra). As regards the sign of proton-proton couplings, an additional experiment, i.e. selective HSQMBC, is necessary.

4.4.2 Interproton distances' measurements

4.4.2.1 Quantitative NOE (qNOE) experiments

The Nuclear Overhauser Effect (NOE) is a nuclear relaxation phenomenon arising from dipole-dipole coupling between two nuclei that are close in space between each other.

In transient NOE experiments, the inversion of the population of spin I causes the perturbation of another spin S which is close to it in space. The NOE effect induces a magnetization (i.e. a spin polarization) transfer from one nucleus to another via cross-relaxation. Experimentally it is possible to exploit the NOE in different ways:

- 1D experiment: the spin I is selectively inverted, a period corresponding to mixing time is waited, and the spectrum recorded to see the effect on spin I.
- 2D experiment: it is similar to 1D NOE with the addiction of the t1 evolution period to extract the second frequency. In this case, slices of the spectrum can be used in order to extrapolate 1D NOE spectra.

A crucial parameter to set-up is the recycle delay d1, i.e. the time between the end of the acquisition period and the following radio-frequency pulse. Within the experiments mentioned above, it is important that all nuclei are completely relaxed (i.e. the magnetization vector returns to its equilibrium state along the axis of the applied magnetic field B_0) before another RF pulse is applied. For quantitative measurements of interproton distances d1 needs to be between 3 to 5 times the relaxation time T_1 relaxation time. The spin-lattice relaxation T_1 of all spin systems can be calculated through the inversion recovery (IR) sequence;[57] the longest T_1 time is then used to calculate the proper d1 for qNOE measurements.

A very useful sequence used throughout this research project, beyond the standard 1D NOE sequence (selnogpzs, i.e. selective noe with gradients and zero-quantum suppression)[58,59], is 1D-EASY-ROESY pulse sequence (Efficient Adiabatic Symmetrized ROESY).[60]

4.4.2.2 Interproton distances' calculations

NOE intensity between two spins I and S is proportional to the cross-relaxation rate which is proportional to inverse of the sixth power of the distance.[61] The issue is that while the NOE build up during the mixing period, all the spins are also relaxing under the normal process. Some spins relax a little bit faster than others, and this causes some artifacts because NOE intensities are compared in order to extract distances from them; if external relaxation differs between two spins, then that's going to give a difference in the NOE intensities which is nothing to do with distance. The normalization of NOE peak intensities, achieved through the application of "peak amplitude normalization for improved cross-relaxation" (PANIC)[62] method, allows to overcome the aforementioned problem. This normalization technique significantly enhances the accuracy of measured interproton distances; NOE peak intensities are normalized against the selectively irradiated peak if 1D NOE spectra are used or against corresponding diagonal peaks in case of 2D NOESY spectra according to the following equation:

$$I_{Is} = \sqrt{\frac{CP_{IS} * CP_{SI}}{DP_I * DP_s}}$$

(4.2)

where I_{IS} is the normalized intensity of spin H_I and H_S, CP_{IS} and CP_{SI} are the two cross-peaks and DP_I and DP_S are the two diagonal peaks. The normalized intensities, whether obtained directly or through the build-up curve from a series of NOESY spectra with varying mixing times, can be utilized for calculating interproton distances based on a reference proton-pair.

Obtaining quantitative interproton distances from 1D spectra involves acquiring NOESY data with a series of mixing times, typically including at least four different mixing times within the range of 50 and 500 ms. The analysis of these spectra is performed using the initial rate approximation. The intensity of Nuclear Overhauser Effects (NOEs) is plotted against the mixing time for each spectrum, forming the NOE build-up curve. By determining the slope at the onset of the linear signal build-up, corresponding to the build-up rate (σ), it becomes possible to calculate interatomic distances.

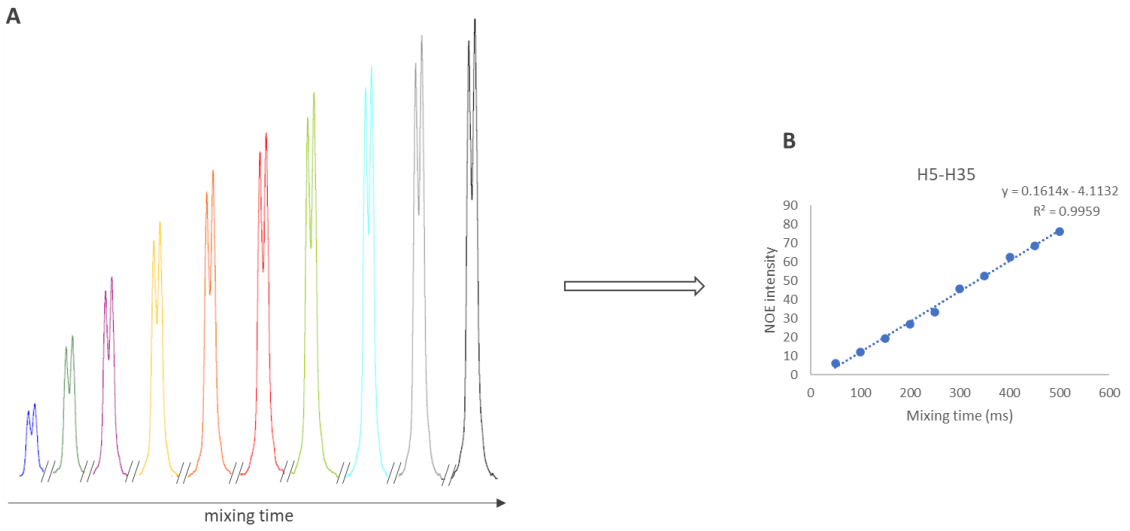


Figure 4.10 – NOE increasing intensities at different mixing time. 1D spectra extracted from proton H35 of Pacritinib in DMSO-d_6 when H5 is selectively inverted (A). Mixing time from 50 ms (blu trace) to 500 ms (black peak). Build-up curve obtained integrating peaks shown in panel A after normalization.

This approach proves more accurate compared to quantifying Nuclear Overhauser Effects (NOEs) from a single spectrum with a fixed mixing time because it allows to exclude some biased distances measured from one spectrum in which some issues

related to relaxation delay, low signal-to-noise ratio, unusual relaxation processes may be present.

The slope of the build-up curve, like the normalized intensities obtained from 2D spectra, serves as a valuable parameter for determining interproton distances through the formula:

$$r_{IS} = r_{ref} * \sqrt[6]{\frac{\sigma_{ref}}{\sigma_{IS}}} = r_{ref} * \sqrt[6]{\frac{I_{ref}}{I_{IS}}}$$

in which r_{ref} is the known interproton distance of the reference proton-pair, I_{ref} and I_{IS} are the normalized intensities of the reference proton-pair and the proton-pair of interest, respectively, and σ_{ref} and σ_{IS} are the slopes of the build-up curves of the reference proton-pair and the proton-pair of interest. The reference proton-pair should be one with a known interproton distance, such as geminal protons (1.78 Å) or the ortho-protons of an aromatic ring (2.54 Å). For methyl groups or, in some cases, methylene groups that rotate quickly enough to yield a single NMR signal for all protons, the interproton-distance is a time-averaged distance from all protons.

4.5 Software

NMR spectra acquisition was performed by means of Bruker TopSpin[63] (versions 3.5 and 3.6). NMR spectra processing and analysis were carried out using both TopSpin and Advanced Chemistry Development, Inc. (ACD/Labs)/Spectrus Processor and NMR Workbook Suite.[64]

Computational analyses were performed using Schrödinger Suite/Maestro software through MacroModel[65] and Jaguar[66].

Multiparametric fittings between experimental NMR parameters and back-calculated information were accomplished through StereoFitter[67] plugin into Mnova (Mestrelab Research) software[68].

4.5.1 StereoFitter

StereoFitter is a Mnova plugin useful for 3D conformational and configurational analysis exploiting NMR constraints. It allows to use various NMR parameters such

as $^{3}J_{XH}$, NOE distances, chemical shifts, residual dipolar couplings, residual chemical shifts anisotropies. StereoFitter analysis is based on the fitting of experimental information against theoretical data derived by the conformational ensemble.

The fitting of experimental NMR properties to the conformation ensemble was proposed by Cicero et al. already in 1995[69] through the so called NAMFIS (NMR Analysis of Molecular Flexibility in Solution) approach.

4.5.1.1 Conformational deconvolution's issue

All properties measured by NMR spectroscopy can be considered as time-average properties. This time-averaging approach is particularly relevant in situations where rapid conformational changes occur, making the observation of distinct resonances for individual conformations unfeasible. Consequently, the measured NMR property T provides a representative average, taking into account the varying probabilities of the molecule adopting different conformations across its potential energy surface (PES). In general, the property T is an average over the entire potential surface, where the value of that specific property in a particular position r is weighted by the probability of the molecule existing in a specific conformation.[30]

Starting from all observed NMR properties, StereoFitter can be employed to try to establish the individual contributions of each conformation within the population by fitting experimental data to individual values calculated from theoretical conformers. This calculation is performed through the non-linear least square Levenberg-Marquadt algorithm. One of the requirements is that the conformational space is completely explored not to miss any relevant conformation.

The maximum number of conformers to be combined for each solution can be limited. Nevertheless, it is advisable to run the fitting without this limitation and use the automation mode.

4.5.1.2 Statistical parameters

In Stereofitter the Akaike Criterion (AIC)[70] is employed to balance model complexity and goodness of fit, aiming to avoid excessively long computation times. The AIC is calculated for models with an increasing number of conformations until a minimum

AIC, with respect to the number of conformations, is identified. If increasing the number of conformations results in higher AIC values for all ensembles, the computation is halted, and the model with the lowest AIC is presented as the solution to the structural problem. The solution with the lowest AIC is considered the simplest model that aligns with the experimental data within experimental uncertainties.

$$AIC = \chi^2 + 2(m - 1)$$

The AIC is composed of the χ^2 term (sum of deviations normalized by variances) and another term that penalizes models with a high number of conformations. Hence, the best-fitting model in StereoFitter calculation, i.e. the one with the lowest differences between experimental and back-calculated parameters (NOEs, RDCs, $^3J_{HH}$, etc.) and the one with the lowest number of conformations per solution.

As regard to experimental uncertainties, error values can be inserted for each measurement or a common standard error can be used for measurements related to each parameter. Unless otherwise stated, standard errors of 0.2 Å and 1.2 Hz were used in fitting analyses presented in this thesis for interproton distances and residual dipolar couplings respectively.

5. Results and Discussion

5.1 Conformational analysis of Lorlatinib

Lorlatinib, (16R)-19-amino-13-fluoro-4,8,16-trimethyl-9-oxo-17-oxa-4,5,8,20-tetrazatetracyclo[16.3.1.$0^{2,6}$.$0^{10,15}$]docosa-1(22),2,5,10(15),11,13,18,20-octaene-3-carbonitrile, is a selective and reversible 12-membered macrocyclic drug approved for treatment of non-small cell lung cancer (NSCLC). It was first approved by the FDA in November 2018[71] and subsequently by the EMA in 2019[72] for the treatment of patients who were previously treated for advanced ALK-positive non-small cell lung cancer. It presents a good profile of effectiveness, safety, selectivity and bioavailability. The target proteins of Lorlatinib are anaplastic lymphoma kinase (ALK) and c-ros oncogene1 kinase (ROS1).[73] As compared to linear analogues, the macrocyclic skeleton reduces the number of rotatable bonds, leading to a more compact bioactive conformer when bound to its target. X-ray structures revealed that Lorlatinib adopts the same bioactive conformer in the two target protein crystals. An interesting property of this molecule is the permeability: it exhibits high permeability of both cell membrane and blood brain barrier[74]. Hence, Lorlatinib may display molecular chameleonicity and change its conformation depending on the polarity of the environment, as demonstrated by Erdélyi et al. in a study in which NOE distances and molecular dynamics simulations were used.[75] For this feature Lorlatinib was chosen as a good candidate to be investigated using polar and apolar solvents. DMSO-d_6 was used to mimic body fluids and blood brain barrier, whereas $CDCl_3$ (dielectric constant $\varepsilon = 4.71$) was utilized to simulate the phospholipid bilayer.

The initial step of the analysis consists of conformational search. The conformational sampling was conducted in water and chloroform at first. Two Monte Carlo methods were compared, MCMM[76] (Monte-Carlo Multiple Minimum) and SPMC[77] (Systematic Pseudo-Monte Carlo) torsional sampling; OPLS-2005[78,79] force field was employed with GB/SA solvation model[80] using water and chloroform. Conformational searches were carried out using an enhanced torsion sampling using 10000 maximum steps and maximum atom deviation cut-off set to 0.5 Å. Conformations within 10 kcal/mol from the global minimum were saved. Since only five conformations were found, the potential energy window for saving structure was set to 20 kcal/mol in order not to miss any relevant conformation. In the following step, the more accurate

electronic energy calculated through quantum mechanics will be considered to eliminate all the conformations above 10 kcal/mol. All conformers obtained through the two methods and solvents were compared between each other; they resulted to be the same, except for the ranking of the potential energy. For this reason, the conformational search was performed again in gas phase which resulted in 14 conformers. All the conformations were geometry optimized by means of quantum mechanics using DFT/B3LYP[81,82,83] theory and 6-31G+* functional. The simulated solvents were added during the single point energy calculations, needed to obtain isotropic shieldings of proton, and hence predicted chemical shifts.

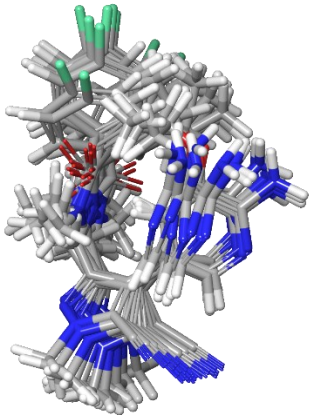


Figure 5.1 – Superposition of 14 conformers found through MCMM and SPMC sampling.

5.1.1 Conformational analysis in DMSO-d_6

5.1.1.1 NMR assignment

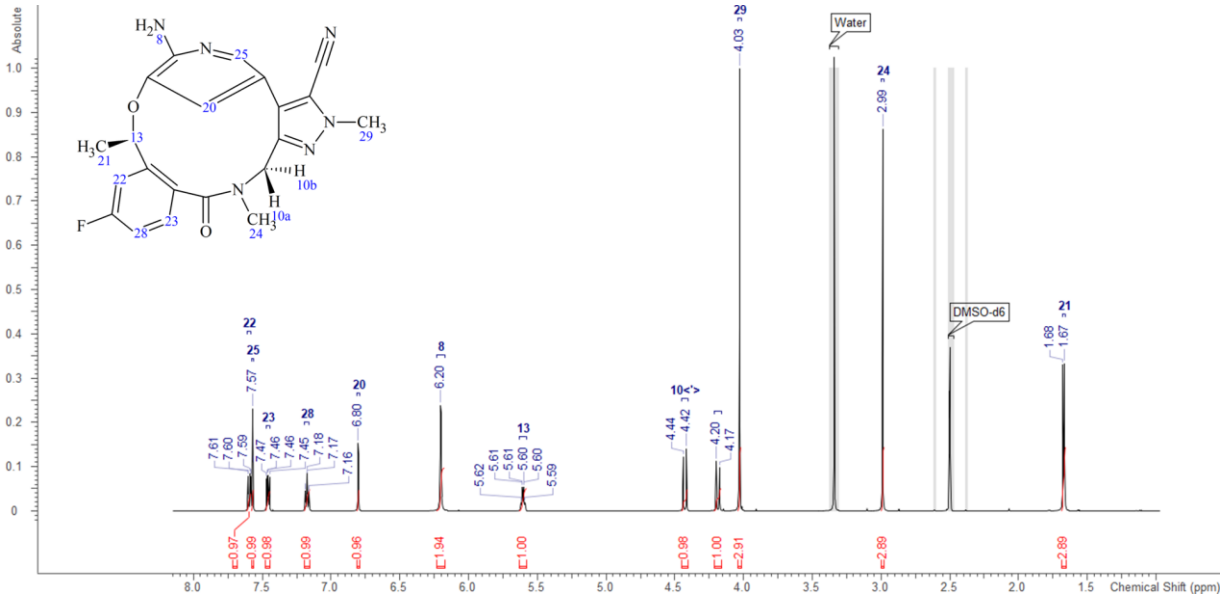


Figure 5.2 – ^{1}H spectrum of Lorlatinib in DMSO-d_6.

NMR assignments were carried out through acquisition and analysis of ^{1}H, COSY, TOCSY, ^{1}H-^{13}C HSQC, ROESY, ^{1}H-^{13}C HMBC, ^{1}H-^{15}N HMBC spectra at 600 MHz AVANCE IIID TBI spectrometer.

No.	F2 atom (^{1}H)	F1 atom (^{13}C)	F2 (ppm)	F1 (ppm)
1	38	22	7.6	113.19
2	43	25	7.57	136.12
3	39	23	7.46	128.02
4	44	28	7.18	114.49
5	34	20	6.8	118
6	33	13	5.6	70.17
7	31	10	4.43	45.89
8	32	10	4.19	45.88
9	45,46,47	29	4.03	38.08
10	40,41,42	24	2.99	30.35
11	35,36,37	21	1.68	21.46

Table S1. ^{1}H and ^{13}C NMR assignments of Lorlatinib in DMSO-d_6.

5.1.1.2 QM-based analysis

As a first step analysis, quantum mechanics simulations were performed to calculate chemical shifts, δ, to compare them to experimental ones in order to check whether or not it was possible to identify some diagnostic features.

Moreover, different combinations of QM methods, functionals, basis sets and solvent models were tested to obtain chemical shifts' predictions as accurate as possible. Ab initio method Hartree-Fock, HF, was tested in combination with cc-pVDZ functional, whereas DFT method was used with different level of theory, i.e. B3LYP-d3 and M06-2X, and basis sets, namely 6-311G++** and cc-pVDZ.

RMSEs' analysis was conducted taking into account all protons except for exchangeable protons or protons that could present strong interaction with the solvent, i.e. amine or alcoholic protons.

In fact, it was evident that NH_2 was badly predicted in all conformers. The predicted value was 5 ppm, whereas the experimental chemical shift was 6.2 ppm. Amine protons are exchangeable protons, which means that they easily exchange with other labile protons or deuterium in solution; this process impacts the shape of their peaks, which are usually broader, their detection and chemical shift. Specifically, the chemical shift of signals arising from exchangeable protons is affected by pH, temperature, hydrogen bonding, i.e. it can be down or up-shifted. For example, lone electron pairs of sulfoxide groups of DMSO-d_6 can electrostatically interact with the protons of amine groups; this interaction causes a deshielding effect, leading to a shift of the signal to higher resonance frequencies (low-field) relative to the signal of the same proton in a solvent without interactions. Since the polarizable continuum model (PCM)[84] used as implicit solvent model for quantum-mechanics simulations presented throughout this thesis does not take into account electrostatic interactions, such as hydrogen bonding, the chemical shifts of exchangeable protons (−OH and −NH) were excluded from RMSEs' analysis.

Density Functional Theory calculations with B3LYP or M06-2X functionals coupled with two different basis sets (6-311G++**, cc-pVDZ) showed almost the same trend in chemical shifts prediction except for few conformers (e.g. conformer 14). Conformers 1 and 5 presented the lowest RMSE values. On the other hand, Hartree-Fock calculations led to large errors in isotropic shieldings of all protons of all

conformers. This could be due to the lack of electronic correlation contributions in HF theory since Hartree-Fock theory assumes that electrons move independently from each other in an averaged field potential caused by the other electrons.

It was noted that conformers 1 and 5 have similar errors between all experimental and calculated chemical shifts because they are pseudo-symmetric. In fact, superimposing the pyrazole it's evident that conformations 1 and 5 differ for the dihedral N6-C11-C10-N4 (highlighted in Figure 5.3) which is -96.4 and 96.4 respectively. For this reason, all protons of the two conformers have the same chemical surroundings determining isotropic shieldings, and hence chemical shifts.

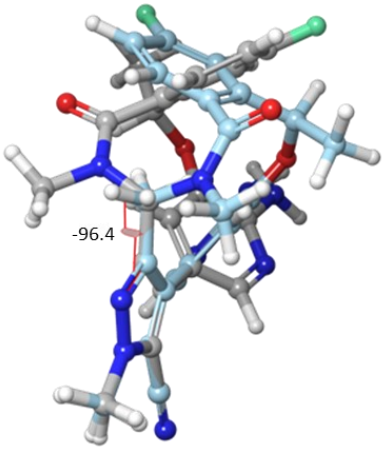


Figure 5.3 – Superposition of conformer 1 (grey) and conformer 5 (light blu). Dihedral angle N6-C11-C10-N4 of conformation 1 is highlighted in red.

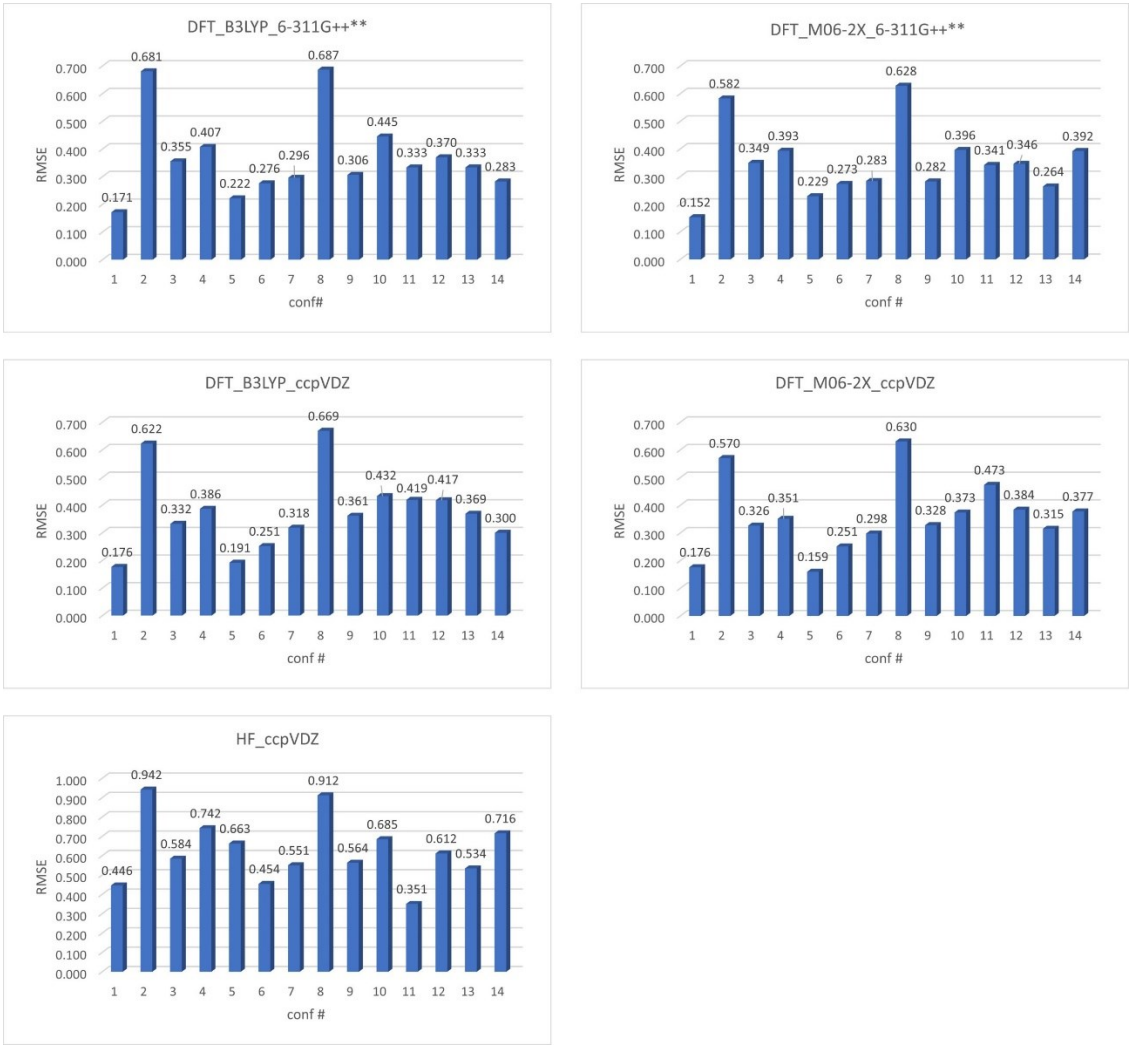


Figure 5.4 – QM calculations of ^{1}H chemical shifts of Lorlatinib in DMSO. Histogram bars represent root mean square error of predicted chemical shifts.

An important feature found in some conformers was that the predicted chemical shift of one of the two diastereotopic protons, H10'; it was higher (deshielded, or down-field) in conformers n. 2, 8 and 10, leading to very large errors, between 1.08 and 1.57 ppm.

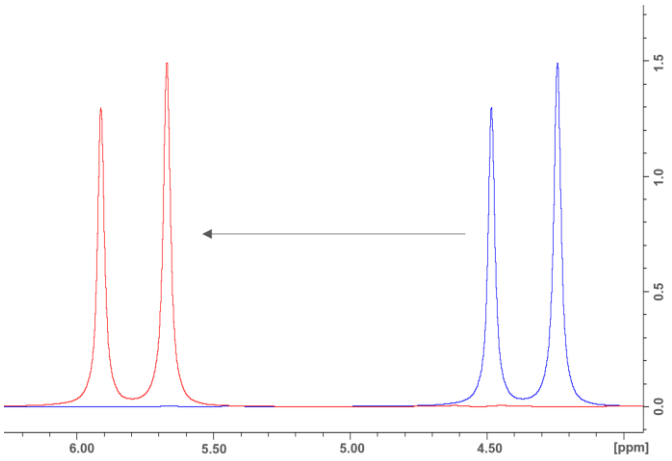


Figure 5.5 – Deshielding of proton H10' of Lorlatinib in DMSO-d_6. Experimental resonance peak is represented by the blue line at 4.4 ppm, whereas calculated chemical shift by the red line at 5.8 ppm.

A common characteristic of these conformations is that the amide configuration is trans, and diastereotopic proton H10' lies on the same plane of carbonyl group. Hence, the down-field shifting could be due to the cone of the carbonyl group. The electrons of a double bond are free to move between the two atoms; the applied magnetic field stimulates the motion of these electron which leads to a generation of a complementary magnetic field. This induced magnetic field induces shielding (up-field) or deshielding (down-field) of surrounding nuclei depending on their relative position.[85]

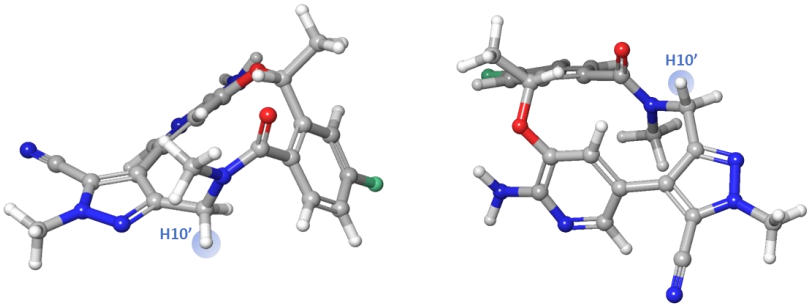


Figure 5.6 – Conformer 1 (left) with cis-amide configuration compared to conformer 2 (right) with trans-amide.

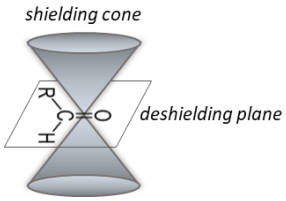


Figure 5.7 – Anisotropic effect of carbonyl group. The electronegativity of carbonyl can cause shielding and deshielding effects.

All the other conformers with trans amide were carefully inspected because also conformers n. 3, 4, 7, 9, 12, 13 presented an error on the prediction of H10'. The magnitude of the error was lower (between 0.4 and 0.9 ppm) than the one found in the above structure. In this case in addition to the electronic effect of the carbonyl group, the chemical shift of the diastereotopic proton can be also affected by the pyridine ring. In fact, with regard to aromatic rings, the ring-current causes a deshielding effect in protons lying on the plane of the ring and a shielding effect in protons above and below the aromatic ring (see Figure 5.4).

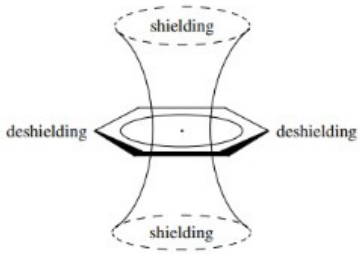


Figure 5.8 – Aromatic ring anisotropic effect with deshielding plane and shielding cone.

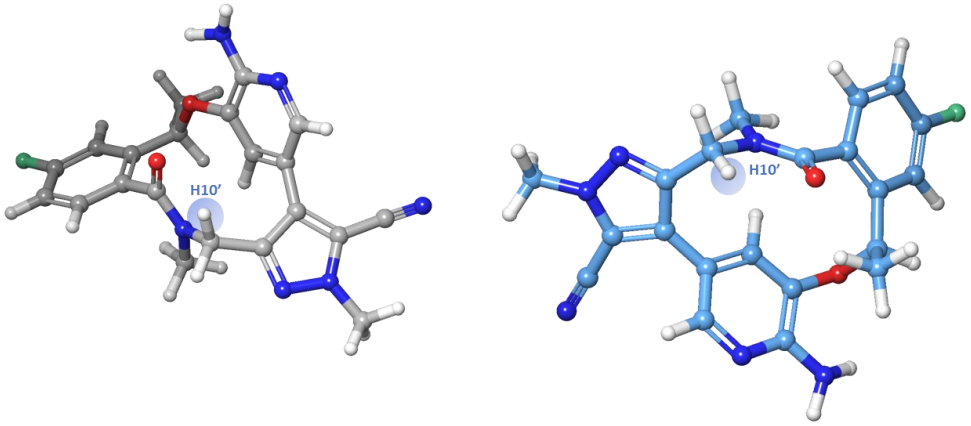


Figure 5.9 – Conformer 9 (grey) vs 10 (blue). The configuration of the amide is trans in both cases, but in conformer 9 proton H10' can be shielded by the aromatic shielding cone of the pyridine ring.

5.1.1.3 Analysis by means of interproton distances

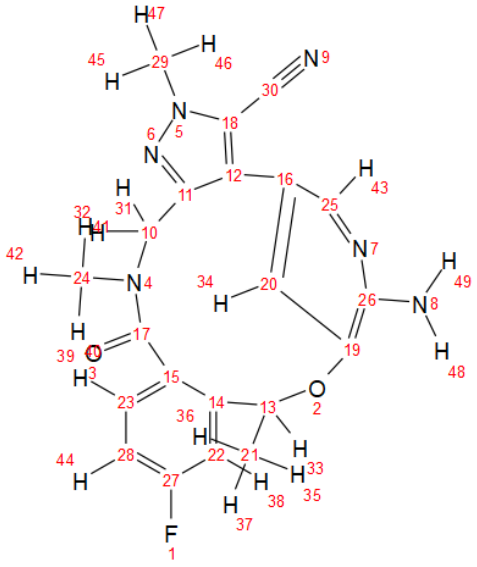


Figure 5.10 – 2D structure of Lorlatinib.

NMR measurements of interproton distances were conducted using both 1D and 2D pulse sequences. T_1 measurement, which is necessary to calculate d1 delay, was performed through T_1 Inversion Recovery (T1IR) pulse sequence; aromatic proton H25 at 7.57 ppm presented the slowest T_1 relaxation time (4.22 s), hence a d1 delay of 22 s was used.

Interproton distances derived from build-up rate obtained from 1D ROE spectra (PANIC approach, see chapter 4.4.2) were compared to both distances obtained from 2D ROESY and distances measured through a single 1D ROE spectrum using 500 ms mixing time. Distance between methylene protons were used as reference.

		build-up curve	500 ms	
			1D ROESY	2D ROESY
Proton i	**Proton j**	**distance (Å)**	**distance (Å)**	**distance (Å)**
H22				
	H21	2.756	2.370	2.662
H25				
	H21	3.723		
H23				
	H28	2.550	-	-
	H10'	3.162	2.870	2.909
	H10''	2.946	2.600	2.650

		build-up curve	500 ms	
			1D ROESY	2D ROESY
Proton i	**Proton j**	**distance (Å)**	**distance (Å)**	**distance (Å)**
H23	H24	4.075	-	4.252
H20				
	H25	4.147	-	-
	H13	2.098	1.980	2.029
	H10"	2.795	2.550	2.617
	H24	3.283	3.040	3.491
H13				
	H22	3.800	-	3.74
	H20	2.036	1.972	
	H24	4.036	-	4.349
	H21	2.435	2.412	2.471
H10'				
	H23	2.741	2.710	-
	H10"	1.825	1.768	1.792
	H24	2.620	2.556	2.921
H10"				
	H23	2.497	2.453	-
	H20	2.527	2.472	-
	H10'	1.810	1.731	-
	H24	3.792	3.516	4.042
H29				
	H24	3.931		4.707
H24				
	H23	3.826	3.254	-
	H20	3.248	3.000	-
	H13	3.826	3.262	-
	H10'	2.802	2.586	2.921
	H10"	3.921	3.249	4.042
	H29	3.975		
H21				
	H22	2.258	2.208	-
	H13	2.175	2.105	-

Table S2. Comparison between distances extracted from 1D and 2D ROESY using a single point (500 ms as mixing time), and 1D ROESY using 10 mixing times to create build-up curves.

All distances were under-estimated when extracted from a single 1D ROE experiment with a single mixing time (500 ms) when compared with distances resulted from build-up rates arising from the collection of ten spectra for each single resonance varying the mixing time from 50 ms to 500 ms. Specifically, distances derived from the build-up curve were 3.16% larger (ranging from 0.96% to 5.96%) for distances up

to 2.5 Å and 9.77% (from 1.14% to 20.68%) for distances exceeding 2.5 Å. On the other hand, the deviation between distances derived from 2D ROESY spectrum in comparison with distances obtained through PANIC method was 2.23% (between 1.47% and 3.4%) for distances up to 2.5 Å and 7.02% (ranging from 1.61% and 19.73%) for distances greater than 2.5 Å.

Since PANIC approach allows to obtain more accurate measurements, interproton distances calculated through this approach were used for the following analyses. Specifically, only distances coming from linear regression curve ($R^2 > 0.95$) obtained through at least 6 points (out of 10 mixing times used, ranging from 50 to 500 ms).

Proton i	**Proton j**	**δ_i (ppm)**	**δ_j (ppm)**	**Atom$_i$ #**	**Atom$_j$ #**	**σ_{ij}**	**R^2**	**distance r_{ij} (Å)**	**average distance (Å)**
22	21	7.6	1.68	38	35, 36, 37	0.0522	0.974	2.756	**2.507**
25	21	7.57	1.68	43	35, 36, 37	0.0086	0.95	3.723	
23	28	7.46	7.18	39	44	0.0832	0.998	2.550	
23	10'	7.46	4.43	39	31	0.0229	0.988	3.162	**2.952**
23	10"	7.46	4.19	39	32	0.035	0.976	2.946	**2.722**
23	24	7.46	2.99	39	40, 41, 42	0.0054	0.985	**4.023**	
20	25	6.8	7.57	34	43	0.0045	0.963	4.147	
20	13	6.8	5.6	34	33	0.2685	0.999	2.098	**2.067**
20	10"	6.8	4.19	34	32	0.048	0.999	2.795	**2.661**
20	24	6.8	2.99	34	40, 41, 42	0.0183	0.993	3.283	**3.265**
13	22	5.6	7.6	33	38	0.0076	0.968	**3.800**	
13	20	5.6	6.8	33	34	0.3211	0.999	2.036	-
13	24	5.6	2.99	33	40, 41, 42	0.0053	0.96	**4.036**	
13	21	5.6	1.68	33	35, 36, 37	0.1098	0.991	2.435	**2.305**
10'	23	4.43	7.46	31	39	0.054	0.995	2.741	-
10'	10"	4.43	4.19	31	32	0.6202	0.999	1.825	-
10'	24	4.43	2.99	31	40, 41, 42	0.0707	0.997	2.620	**2.711**
10"	23	4.19	7.46	32	39	0.0945	0.995	2.497	-
10"	20	4.19	6.8	32	34	0.0879	0.996	2.527	-
10"	10'	4.19	4.43	32	31	0.6511	0.998	1.810	ref
10"	24	4.19	2.99	32	40, 41, 42	0.0077	0.959	**3.792**	
29	24	4.03	2.99	45, 46, 47	40, 41, 42	0.0062	0.97	3.931	**3.953**

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	$Atom_i$ #	$Atom_j$ #	σ_{ij}	R^2	distance r_{ij} (Å)	average distance (Å)
24	20	2.99	6.8	40, 41, 42	34	0.0195	0.9974	3.248	-
24	10'	2.99	4.43	40, 41, 43	31	0.0473	0.993	2.802	-
24	29	2.99	4.03	40, 41, 44	45, 46, 47	0.0058	0.964	3.975	-
21	22	1.68	7.6	35, 36, 37	38	0.1729	0.988	2.258	-
21	13	1.68	5.6	35, 36, 38	33	0.2162	0.993	2.175	-

Table S3. Interproton distances (Å) derived from build-up curves of Lorlatinib in DMSO-d_6. Atom numbers of StereoFitter and Maestro. Average distance indicates distances mutually obtainable from the two protons of interest for which the average was used for StereoFitter analysis. Values in bold were employed in following analyses.

ROE-based analysis was performed fitting 11 interproton distances on StereoFitter.

The analysis was conducted without restricting the number of conformations that the software could combine to give a solution because there is no evidence to assume that only one conformer exists in solution.

Interproton distances' analysis did not lead to a unique and clear solution. They were evaluated through the relative probability of each solution and the Akaike Criterion (AIC) which gives the goodness-of-fit. Since more than one solution has a high relative probability (near to 1) and similar AIC number (between 3.48 and 4.48) there is less confidence that the first solution can be considered as the best and the only result. As it is shown in table S4, the first five solution have relative probability greater than 60% and all linear regressions of experimental vs calculated distances present R^2 values around 0.9 (figure 5.11).

	Solution	1	2	3	4	5
		Conformer 1: 100%	Conformer 3: 52.21%	Conformer 1: 72.36%	Conformer 1: 74.02%	Conformer 1: 64.35%%
		-	Conformer 9: 47.49%	Conformer 9: 27.64%	Conformer 4: 25.98%	Conformer 2: 35.65%
	AIC	3.478	3.627	4.226	4.428	4.486
	χ^2	3.478	1.627	2.226	2.428	2.486
	Relative probability	1.000	0.928	0.688	0.622	0.604
	NOE distances (Å)					
Atoms (^{1}H-^{1}H)	Experimental	Calculated	Calculated	Calculated	Calculated	Calculated
31,(40,41,42)	2.711	3.027	3.037	3.014	3.039	3.041
31,39	2.952	2.836	2.904	2.975	2.967	3.029
32,(40,41,42)	3.792	3.849	3.828	3.843	3.84	3.839
32,39	2.722	2.482	2.658	2.613	2.606	2.665
32,34	2.661	2.991	2.607	3.086	3.084	2.951
33,38	3.8	3.724	3.601	3.649	3.726	3.728
33,(40,41,42)	4.036	4.447	3.79	3.984	4.555	4.593
33,(35,36,37)	2.305	2.602	2.599	2.6	2.602	2.603
33,34	2.067	2.039	2.078	1.97	2.133	2.093
34,(40,41,42)	3.265	3.605	3.511	3.389	3.286	3.781
(35,36,37),38	2.507	2.964	2.609	2.81	2.892	2.869

Table S4. Experimental vs calculated distances comparison obtained for the 5 best-fitting solutions. All the solutions have a similar AIC value and their relative probability value indicates that all of them can be considered as possible solutions.

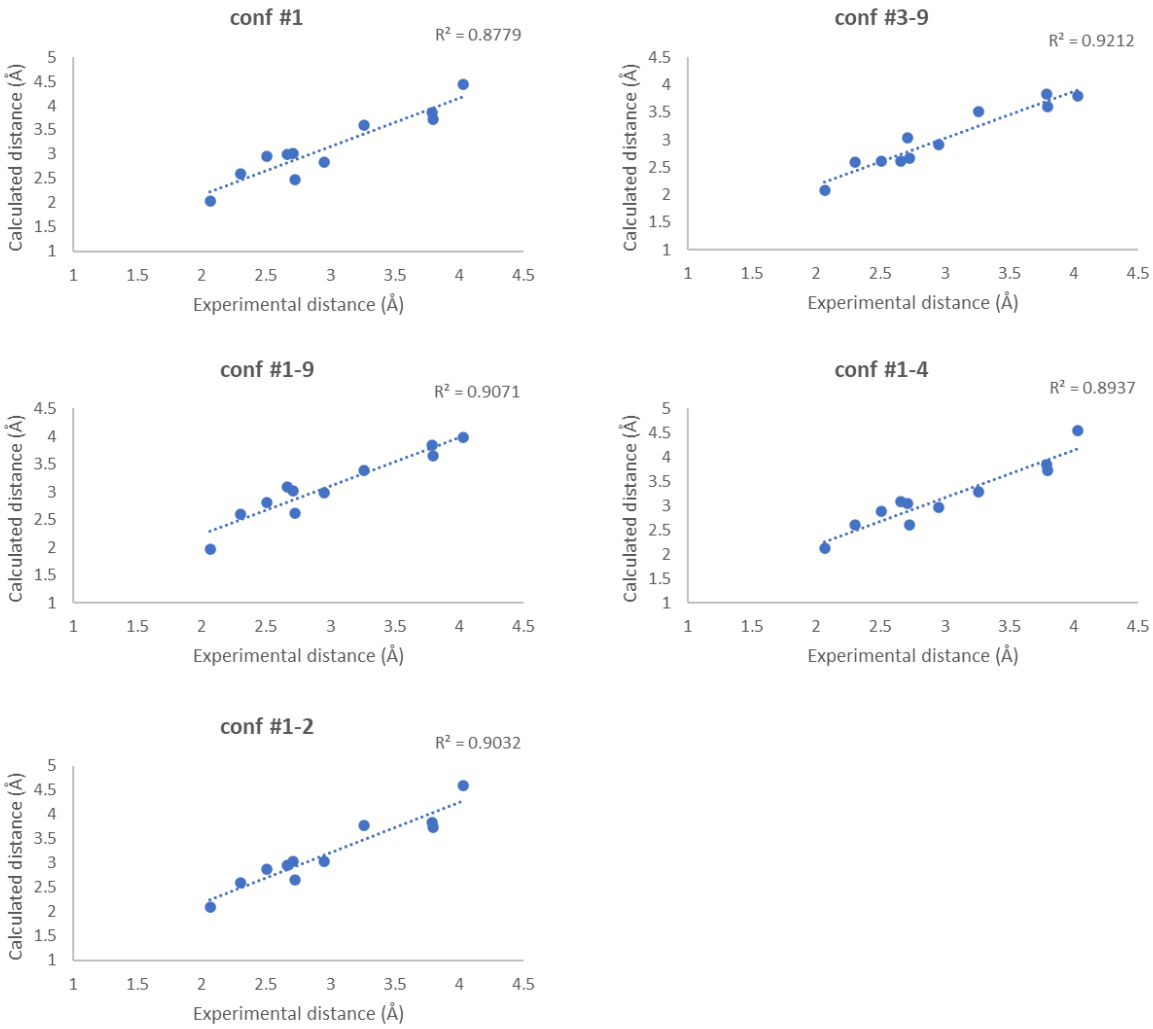


Figure 5.11 - Linear relationships between experimental and back-calculated distances derived from results shown in Table S4.

Even if NOE-based analysis was not able to provide an unambiguous solution, it proved to be remarkably effective in the assignment of diastereotopic protons H10' and H10'' (corresponding to atoms 31 and 32). These protons were permutated, and the fitting analysis performed again. Parameters indicating the goodness-of-fit increased a lot compared with the previous analysis; AIC value of the first five solutions was between 13.6 and 14.4, which is 3.9 times larger than AIC value obtained in the previous analysis with the opposite assignment of diastereotopic protons. Moreover, it is important to note that at least 4 out of 5 distances of involved diastereotopic protons of all theoretical models presented large errors with respect to experimental values.

	Solution	1	2	3	4	5
		Conformer 2: 67.32%	Conformer 3: 44.32%	Conformer 2: 100%	Conformer 1: 36.91%	Conformer 1: 50.97%
		Conformer 3: 32.68%	Conformer 4: 55.68%	-	Conformer 2: 63.09%	Conformer 4: 49.03%
	AIC	13.595	13.928	13.957	14.044	14.414
	χ^2	11.595	11.928	13.957	12.044	12.414
	Relative probability	1.000	0.847	0.834	0.799	0.664
		NOE distances (Å)				
Atoms (^{1}H-^{1}H)	Experimental	Calculated	Calculated	Calculated	Calculated	Calculated
32,(40,41,42)	2.711	3.824	3.822	3.821	3.831	3.832
32,39	2.952	2.868	2.733	4.593	2.91	2.767
31,(40,41,42)	3.792	3.076	3.083	3.069	3.053	3.05
31,39	2.722	3.114	2.978	4.272	3.274	3.131
31,34	2.661	4.117	4.08	4.213	4.171	4.149
33,38	3.8	3.732	3.73	3.735	3.731	3.728
33,(40,41,42)	4.036	4.814	4.79	4.912	4.719	4.66
33,(35,36,37)	2.305	2.604	2.602	2.604	2.603	2.601
33,34	2.067	2.365	3.215	2.226	2.143	2.252
34,(40,41,42)	3.265	4.324	3.179	4.214	3.943	3.079
(35,36,37),38	2.507	2.717	2.712	2.723	2.803	2.833

Table S5. StereoFitter results of distances' fitting when H31 and H32 were permutated. Strongly deviating distances of the first solution are boxed in blue.

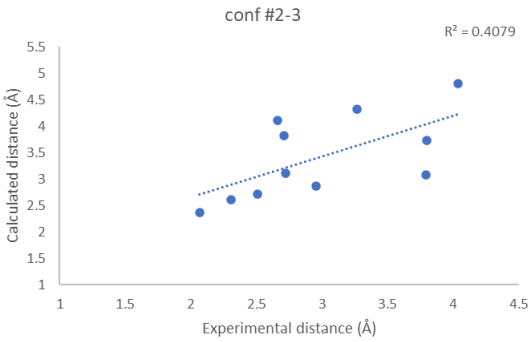


Figure 5.12 – Linear relationship between experimental and back-calculated distances from solution 1 presented in stable S5.

Results shown in Table S5 allowed the correct assignment of diastereotopic protons: downfield peak corresponds to H10' (atom 31) whereas upfield peak corresponds to H10'' (atom 32).

Another important aspect of NOE-based analysis was that it was useful to exclude from the possible solutions one of the best conformers found from QM predictions'

analysis of chemical shifts, i.e. conformer 5. AIC value of distances' fitting for conformation 5 was very high (273.3); 7 out of 11 distances of the theoretical model presented errors between 0.82 and 2.38 Å with respect to the experimental data.

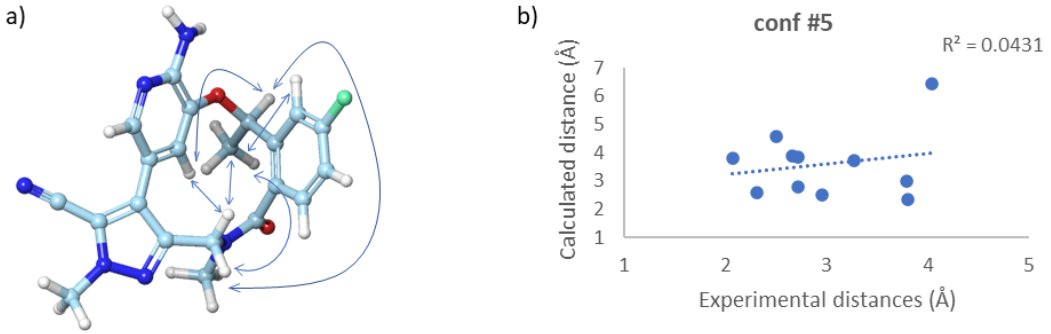


Figure 5.13 – Distances' analysis of conformer 5. a) Conformer 5 with drawn distances of back-calculated interproton distances presenting large errors. b) Linear regression of experimental vs calculated distances of conformer 5.

5.1.1.4 Analysis through Residual Dipolar Couplings

A solution 20.5 mM of Lorlatinib was added into an NMR tube containing a dried stick of polymeric gel p-HEMA (poly(2-hydroxyethyl methacrylate)). After 48 hours, the gel had swollen radially and longitudinally; this sample was used for the acquisition of spectra to extract Residual Dipolar Couplings.

Deuterium spectrum was recorded to assess the degree of alignment after compressing the gel with the specific device. A quadrupolar splitting of 9.52 Hz was achieved, a proper degree of alignment for this particular gel.

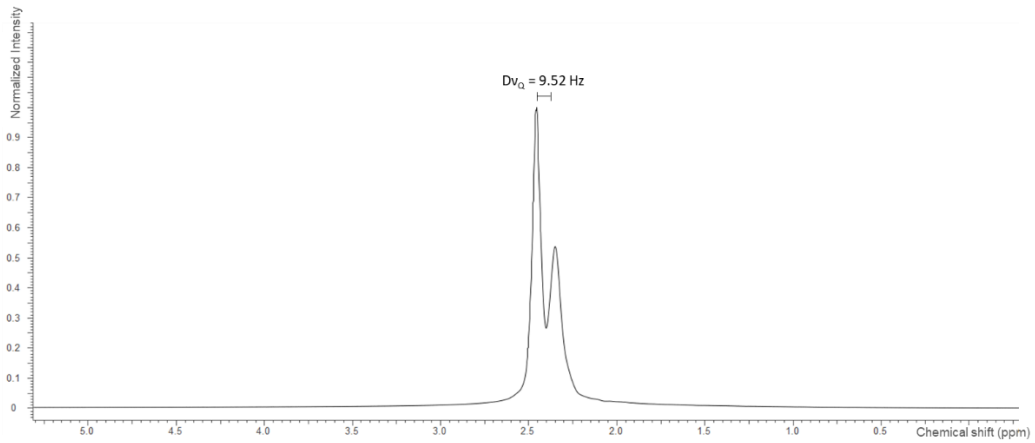


Figure 5.14 – Deuterium spectrum of poly(2-hydroxyethyl methacrylate) gel in DMSO-d_6 after compression of the piston. The quadrupolar splitting (Dv_Q) of the solvent peak after the compression is shown. The peak on the left side results from the overlapping of the residual solvent peak in solution and solvent peak inside the gel.

^{1}H proton spectra were recorded both when the polymeric gel was completely relaxed and when it was compressed through the compression device to follow the inclusion of the analyte into the gel. As it is shown in the inset in Figure 5.15, when the gel is stretched, the ratio between the analyte's signals outside the gel and inside the gel is 0.7:1. It can be also noted that multiplicities of peaks of signals coming from the analyte diffused into the gel are well recognizable, meaning that the sample is well shimmed.

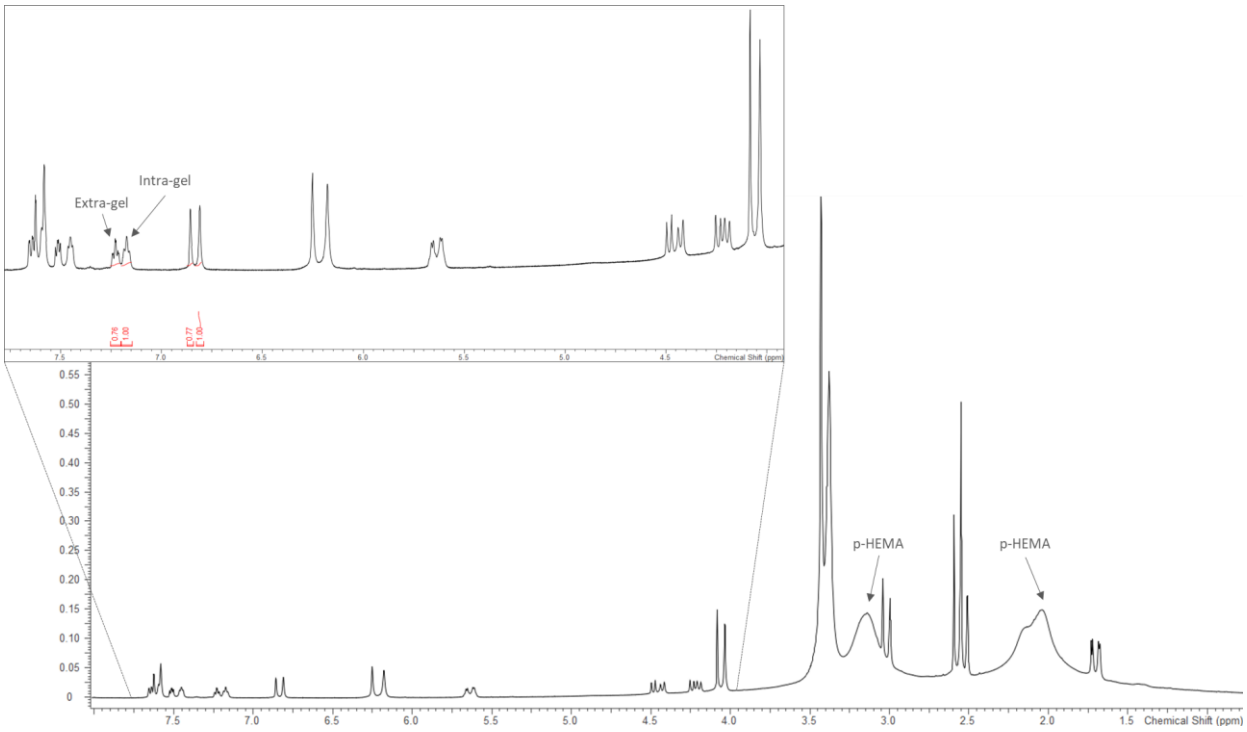


Figure 5.15 – ^{1}H spectrum acquired on a stretched p-HEMA gel in DMSO-d_6 gel with ≈ 5 mg Lorlatinib. In the inset integrals of two aromatic protons of signals inside and outside the gel are shown. In the proton spectrum broad peaks belonging to the gel are visible in the aliphatic region at 3.10 and 2.06 ppm. However, RDC measurements from 2D spectra was not affected by signal overlap with the polymer.

After the sample was compressed through the specific compatible plunger to reach a weak alignment degree ($D\nu_Q$ = 9.52 Hz, see Figure 5.16), the amount of residual isotropic signals decreased, and the ratio was 0.5:1. The compression led to peak broadening, as shown in Figure 5.16.

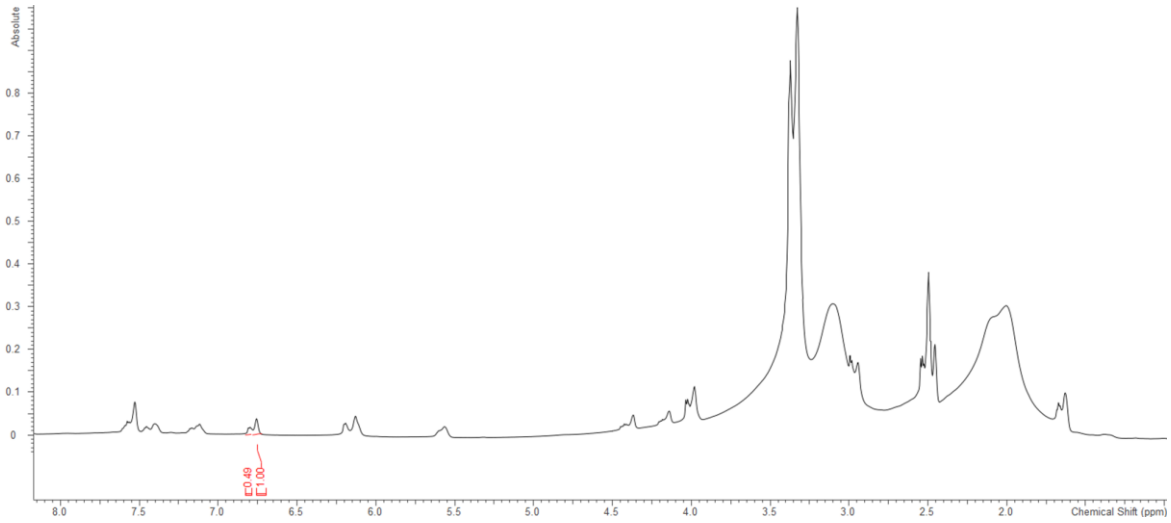


Figure 5.16 – ^{1}H spectrum of Lorlatinib included in the compressed polymeric gel p-HEMA. The residual isotropic signal still visible is half the amount of signals coming from the anisotropic environment.

HD-J-HSQC spectrum was acquired to extract one-bond heteronuclear $^{1}D_{CH}$.

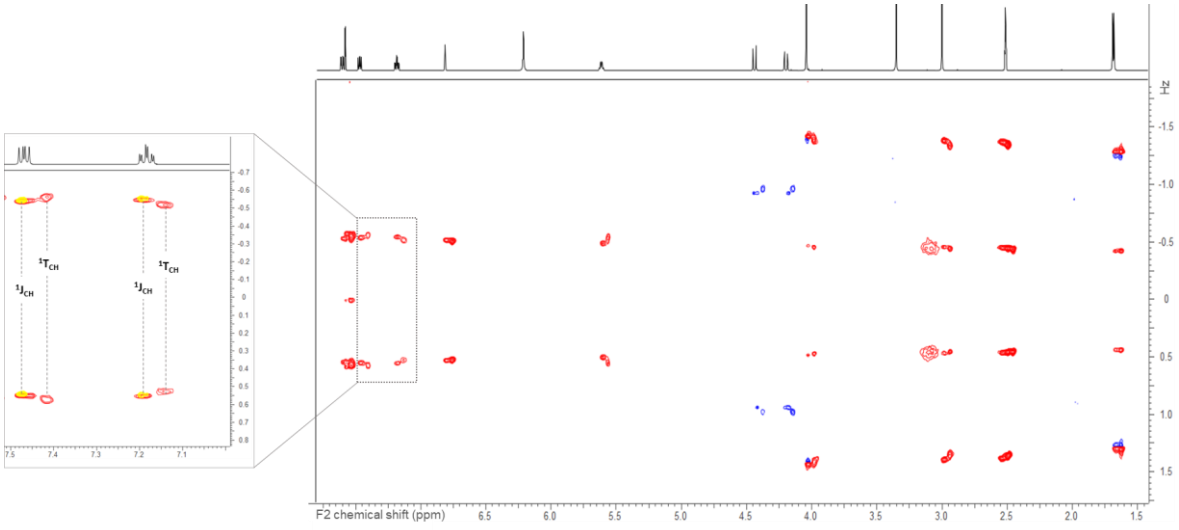


Figure 5.17 - HD-J-HSQC spectrum of Lorlatinib diffused into p-HEMA gel recorded to extract total dipolar couplings ($^{1}T_{CH}$). In the inset overlap between isotropic (yellow peaks) and anisotropic (red peaks) couplings of C23-H39 at 7.46 ppm and C28-H44 at 7.18 are shown. Isotropic spectrum was acquired to facilitate the distinction between $^{1}J_{CH}$ and $^{1}T_{CH}$ since they are both visible in the anisotropic spectrum and the signals' intensity is quite similar.

The extracted one-bond heteronuclear isotropic, total and dipolar couplings are shown in Table S6.

A total of 9 residual dipolar couplings were measured and put into the input file of StereoFitter. Since only the sum of methylene protons is measurable in the F1 dimension, half the sum of dipolar couplings was used for the analysis. Similarly, dipolar couplings of methyl groups are extracted as the sum of the contribution of each CH, and hence dipolar couplings arising from CH_3 groups need to be divided by 3.

^{13}C	^{1}H	δ ^{13}C (ppm)	δ ^{1}H (ppm)	$^{1}J_{CH}$ (Hz)	$^{1}T_{CH}$ (Hz)	$^{1}D_{CH}$ (Hz)	$^{1}D_{CH2}$ (Hz)	$^{1}D_{CH3}$ (Hz)
25	43	136.12	7.57	174.960	169.810	-5.15		
23	39	128.02	7.46	164.000	171.170	7.17		
28	44	114.49	7.18	166.000	157.530	-8.47		
20	34	118	6.8	157.350	155.980	-1.37		
13	33	70.17	5.6	149.530	164.210	14.68		
10	31,32	45.88	4.43, 4.19	281.039	292.204	11.16	5.58	
29	45,46,47	38.08	4.03	429.410	421.790	-7.62		-2.54
24	40,41,42	30.35	2.99	417.860	404.870	-12.99		-4.33
21	35,36,37	21.46	1.68	388.590	389.850	1.26		0.42

Table S6 – Experimental $^{1}J_{CH}$ and $^{1}T_{CH}$ values extracted from an HD-J-HSQC spectrum. Dipolar couplings are calculated according to the equation $^{1}D_{CH} = {}^{1}T_{CH} - {}^{1}J_{CH}$.

Multiparametric fitting, using both RDCs and NOE distances was performed in StereoFitter. No restrictions regarding the maximum number of conformations used per solution was added. This analysis resulted in two best-fitting solutions, both having the conformer 1 as the most populated. In the first solution, conformer 1 represents 71.8% of the population in solution, whereas conformer 4 represents 28.2%. It is important to note that the improvement of the AIC value when switching from 1-conformer solution to 2-conformers solution is very low (from 6.127 to 5.628). For this reason, the relative probability of the 1-conformer solution is high (0.78%).

From linear regression functions (Figure 5.18), it is clear that R^2 values of linear fittings regarding residual dipolar couplings are 0.99 for both solutions, whereas R^2 is slightly higher when conformers 1 and 4 are combined respect to the solution in which conformer 1 alone represents 100% of the population; R^2 values are 0.89 and 0.88 respectively.

	Solution	1	2
		Conformer 1: 71.83%	Conformer 1: 100%
		Conformer 4: 28.17%	-
	AIC	5.628	6.127
	χ^2	3.628	6.127
	Relative probability	1.000	0.779
	NOE distances (Å)		
Atoms (^{1}H-^{1}H)	Experimental	Calculated	Calculated
32,(40,41,42)	2.711	3.040	3.027
32,39	2.952	2.980	2.836
31,(40,41,42)	3.792	3.839	3.849
31,39	2.722	2.619	2.482
31,34	2.661	3.093	2.991
33,38	3.8	3.727	3.724
33,(40,41,42)	4.036	4.565	4.447
33,(35,36,37)	2.305	2.602	2.602
33,34	2.067	2.143	2.039
34,(40,41,42)	3.265	3.264	3.605
(35,36,37),38	2.507	2.886	2.964
	RDCs (Hz)		
Atoms (^{13}C-^{1}H)	Experimental	Calculated	Calculated
(10,31) (10,32)	5.58	5.763	5.788
13,33	14.68	14.981	14.583
20,34	-1.37	-1.268	-0.962
(21,35) (21,36) (21,37)	0.42	0.115	0.624
(24,40) (24,41) (24,42)	-4.33	-3.82	-3.863
25,43	-5.15	-4.868	-5.17
28,44	-8.47	7.197	7.046
23,39	7.17	-8.363	-8.146
(29,45) (29,46) (29,47)	-2.54	-1.757	-3.972

Table S7 - StereoFitter results of experimental vs back-calculated distances and residual dipolar couplings fitting of Lorlatinib in DMSO-d_6 in p-HEMA gel.

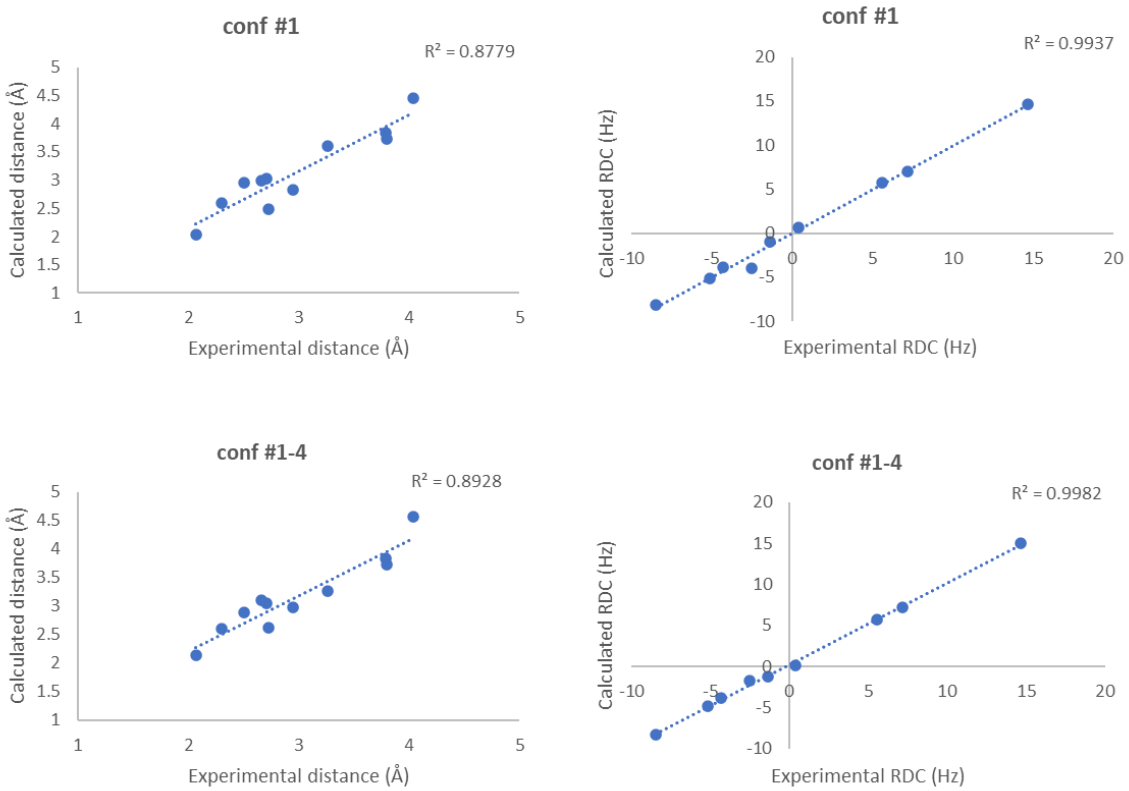


Figure 5.18 – Linear regressions of StereoFitter results presented in Table S7.

Conformer 1 was interestingly superimposable to crystal structure (PDB: 5AA9)[86] obtained in complex with L1198F Mutant Human Anaplastic Lymphoma Kinase. Superposition based on heavy atoms led to a RMSD of 0.0348 Å.

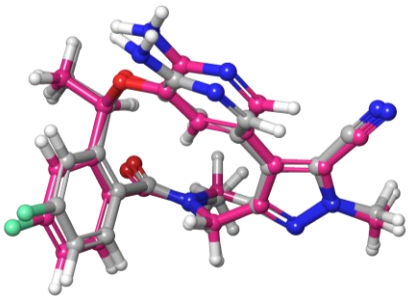


Figure 5.19 – Superposition between conformer 1 (grey) and crystal structure (pink).

In conclusion, results obtained through the multiparametric analysis performed in polar aprotic solvent, DMSO-d_6, agreed with those coming from the QM simulations concerning the prediction of chemical shifts, through which two probable

conformations were identified. One of the two, conformer 1, was found to be superimposable to the Lorlatinib crystal obtained in presence of the target protein ALK (Anaplastic Lymphoma Kinase). Hence, NMR-based conformational analysis in solution allowed to find a conformation similar to the one found in the crystalline form.

5.1.2 Conformational analysis in $CDCl_3$

5.1.2.1 NMR Assignment

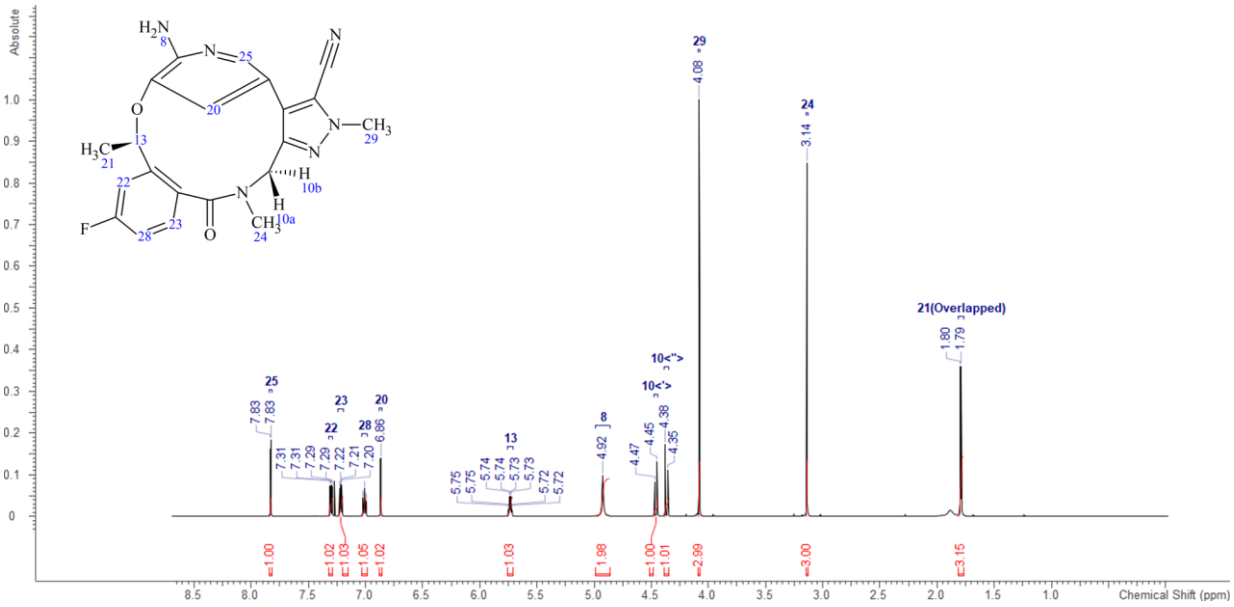


Figure 5.20 – ^{1}H spectrum of Lorlatinib in $CDCl_3$.

NMR assignments were carried out through acquisition and analysis of ^{1}H, COSY, TOCSY, ^{1}H-^{13}C HSQC, ROESY, ^{1}H-^{13}C HMBC, ^{1}H-^{15}N HMBC spectra at 600 MHz AVANCE III HD TBI spectrometer.

No.	F2 atom (^{1}H)	F1 atom (^{13}C)	F2 (ppm)	F1 (ppm)
1	43	25	7.83	136.87
2	38	22	7.29	128.12
3	39	23	7.22	128.12
4	44	28	7.01	115.32
5	34	20	6.87	118.53
6	33	13	5.73	71.44
7	31	10	4.45	47.47
8	32	10	4.37	47.47
9	45,46,47	29	4.07	38.6
10	40,41,42	24	3.14	31.55
11	35,36,37	21	1.8	22.48

Table S8 – NMR assignments of Lorlatinib in $CDCl_3$.

5.1.2.2 Analysis by means of interproton distances

Interproton distances were measured through one-dimensional spectra of each proton of interest using 10 mixing times (from 50 ms to 500 ms).

		build-up curve	500 ms	
			1D ROESY	2D ROESY
		distance (Å)	distance (Å)	
H22				
	H13	4.240	3.878	3.582
	H21	2.705	2.547	2.707
	H22	4.270		
H23				
	H10'	2.901	2.737	2.558
	H10"	3.135	2.967	-
	H24	-	4.035	3.986
H20				
	H25	-	4.270	
	H13	2.077	2.003	2.064
	H10'	2.757	2.647	2.620
	H10"	3.910	3.778	-
	H24	3.332	3.197	3.555
H13				
	H22	4.157	3.917	-
	H20	2.092	2.019	-
	H24	4.240	3.947	3.699
	H21	2.399	2.274	2.489
H10'				
	H23	2.688	-	-
	H20	2.597	-	-
H10"				
	H23	2.897	-	-
	H24	2.737	-	-
H29				
	H24	4.371	3.512	4.334
H24				
	H28	4.085	3.579	-
	H20	3.337	3.169	-
	H13	4.063	3.520	-
	H10'	3.556	3.211	-
	H10''	2.856	2.730	2.905
	H29	-	3.727	-
H21				
	H22	2.335	2.262	-
	H13	2.163	2.093	-

Table S9 - Comparison between distances extracted from 1D and 2D ROESY using a single

point measurement (500 ms as mixing time), and 1D ROESY using 10 points (mixing times from 50 ms to 500 ms) to create build-up curves of Lorlatinib in $CDCl_3$.

From distances' analysis it was clear that all distances extracted from a single 1D ROE experiment with a single mixing time (500 ms) were under-estimated compared with distances resulted from build-up rates arising from the collection of 10 1D spectra for each resonance of interest varying the mixing time from 50 ms to 500 ms. Particularly, distances derived from PANIC method were 3.88% larger on average (ranging from 3.25% to 5.47%) for distances up to 2.5 Å and 8.8% (from 4.13% to 24.45%) for distances greater than 2.5 Å. On the other hand, distances obtained from 2D ROESY were under-estimated in 6 out of 10 cases. The deviation between distances derived from 2D ROESY spectrum with respect to distances obtained through PANIC method was 2.13% (from 0.63% to 3.63%) for distances up to 2.5 Å and 6.02 % for distances exceeding 2.5 Å (ranging from 0.85% to 14.62%).

Proton i	**Proton j**	**δ_i (ppm)**	**δ_j (ppm)**	**Atom$_i$ #**	**Atom$_j$ #**	**σ_{ij}**	**R^2**	**distance r_{ij} (Å)**	**average distance r_{ij} (Å)**
22	20	7.29	6.87	38	34	0.0010	0.903	4.906	
22	13	7.29	5.73	38	33	0.0024	0.945	4.240	
22	8	7.29	4.92	38	48,49	0.0023	0.999	**4.270**	
22	21	7.29	1.8	38	(35,36,37)	0.0356	0.991	2.705	**2.520**
23	10'	6.87	4.45	39	31	0.0234	0.996	2.901	**2.795**
23	10"	7.22	4.37	39	32	0.0147	0.988	3.135	**3.016**
20	25	6.87	7.83	34	43	0.0023	0.977	4.270	ref
20	22	6.87	7.29	34	38	0.0017	0.883	4.491	
20	13	6.87	5.73	34	33	0.1740	0.999	2.077	**2.085**
20	10'	6.87	4.45	34	31	0.0318	0.999	2.757	**2.677**
20	10"	6.87	4.37	34	32	0.0039	0.945	3.910	
20	24	6.87	3.14	34	(40,41,42)	0.0102	0.996	3.332	**3.334**
13	22	5.73	7.29	33	38	0.0027	0.954	**4.157**	-
13	20	5.73	6.87	33	34	0.1665	0.999	2.092	-
13	24	5.73	3.14	33	(40,41,42)	0.0024	0.981	4.240	**4.151**
13	21	5.73	1.8	33	(35,36,37)	0.0733	0.993	2.399	**2.281**
10'	23	4.45	7.22	31	38	0.0370	0.997	2.688	-

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	Atom$_i$ #	Atom$_j$ #	σ_{ij}	R^2	distance r_{ij} (Å)	average distance r_{ij} (Å)
10'	20	4.45	6.87	31	34	0.0455	0.995	2.597	-
10''	23	4.37	6.87	32	39	0.0236	0.993	2.897	-
10''	24	4.37	3.14	32	(40,41,42)	0.0332	0.996	2.737	**2.797**
29	24	4.07	3.14	(45,46,47)	(40,41,42)	0.0020	0.977	**4.371**	
24	28	3.14	7.01	(40,41,42)	44	0.0030	0.992	**4.085**	
24	20	3.14	6.87	(40,41,42)	34	0.0101	0.999	3.337	-
24	13	3.14	5.73	(40,41,42)	33	0.0031	0.956	4.063	-
24	10'	3.14	4.45	(40,41,42)	31	0.0069	0.982	**3.556**	
24	10''	3.14	4.45	(40,41,42)	32	0.0257	0.998	2.856	-
24	29	3.14	4.07	(40,41,42)	(45,46,47)	0.0024	0.942	4.240	
21	22	1.8	7.29	(35,36,37)	38	0.0861	0.993	2.335	-
21	13	1.8	5.73	(35,36,37)	33	0.1362	0.996	2.163	-

Table S10 - Distances of Lorlatinib in $CDCl_3$ obtained from build-up rates. σ_{ij} indicates the slope of the curves obtained through PANIC approach. Average distance indicates the average of distances measurable from each spin involved in the interproton distance. Distances in bold were used for ROE-based analysis.

Interproton distances obtained through PANIC approach were used in StereoFitter analysis. ROE-based conformational analysis was carried out using 11 interproton distances on StereoFitter.

The analysis was conducted without any limitation in the number of conformations that the software could combine to give a solution, hence fitting run until the model combining n+1 conformers resulted in an AIC value greater than the solution with n conformers.

Interproton distances' analysis did not result in an unambiguous solution. In fact, five solutions presented relative probabilities greater than 50%, as shown in Table S12. StereoFitter printed other 5 solutions (data not shown) in which the relativity probability was between 45% and 49%. Since more than one solution has a very low AIC number (between 8.275 and 9.554) the first solution cannot be considered as the best and the only result. Linear regression analyses gave R^2 values of about 0.8. All

these solutions can explain experimental data regarding interproton distances. Hence, the single-parametric fitting failed in providing a result of the time-averaged conformation ensemble in solution.

	Solution	1	2	3	4	5
		Conformer 1: 51.87%	Conformer 2: 40.18%	Conformer 2: 66.31%	Conformer 1: 68.86%	Conformer 2: 42.39%
		Conformer 2: 48.13%	Conformer 3: 35.64%	Conformer 3: 33.69%	Conformer 4: 31.14%	Conformer 3: 34.71%
		-	Conformer 9: 29.18%	-	-	Conformer 4: 22.90%
	AIC	8.275	8.857	8.936	9.392	9.554
	χ^2	6.275	4.857	6.936	7.392	5.554
	Relative probability	1.000	0.748	0.719	0.572	0.528
	NOE distances (Å)					
Atoms (^{1}H-^{1}H)	Experimental	Calculated	Calculated	Calculated	Calculated	Calculated
33,38	4.157	3.729	3.663	3.732	3.727	3.732
(35,36,37),38	2.52	2.838	2.661	2.717	2.878	2.715
32,39	2.795	2.759	2.826	2.854	2.637	2.841
31,39	3.016	3.124	3.073	3.1	2.999	3.087
33,34	2.085	2.115	2.168	2.37	2.156	2.515
32,34	2.677	2.938	2.66	2.628	3.105	2.664
34,(40,41,42)	3.334	3.851	3.815	4.327	3.234	3.674
33,(40,41,42)	4.151	4.648	4.188	4.811	4.578	4.812
31,(40,41,42)	2.797	3.047	3.055	3.077	3.042	3.078
(40,41,42),44	4.085	5.296	5.2	4.972	5.688	4.99
32,(40,41,42)	3.556	3.836	3.826	3.825	3.838	3.823

Table S11 – StereoFitter results of interproton distances' fitting.

5.1.2.3 Analysis through Residual Dipolar Couplings in PMMA gel

A PPMA ((poly(methyl methacrylate)) stick of about 2.5 cm was placed into a high-resolution NMR tube. A solution of 6 mg Lorlatinib in 0.6 mL $CDCl_3$ was added to the gel. Sample was let to swell and equilibrate for 24 hours. After that, it was compressed with the compression apparatus. The alignment degree was verified through the acquisition of deuterium spectra. The quadrupolar splitting of the ^2H signal of the solvent inside the gel was 44.81 Hz. It can be noted that the magnitude of the quadrupolar splitting of this particular gel compatible with $CDCl_3$.

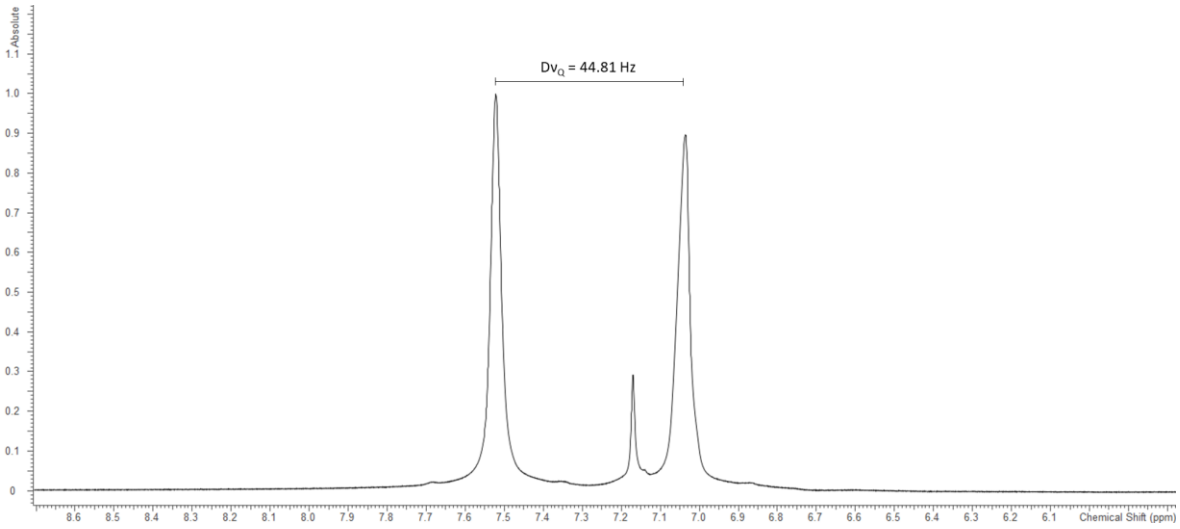


Figure 5.21 – Deuterium spectrum of $CDCl_3$ in poly(methyl methacrylate) gel after compression of the plunger. The quadrupolar splitting (Dv_Q) of the solvent peak after the compression is shown. The less intense peak in at 7.17 ppm comes from the residual solvent peak of the solution remaining between the gel and the wall of the NMR tube.

^1H spectrum acquisition was not useful to assess the diffusion of Lorlatinib inside the gel; for this reason, it was assessed through ^1H-^{13}C HSQC spectrum. After that, HD-J-HSQC was recorded to extract total dipolar couplings.

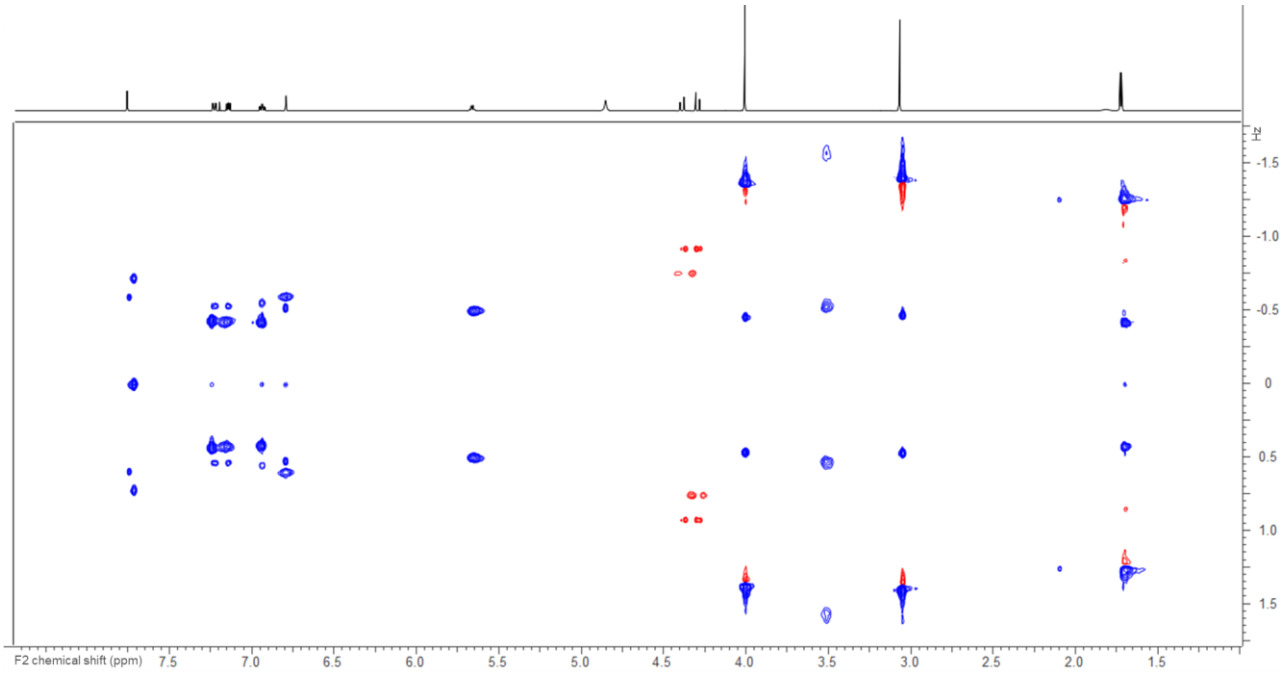


Figure 5.22 – HD-J-HSQC spectrum of Lorlatinib diffused into PMMA gel/$CDCl_3$ recorded to measure total dipolar couplings ($^1T_{CH}$). On the one hand, in the aromatic region isotropic and anisotropic signals can be distinguished (the latter are more intense), on the other hand acquisition of isotropic spectrum of analyte in solution was necessary for measuring aliphatic residual dipolar couplings.

Isotropic and total splittings to obtain RDCs were measured prior in solution and subsequent to alignment in the gel (Table S12).

^{13}C	^1H	δ ^{13}C (ppm)	δ ^1H (ppm)	$^1J_{CH}$ (Hz)	$^1T_{CH}$ (Hz)	$^1D_{CH}$ (Hz)	$^1D_{CH2}$ (Hz)	$^1D_{CH3}$ (Hz)
25	43	136.87	7.83	178.63	217.94	39.31		
22	38	128.12	7.29	163.39	131.15	-32.24		
23	39	128.12	7.22	163.37	128.08	-35.29		
28	44	115.32	7.01	163.51	125.04	-38.47		
20	34	118.53	6.87	158.24	180.71	22.47		
13	33	71.44	5.73	150.1	151.04	0.94		
10	31,32	47.47	4.45, 4.37	277.34	227.8	-49.54	-24.7	
29	45,46,47	38.6	4.07	425.87	413.5	-12.37		-4.12
24	40,41,42	31.55	3.14	416.63	422.71	6.08		2.03
21	35,36,37	22.48	1.8	386.11	378.81	-7.3		-2.43

Table S12 – Experimental $^1J_{CH}$ and $^1T_{CH}$ values extracted from an HD-J-HSQC spectrum. Dipolar couplings are calculated according to the equation $^1D_{CH} = {}^1T_{CH} - {}^1J_{CH}$. $^1D_{CH2}$

represents the value of the residual dipolar coupling extracted from HD-J-HSQC spectrum divided by 2 since from that spectra only the sum of the couplings between carbon atom and two diastereotopic protons can be obtained, whereas $^{1}D_{CH3}$ *represents the value of the residual dipolar coupling divided by 3 for the methyl group.*

A total of 10 Residual Dipolar Couplings were used in StereoFitter. The analysis was conducted without any restriction on the maximum number of conformers per solution. Levenberg-Marquardt algorithm was used as least squares method. As it is shown in Table S13, at least 5 solutions can be identified as possible results because of the value of the normalized relative probability. Moreover, all correlation plots of back-calculated against experimental Residual Dipolar Couplings showed almost the same coefficient of determination value (see Figure 5.23).

	Solution	**1**	**2**	**3**	**4**	**5**	**6**
		Conformer 1: 22.99%	Conformer 1: 20.98%	Conformer 5: 72.11%	Conformer 5: 36.65%	Conformer 2: 15.97%	Conformer 3: 13.92%
		Conformer 5: 77.01%	Conformer 4: 40.46%	Conformer 11: 14.94%	Conformer 6: 45.00%	Conformer 5: 84.03%	Conformer 5: 51.42%
		-	Conformer 10: 38.56%	Conformer 12: 12.95%	Conformer 13: 18.35%	-	Conformer 6: 34.66%
	AIC	4.898	5.113	5.354	5.615	5.791	6.233
	χ^2	2.898	1.113	1.354	1.615	3.791	2.233
	Relative probability	1.000	0.898	0.796	0.699	0.640	0.513
	RDCs (Hz)						
Atoms (^{13}C-^{1}H)	**Experimental**	**Calculated**	**Calculated**	**Calculated**	**Calculated**	**Calculated**	**Calculated**
(10,31) (10,32)	-24.7	-24.715	-24.709	-25.12	-24.909	-25.674	-25.184
13,33	0.94	0.753	1.123	0.758	1.078	0.983	1.123
20,34	22.47	22.197	22.412	22.357	22.538	22.398	22.576
(21,35) (21,36) (21,37)	-2.43	-1.672	-2.47	-2.07	-2.665	-1.798	-2.594
22,38	-32.24	-33.116	-32.642	-32.498	-33.269	-33.021	-33.357
23,39	-35.29	-34.415	-34.701	-34.557	-34.252	-34.249	-34.057
(24,40) (24,41) (24,42)	2.027	0.799	1.903	1.108	2.126	0.515	2.186
25,43	39.31	39.418	39.628	39.574	39.227	39.095	39.124
28,44	-38.47	-38.655	-38.438	-38.507	-38.366	-38.413	-38.348
(29,45) (29,46) (29,47)	-4.12	-3.486	-3.151	-3.845	-4.317	-3.846	-4.38

Table S13 – StereoFitter results of RDCs fitting of Lorlatinib in $CDCl_3$*/PMMA. Back-calculated Residual Dipolar Couplings resulted from the orientation predicted by StereoFitter through SVD method.*

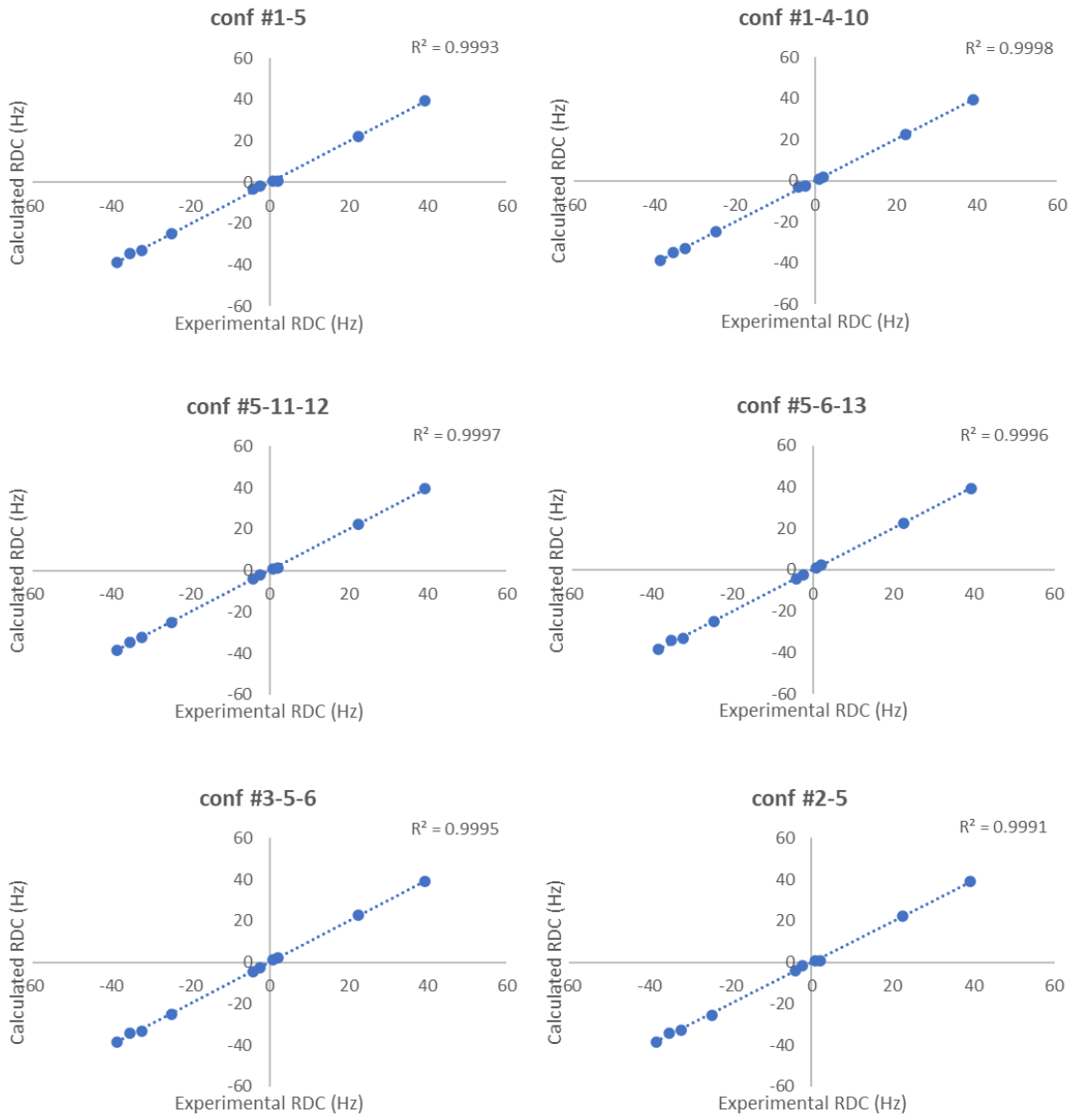


Figure 5.23 - Correlation plot of the 10 CH theoretical RDCs versus experimental RDCs provided in Table S13.

Since both distances-based analysis and RDCs-based analysis didn't result in a unique solution, multiparametric approach was performed.

A total of 10 constraints coming from Residual Dipolar Couplings and 11 constraints coming from interproton distances were used for multiparametric fitting in StereoFitter. Interestingly, this analysis led to two likely solutions which share the most populated conformations (conformer 1 and conformer 9) and differ for the least populated conformation (conformer 4) which only represents 4.67% of the population.

	Solution	1	2
		Conformer 1: 79.41%	Conformer 1: 74.98%
		Conformer 9: 20.59%	Conformer 9: 20.35%
		-	Conformer 4: 4.67%
	AIC	16.815	18.023
	χ^2	14.815	16.815
	Relative probability	1	0.547
	NOE distances (Å)		
Atoms (^{1}H-^{1}H)	**Experimental**	**Calculated**	**Calculated**
33,38	4.157	3.668	3.667
(35,36,37),38	2.52	2.836	2.846
32,39	2.795	2.599	2.575
31,39	3.016	2.96	2.935
33,34	2.085	1.999	1.986
32,34	2.677	3.077	3.06
34,(40,41,42)	3.334	3.385	3.439
33,(40,41,42)	4.151	4.101	4.084
31,(40,41,42)	2.797	3.019	3.017
(40,41,42),44	4.085	6.179	6.355
32,(40,41,42)	3.556	3.843	3.844
	RDCs (Hz)		
Atoms (^{13}C-^{1}H)	**Experimental**	**Calculated**	**Calculated**
(10,31) (10,32)	-24.7	-23.775	-23.989
13,33	0.94	-0.137	-0.103
20,34	22.47	22.119	21.878
(21,35) (21,36) (21,37)	-2.43	-1.953	-1.682
22,38	-32.24	-33.392	-33.347
23,39	-35.29	-34.596	-34.593
(24,40) (24,41) (24,42)	2.027	0.063	0.021
25,43	39.31	39.228	39.093
28,44	-38.47	-38.85	-39.11
(29,45) (29,46) (29,47)	-4.12	-5.043	-4.527

Table S14 – StereoFitter results of multiparametric fitting of Lorlatinib in $CDCl_3$/PMMA. All experimental Residual Dipolar Couplings and interproton distances were used.

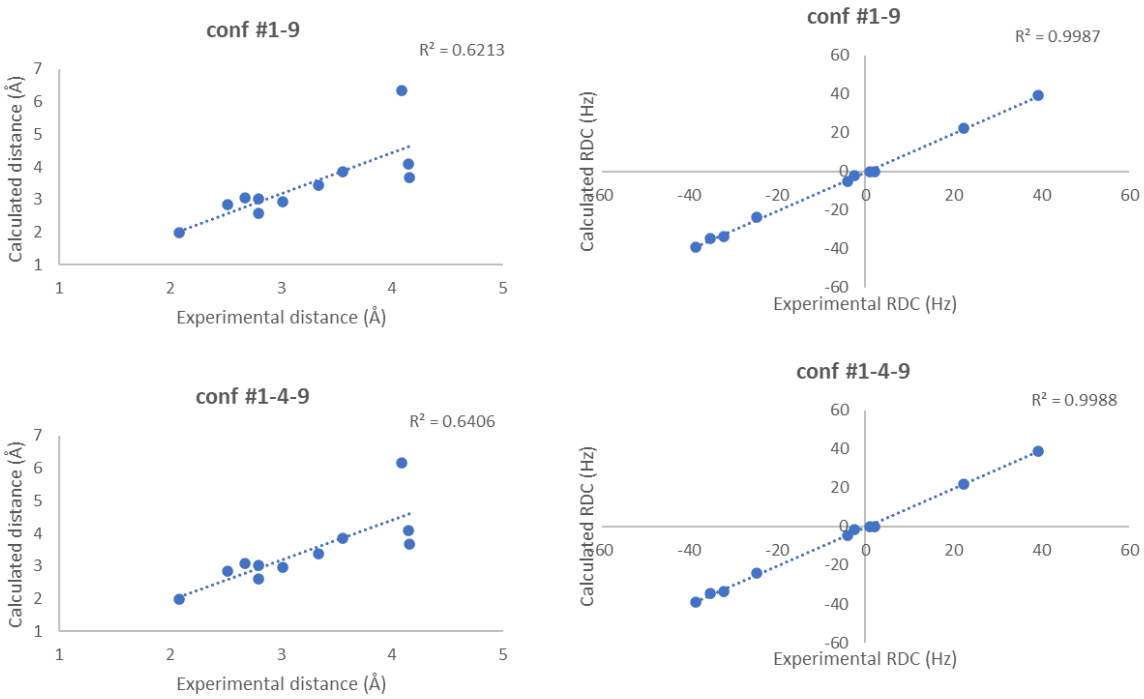


Figure 5.24 – Linear regressions of results provided in Table S14. Both plots of calculated distances versus experimental distances highlight the presence of some outliers leading to R^2 values of 0.6.

3D structures in $CDCl_3$ represented in Figure 5.25 show that carbonyl group of conformer 9 could be shielded by the pyridine ring, and hence less exposed to the solvent when compared to the carbonyl group of conformer 1. This aspect could explain the high permeability of Lorlatinib in apolar environments.

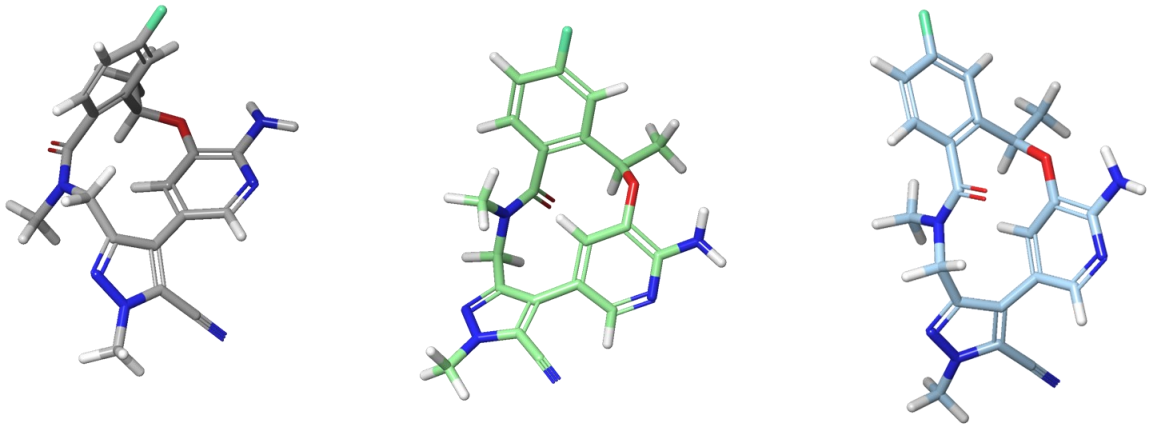


Figure 5.25 – Solution ensemble of Lorlatinib in PMMA/$CDCl_3$. Conformer #1 (grey), conformer #4 (light green), conformer #9 (light blue).

Multiparametric fitting was also performed using a set of interproton distances without the distance between the aromatic proton H_{44} (belonging to the fluorobenzene) and methyl group $Me_{40,41,42}$ since it was the outlier with the biggest error in both solution (see Figure 5.24). The hypothesis was that the distance under discussion was affected by experimental errors somehow. This analysis resulted in the same solutions shown in Table S14, demonstrating that this single distance didn't affect the reliability of the previous results.

In conclusion, single-parametric fitting of distances and Residual Dipolar Couplings led to various likely solutions of conformation of Lorlatinib in chloroform. A unique solution was obtained only integrating information coming from both local constraints, i.e. interproton distances, and global constraints, i.e. Residual Dipolar Couplings.

5.1.2.4 Analysis through Residual Dipolar Couplings in LLC phase

Lorlatinib in $CDCl_3$ was tested in a different alignment medium than the previous analysis. It was aligned in lyotropic liquid crystalline phase.

Poly-γ-p-biphenylmethyl-L-glutamate (PBPMLG)[48] was placed into an NMR tube, then analyte and chloroform were added to obtain a weight-percent concentration (w/w) of 9.9%. At this concentration the polymer reaches the critical concentration to form the LC phase. A capillary filled with acetone-d_6 was put into the tube to facilitate the lock. The sample was homogenized through centrifugation (up and down at about 1000 rpm for 10 times). The homogeneity and shimming of the lyotropic liquid crystals were checked by 2H-IMAGE pulse sequence. The quadrupolar splitting of the deuterium was assessed prior to seal the NMR tube in order not to alter the concentration.

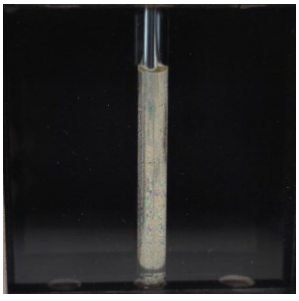


Figure 5.26 – LLC phase formation checked using polarizer filter.

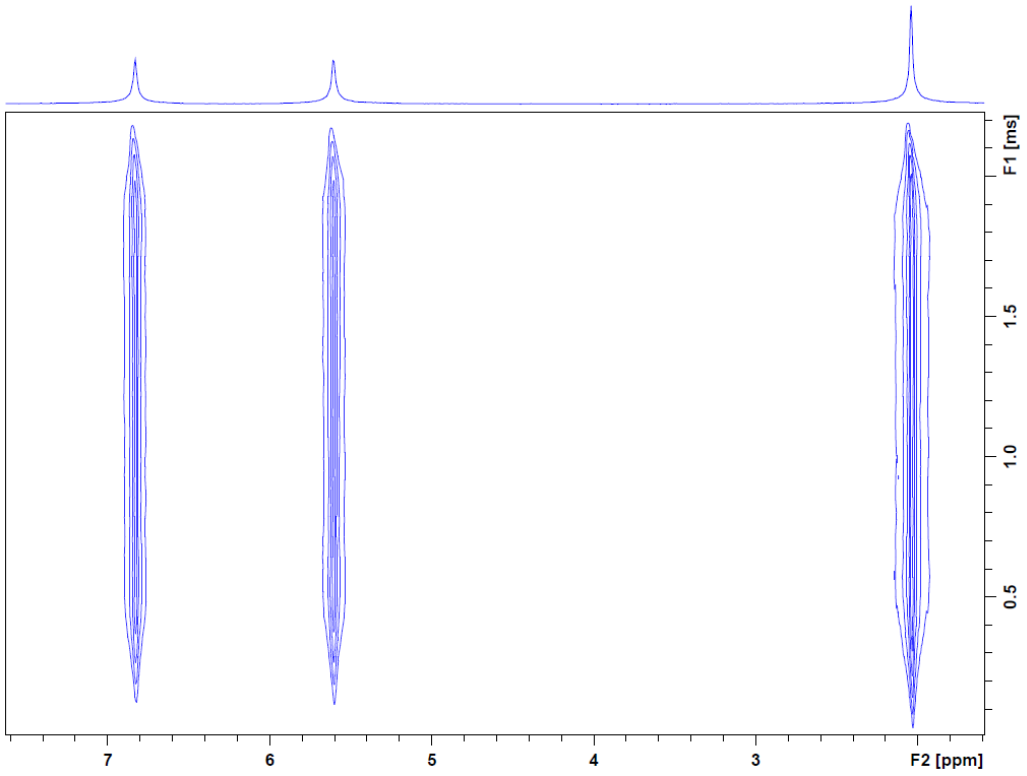


Figure 5.27 – 2H-IMAGE spectrum of Lorlatinib in PBPMLG/$CDCl_3$. Signals at 5.6 and 6.82 ppm represent the quadrupolar splitting of deuterium of $CDCl_3$. Signal at 2.03 ppm is deuterium signal of acetone-d_6. Since any tilt is present neither in the upper part or at the bottom of the signal, the sample is homogenous and well shimmed.

After assessing the homogeneity of the sample, ^{2}H NMR spectrum was recorded to measure the quadrupolar splitting of the solvent.

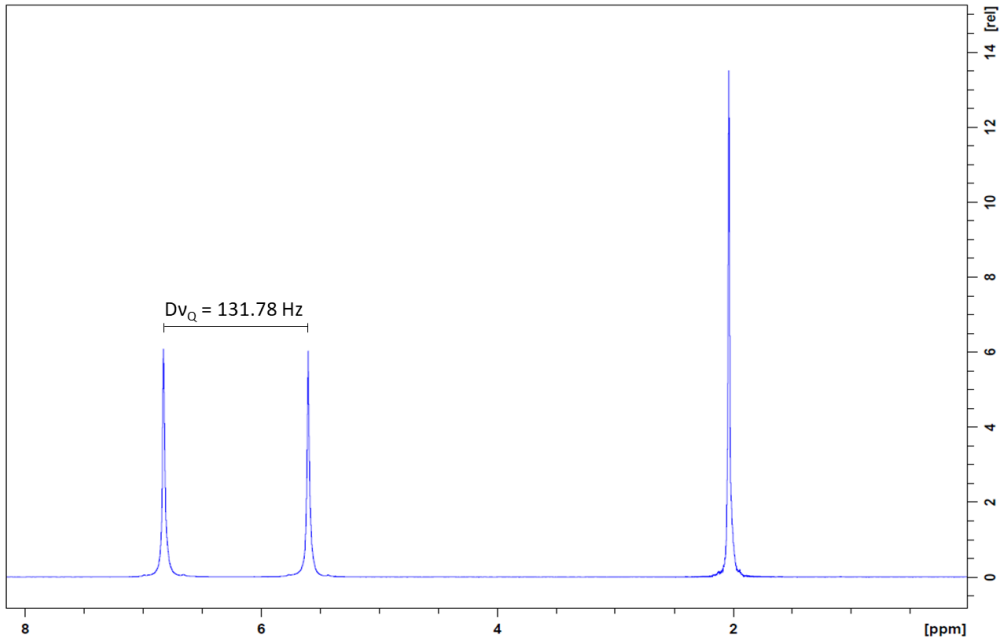


Figure 5.28 – ^{2}H NMR spectrum (107.49 MHz, ^{1}H NMR frequency 700 MHz) of the LLC

phase sample in $CDCl_3$. A ^2H quadrupolar splitting of 131.78 Hz for $CDCl_3$ was obtained.

CLIP-HSQC spectrum was acquired to extract total dipolar couplings since only anisotropic signals are detectable in lyotropic liquid crystalline phases; in fact, isotropic couplings were previously measured through a CLIP-HSQC spectrum recorded in solution.

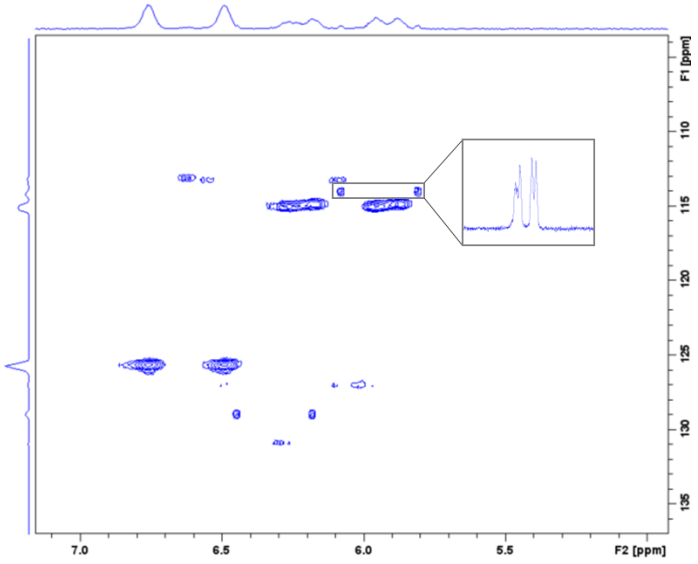


Figure 5.29 – CLIP-HSQC spectrum of Lorlatinib in PBPMLG/$CDCl_3$. Signals from aromatic region are shown. In the inset it can be noted that the spectra quality is not high. In fact, since $^1T_{CH}$ splitting is not well resolved as a doublet, the uncertainty regarding the accuracy of the measured value of the coupling.

Total and isotropic dipolar couplings were extracted and fitting was performed using both RDC@hotFCHT (software developed by V. Schmidts, R. Berger, C. M. Thiele)[28,29] and StereoFitter software. RDC@hotFCHT performs two analyses in parallel: single-conformer single-tensor fit and multi-conformer single-tensor fit.

^{13}C	^1H	δ ^{13}C (ppm)	δ ^1H (ppm)	$^1J_{CH}$ (Hz)	$^1T_{CH}$ (Hz)	$^1D_{CH}$ (Hz)	$^1D_{CH3}$ (Hz)	$^1D_{CC/CN}$ (Hz)	error
25	43	136.87	7.851	178.540	186.384	7.844			0.178
23	39	128.12	7.229	161.409	187.055	25.646			0.656
28	44	115.41	7.024	166.052	210.648	44.596			0.183
20	34	118.53	6.882	156.727	188.732	32.004			1.374
13	33	71.44	5.751	149.731	155.603	5.872			0.459
10	31	47.47	4.476	137.127	177.342	40.215			0.170

^{13}C	^{1}H	δ ^{13}C (ppm)	δ ^{1}H (ppm)	$^{1}J_{CH}$ (Hz)	$^{1}T_{CH}$ (Hz)	$^{1}D_{CH}$ (Hz)	$^{1}D_{CH3}$ (Hz)	$^{1}D_{CC/CN}$ (Hz)	error
10	32	47.47	4.378	142.389	176.993	34.603			0.198
29	45,46,46	38.74	4.093	142.052	102.305	-39.747	-13.249	-1.699	0.155
24	40,41,42	31.55	3.154	139.477	159.121	19.644	6.548	0.839	0.164
21	35,36,37	22.48	1.81	128.732	138.163	9.431	3.144	-2.699	0.384

Table S15. Experimental $^{1}J_{CH}$ and $^{1}T_{CH}$ values extracted from an HD-J-HSQC spectrum. Dipolar couplings are calculated according to the equation $^{1}D_{CH} = {}^{1}T_{CH} - {}^{1}J_{CH}$. $^{1}D_{CH3}$ represents the value of the residual dipolar coupling $^{1}D_{CH}$ divided by 3 for the methyl group for use in StereoFitter, whereas $^{1}D_{CC/CN}$ indicates D_{CH}-couplings of methyl groups converted to the corresponding D_{CC}-couplings or D_{CN}-couplings according to Griesinger formula.[87]

A total of 3 long-range $^{n}D_{CH}$ were measured in HECADE spectrum in solution, whereas only one long-range $^{n}D_{CH}$ could be extracted from the same spectrum in anisotropic conditions. This coupling wasn't used in the following RDCs' fittings because of the poor quality of the spectrum.

Results of single-conformer single-tensor fit using 8 RDCs are shown in Table S16. Analyzing χ^2 and quality factor by Cornilescu et al. (Q-factor), the first four solutions can be considered as probable solution. Anyway, the condition number of the cosine matrix built to calculate the alignment tensor (referred to as SVD # in Table S16) was reasonable. In fact, for macrocyclic compounds, the condition number should be less than 10.[24]

Conf#	χ^2	Q-factor	R^2	SVD #
13	7.77	0.081	0.993	9.054
9	8.054	0.082	0.994	16.263
7	9.154	0.088	0.991	36.676
10	10.502	0.094	0.991	23.282
12	14.039	0.109	0.99	15.072
14	49.678	0.205	0.969	42.374
2	52.483	0.21	0.961	13.212
6	54.757	0.215	0.96	20.933
1	72.198	0.247	0.951	16.933
8	90.388	0.276	0.924	15.378
11	150.307	0.356	0.899	14.078
3	163.519	0.372	0.859	32.733

Table S16 – Single-conformer single-tensor fit of Residual Dipolar Couplings obtained for Lorlatinib in PBPMLG/$CDCl_3$.

Alignment tensors can be represented by Sanson-Flamsteed maps showing the orientation of the principal axes.

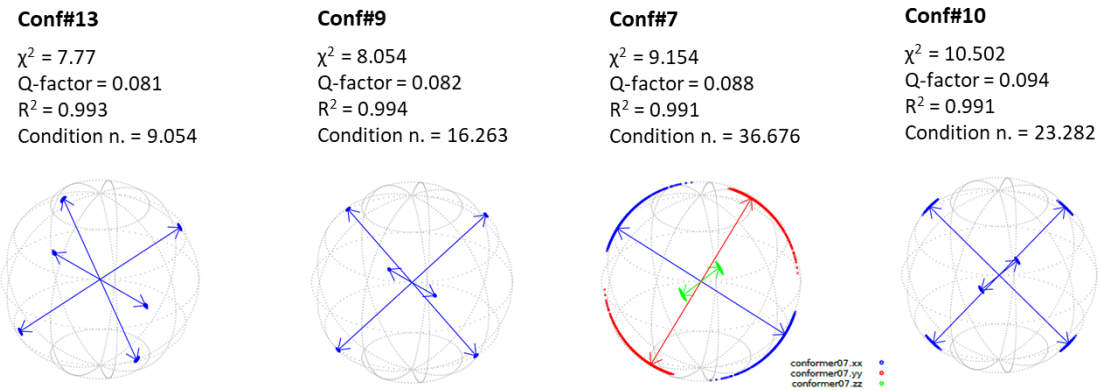


Figure 5.30 - Alignment tensor and quality factors for the different conformers are given: principal axes of the alignment tensor (Axx, Ayy, Azz), quality factors χ^2, determination factor (R^2) and quality factor by Cornilescu.

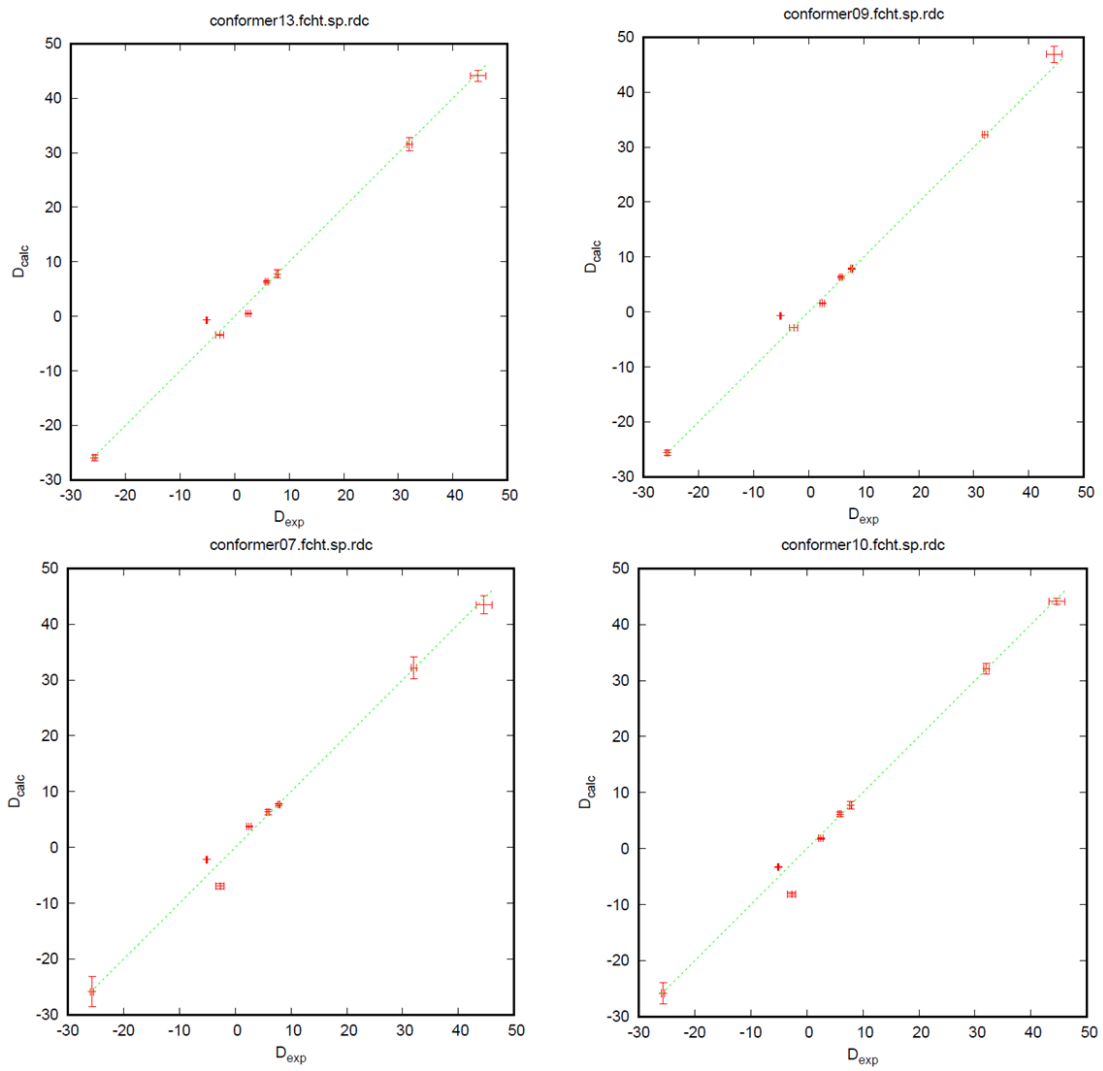


Figure 5.31 – Plots between calculated and observed RDCs of Lorlatinib in PBPMLG/$CDCl_3$ from Table S16. Error bars of both calculated and experimental values are shown.

Multi-conformer single-tensor fit was then performed using stepsize method in which the full population space is sampled recursively and the stepsize defines the percentage change from one conformation to another one within the conformational ensemble. It was set to 20%. This analysis resulted in 8568 combinations that were analysed considering quality factors and condition numbers. Among 8568 combinations, 43 solutions presented SVD values smaller than 10; 7 solutions with lowest values of χ^2 and Q-factor were selected (see Table S17). In particular, solution #15 which combines conformer 9 (20%) with conformer 13 (80%) and solution #23 which combines conformer 11 (20%) and conformer 13 (80%) presented the lowest χ^2 and Q-factor.

Solution	conf #	conf %	χ^2	Q-factor	R^2	SVD #
sol5	conf_10	20%	6.992	0.0768	0.9935	8.2558
	conf_13	80%				
sol11	conf_7	20%	7.47	0.0794	0.9929	8.8508
	conf_13	80%				
sol13	conf_13	100%	7.77	0.081	0.9927	9.0537
sol15	conf_9	20%	6.464	0.0739	0.9941	9.1928
	conf_13	80%				
sol18	conf_3	20%	9.875	0.0913	0.9907	9.4679
	conf_9	60%				
	conf_14	20%				
sol23	conf_11	20%	5.856	0.0703	0.9948	9.8374
	conf_13	80%				
sol27	conf_4	40%	8.563	0.085	0.9932	9.9521
	conf_9	40%				
	conf_11	20%				

Table S17 – Results of multi-conformer single-tensor fit obtained from RDC@hotFCHT.

Residual Dipolar Couplings fitting of experimental measurements obtained through liquid crystals was also performed by means of StereoFitter. Two likely solutions were obtained, the first one averaging two conformers, the other one with averaging three conformers. It is important to note that the two solutions share the two most populated conformers and that the third conformation (conformer #2) of the second solution represents 2.74% of the population only. Hence, solution #1 can be considered as the most likely.

	Solution	1	2
		Conformer 9: 29.85%	Conformer 9: 28.87%
		Conformer 13: 70.15%	Conformer 13: 68.39%
		-	Conformer 12: 2.74%
	AIC	2.511	4.522
	χ^2	0.511	0.522
	Relative probability	1.000	0.366
	RDCs (Hz)		
Atoms (^{13}C-^{1}H)	**Experimental**	**Calculated**	**Calculated**
13,33	5.872	6.119	6.017
20,34	32.005	31.76	31.902
(21,35) (21,36) (21,37)	3.144	3.679	3.265
23,39	-25.646	-25.909	-25.808
(24,40) (24,41) (24,42)	6.548	6.574	6.669
25,43	7.844	7.727	7.773
28,44	44.596	43.165	43.348
(29,45) (29,46) (29,47)	-13.249	-13.053	-13.125

Table S18 – StereoFitter results of RDCs' fitting of Lorlatinib in PBPMLG/$CDCl_3$.

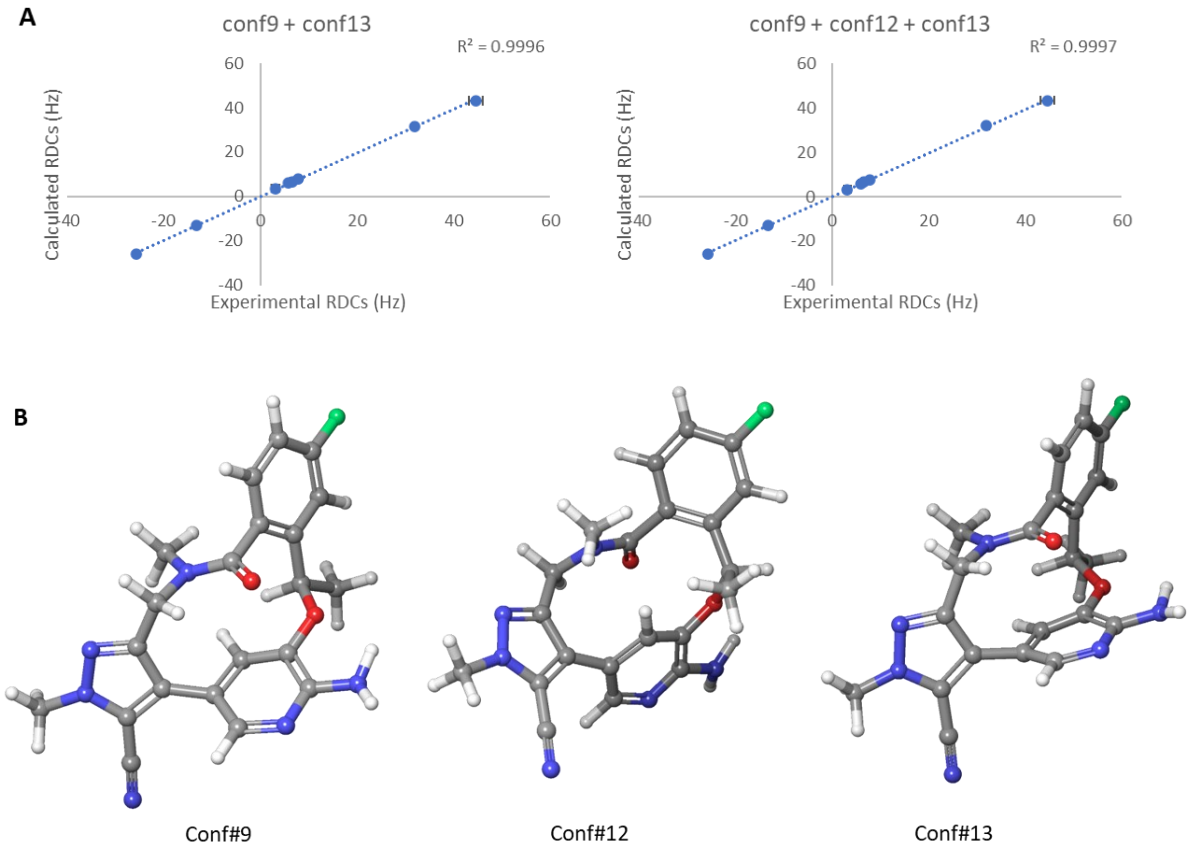


Figure 5.32 – A) Calculated vs experimental RDCs obtained from StereoFitter fitting provided in Table S18. B) Conformations 9 – 12 – 13 as solution of conformational ensemble of Lorlatinib in PBPMLG/$CDCl_3$.

Overall, solution ensembles obtained in PMMA/$CDCl_3$ (see Figure 5.25 – chapter 5.1.2.3) and PBPMLG/$CDCl_3$ share conformation 9, which represents about 20-30% of the conformational ensemble. On the other hand, the remaining two conformations found in liquid crystals both presented trans amide, contrary to the results of the analysis carried out in polymeric gel, in which the amide was cis in conformer #1.

A study on Lorlatinib by means of solid-state NMR has been recently published by Rehman et al.[88] Through the NMR crystallography analysis they found out that two molecules were present in a crystalline unit cell and three out of four amine protons were involved in intramolecular hydrogen bonds; they were able to support this conclusion exploiting quantum-mechanical calculations of chemical shifts compared to experimental ones. Since these conclusions are deduced from a different state than the one used for the conformational analysis explained above, they can add some information to the analysis. Conformer obtained from solid-state NMR analysis and

the one resulted from NMR-based analysis in solution state (in DMSO-d_6) resulted to be nearly superimposable (see Figure 5.33).

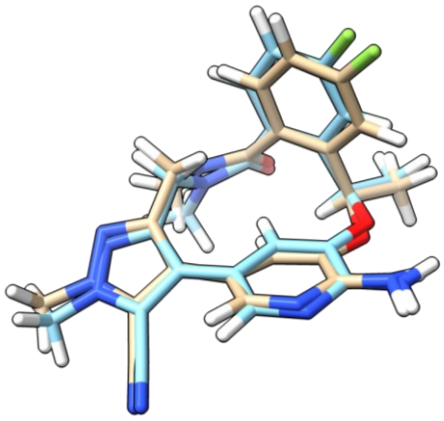


Figure 5.33 – Superposition of conformer #1 (beige) vs conformer (light blue) obtained through NMR crystallography by Rehman et al.[88]

5.2 Conformational analysis of Pacritinib

Pacritinib, ((16*E*)-11-(2-pyrrolidin-1-ylethoxy)-14,19-dioxa-5,7,27-triazatetracyclo [19.3.1.1^{2,6}.1^{8,12}]heptacosa-1(24),2(27),3,5,8(26),9,11,16,21(25),22-decaene), is a potent inhibitor of wild-type JAK2 (IC50=23 nM) and JAK2^{V617F} mutant (IC50=19 nM). It was approved by Food and Drug Administration in 2022 to treat myelofibrosis.[89]

5.2.1. Conformational search

The conformational search was performed through Macrocycle Conformational Sampling[90] tool in MacroModel[65]. It comprised a short (0.5 ps) stochastic dynamics simulation using high temperature, a short simulated annealing step during which the temperature is lowered in 0.5 ps, an energy minimization. The maximum distance for low-mode move was set to 18 Å, RMSD cut-off was set to 0.75 Å and energy window for saving structures to 10 kcal/mol. Force field S-OLPS and 4rdd solvation model were used.

More than 500 conformations were found; their geometry was then optimized through quantum mechanics using DFT/B3LYP theory coupled with 6-31G+* basis set. After this step conformers within 5 kcal/mol electronic energy were analysed. They were divided into 30 clusters superposing the macrocyclic core. All conformers superimposable for the macrocyclic core but different among the flexible side-chain leading to the pyrrolidine were discarded. A total of 63 conformation with different dihedral in the macrocyclic core were selected for the following analyses. Even if macrocyclic compounds are considered to be quite rigid, Pacritinib exhibits a certain degree of flexibility since it has 11 rotatable bonds in the macrocyclic core; in fact, RMSD of the macrocyclic core of 63 conformations was between 0.6 and 3.1 Å.

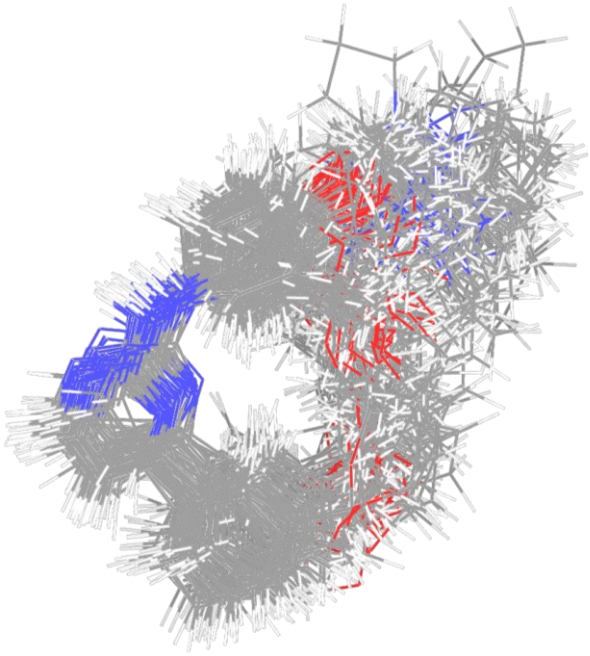


Figure 5.34 – Superposition of 63 conformations obtained through Macrocycle Conformational Sampling.

5.2.2. NMR assignment

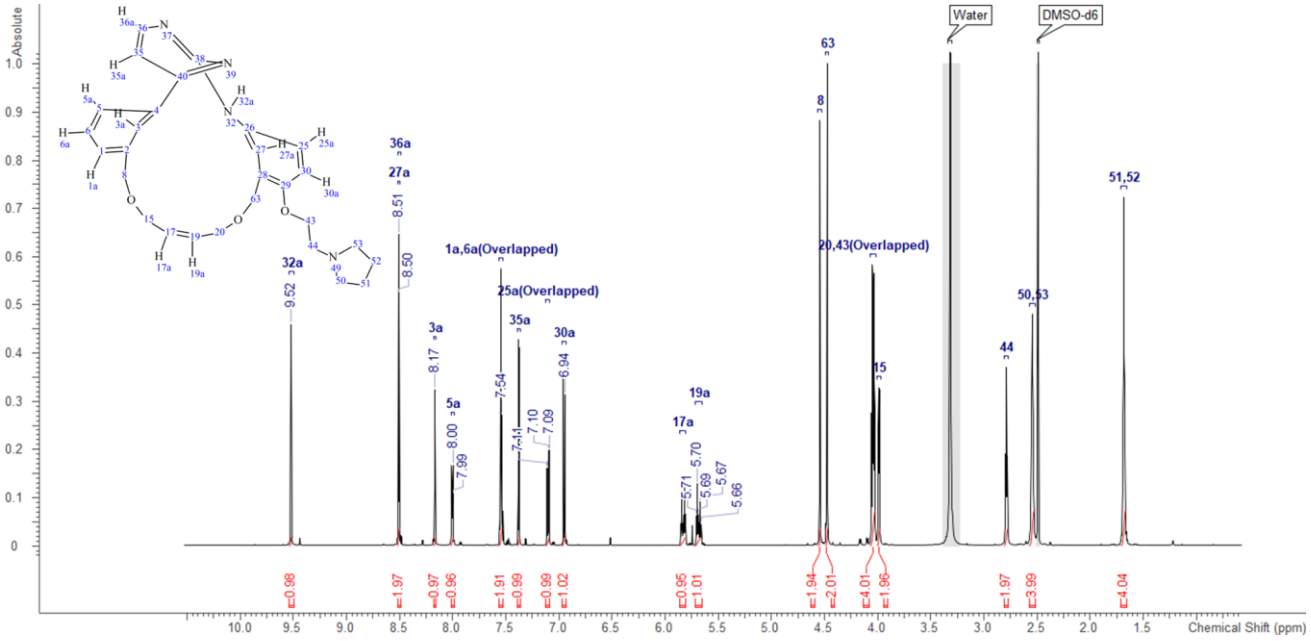


Figure 5.35 – ^{1}H NMR spectrum of 5 mg Pacritinib in 0.6 mL DMSO-d_6.

NMR assignments were performed through acquisition and analysis of ^{1}H, COSY, TOCSY, ^{1}H-^{13}C HSQC, ROESY, ^{1}H-^{13}C HMBC, ^{1}H-^{15}N HMBC spectra on 700 MHz spectrometer Bruker AVANCE III HD.

No.	F2 atom (^{1}H)	F1 atom (^{13}C)	F2 (ppm)	F1 (ppm)
1	33	27	8.52	120.34
2	42	36	8.51	159.06
3	9	3	8.18	126.24
4	10	5	8.01	126.24
5	12	1	7.55	128.86
6	11	6	7.55	130.54
7	41	35	7.39	107.02
8	31	25	7.11	119.37
9	66	30	6.96	112.5
10	24	17	5.84	129.57
11	23	19	5.69	130.8
12	12,13	8	4.55	68.67
13	64,65	63	4.48	65.01
14	45,46	43	4.05	67.79
15	21,22	20	4.04	68.98
16	16,18	15	3.99	66.85
17	47,48	44	2.79	54.24
18	54,55,60,61	50,53	2.55	53.93
19	56,57,58,59	51,52	1.68	22.99

Table S19 – NMR assignments of Pacritinib in DMSO-d_6.

5.2.3. Analysis by means of interproton distances

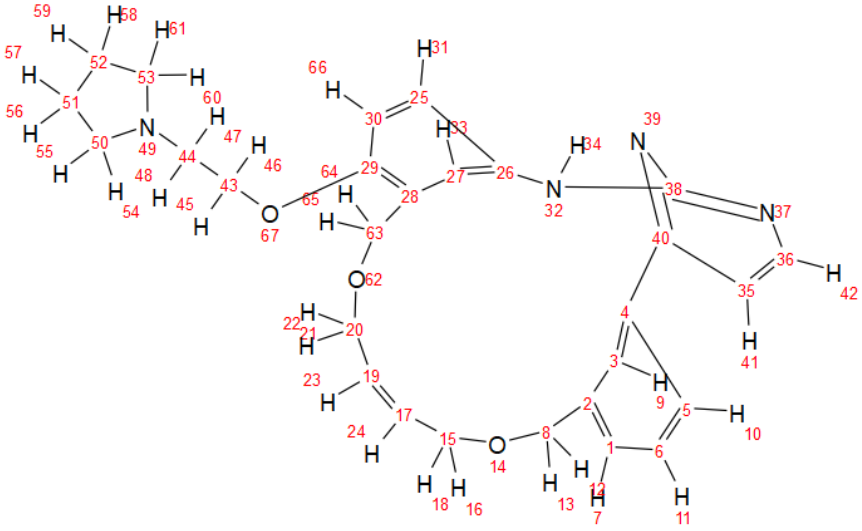


Figure 5.36 – 2D structure of Pacritinib. Atom numbering according to Maestro and StereoFitter.

Interproton distances were calculated using PANIC approach after acquisition of 10 spectra with 10 different mixing times (between 50 ms to 500 ms) for each irradiated proton.

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	Atom # i	Atom # j	σ_{ij}	R^2	distance r_{ij} (Å)	average distance (Å)
H3	H27	8.18	8.52	9	33	0.0895	0.999	**2.495**	
H3	H35	8.18	7.39	9	41	0.0023	0.967	4.592	**4.710**
H3	H17	8.18	5.84	9	24	0.0066	0.992	3.852	**3.906**
H3	H19	8.18	5.69	9	23	0.0131	0.997	3.436	**3.412**
H3	H8	8.18	4.55	9	12,13	0.0661	0.988	2.624	
H3	H63	8.18	4.48	9	64,65	0.0125	0.973	**3.463**	
H3	H15	8.18	3.99	9	16,18	0.0034	0.976	**4.302**	
H5	H35	7.55	7.39	10	41	0.1614	0.996	2.262	**2.295**
H35	H36	7.39	8.51	41	42	0.0804	0.998	2.540	ref
H35	H3	7.39	8.18	41	9	0.0017	0.953	4.829	
H35	H5	7.39	7.55	41	10	0.1354	0.999	2.329	
H30	H25	6.96	7.11	66	31	0.0855	0.992	2.514	
H30	H63	6.96	4.48	66	64,65	0.0015	0.958	**4.931**	
H30	H43	6.96	4.05	66	45,46	0.1406	0.996	2.314	
H17	H27	5.84	8.52	24	33	0.0028	0.961	**4.444**	
H17	H3	5.84	8.18	24	9	0.0056	0.975	3.959	
H17	H8	5.84	4.55	24	12,13	0.0174	0.985	**3.278**	
H17	H63	5.84	4.48	24	64,65	0.0133	0.992	**3.428**	
H17	H15	5.84	3.99	24	16,18	0.056	0.992	**2.698**	

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	Atom # i	Atom # j	σ_{ij}	R^2	distance r_{ij} (Å)	average distance (Å)
H17	H20	5.84	4.04	24	21,22	0.0557	0.989	**2.700**	
H19	H27	5.69	8.52	23	33	0.0082	0.988	**3.715**	
H19	H3	5.69	8.18	23	9	0.0143	0.991	3.387	
H19	H8	5.69	4.55	23	12,13	0.016	0.987	**3.324**	
H19	H63	5.69	4.48	23	64,65	0.0155	0.987	**3.341**	
H19	H20	5.69	4.04	23	21,22	0.0547	0.988	**2.708**	
H19	H15	5.69	3.99	23	16,18	0.0536	0.989	**2.718**	
H8	H3	4.55	8.18	12,13	9	0.1545	0.990	2.278	
H8	H17	4.55	5.84	12,13	24	0.0368	0.972	2.893	
H8	H19	4.55	5.69	12,13	23	0.0315	0.977	2.969	
H8	H15	4.55	3.99	12,13	16,18	0.0817	0.999	2.533	
H63	H27	4.48	8.52	64,65	33	0.0806	0.996	2.539	
H63	H3	4.48	8.18	64,65	9	0.0193	0.995	3.222	
H63	H17	4.48	5.84	64,65	24	0.0223	0.997	3.145	
H63	H19	4.48	5.69	64,65	23	0.0203	0.986	3.195	
H63	H20	4.48	4.04	64,65	21,22	0.0989	0.998	2.454	
H63	H43	4.48	4.05	64,65	45,46	0.0030	0.968	4.393	
H15	H3	3.99	8.18	16,18	9	0.0082	0.945	3.715	
H15	H17	3.99	5.84	16,18	24	0.1273	0.994	2.353	
H15	H19	3.99	5.69	16,18	23	0.1073	0.993	2.421	
H15	H8	3.99	4.55	16,18	12,13	0.0776	0.996	2.555	
H15	H63	3.99	4.48	16,18	64,65	0.0025	0.967	4.528	

Table S20 – Interproton distances of Pacritinib in DMSO-d_6. PANIC approach was employed; measurements in bold were used for fitting analysis.

A total of 18 proton-proton distances were used in StereoFitter input file. Some experimental distances were not included in the analysis because they were considered unnecessary for the conformation (e.g., distances between aromatic proton 9 and methylene 12-13, 66 and methylene 45-46, 33 and methylene 64-65). StereoFitter analysis resulted in 10 solutions with relative probability greater than 0.9. Even if the gain in AIC value between the first solution (comprising two conformation) and the other solutions (comprising three conformations) is very low, there is no confidence that the first solution is the only one.

	Solution	1	2	3	4	5	6
		Conformer 6: 44.51%	Conformer 6: 45.88%	Conformer 6: 46.01%	Conformer 6: 45.24%	Conformer 6: 47.94%	Conformer 6: 48.10%
		Conformer 22: 55.49%	Conformer 22: 48.62%	Conformer 22: 46.43%	Conformer 22: 50.14%	Conformer 18: 44.26%	Conformer 18: 46.29%
		-	Conformer 46: 5.49%	Conformer 1: 7.57%	Conformer 52: 4.62%	Conformer 1: 7.80%	Conformer 46: 5.61%
	AIC	5.222	5.231	5.250	5.276	5.300	5.334
	χ^2	3.222	1.231	1.250	1.276	1.300	1.334
	Relative probability	1	0.995	0.986	0.973	0.962	0.945
	NOE distances (Å)						
Atoms (^{13}C-^{1}H)	**Experimental**	**Calculated**	**Calculated**	**Calculated**	**Calculated**	**Calculated**	**Calculated**
9,33	2.495	2.492	2.49	2.488	2.49	2.441	2.44
9,41	4.71	4.643	4.631	4.63	4.635	4.624	4.624
9,24	3.906	4.962	3.824	3.963	3.889	3.943	3.809
9,23	3.412	3.435	3.49	3.511	3.478	3.493	3.472
9,(64,65)	3.463	3.565	3.578	3.583	3.577	3.554	3.547
9,(16,18)	4.302	4.266	4.267	4.267	4.275	4.183	4.182
10,41	2.295	2.184	2.197	2.197	2.194	2.211	2.211
(64,65),66	4.931	4.935	4.949	4.95	4.94	4.961	4.962
24,33	4.444	5.519	4.696	4.47	4.431	4.441	4.668
(12,13),24	3.278	3.46	3.312	3.283	3.414	3.296	3.329
24,(64,65)	3.428	3.438	3.431	3.317	3.379	3.317	3.437
(16,18),24	2.698	2.627	2.624	2.622	2.628	2.62	2.622
(21,22),24	2.7	2.676	2.676	2.704	2.693	2.705	2.677
23,33	3.715	3.564	3.618	3.64	3.608	3.588	3.566
(12,13),23	3.324	3.358	3.364	3.367	3.357	3.372	3.368
23,(64,65)	3.341	3.469	3.458	3.469	3.477	3.449	3.436
(21,22),23	2.708	2.722	2.722	2.706	2.713	2.708	2.725
(16,18),23	2.718	2.861	2.869	2.869	2.868	2.875	2.876

Table S21 – StereoFitter results of interproton distances' fitting.

Coefficient determination goes from 0.88 for the solution which combines two conformers to 0.98 for solutions combining three conformers.

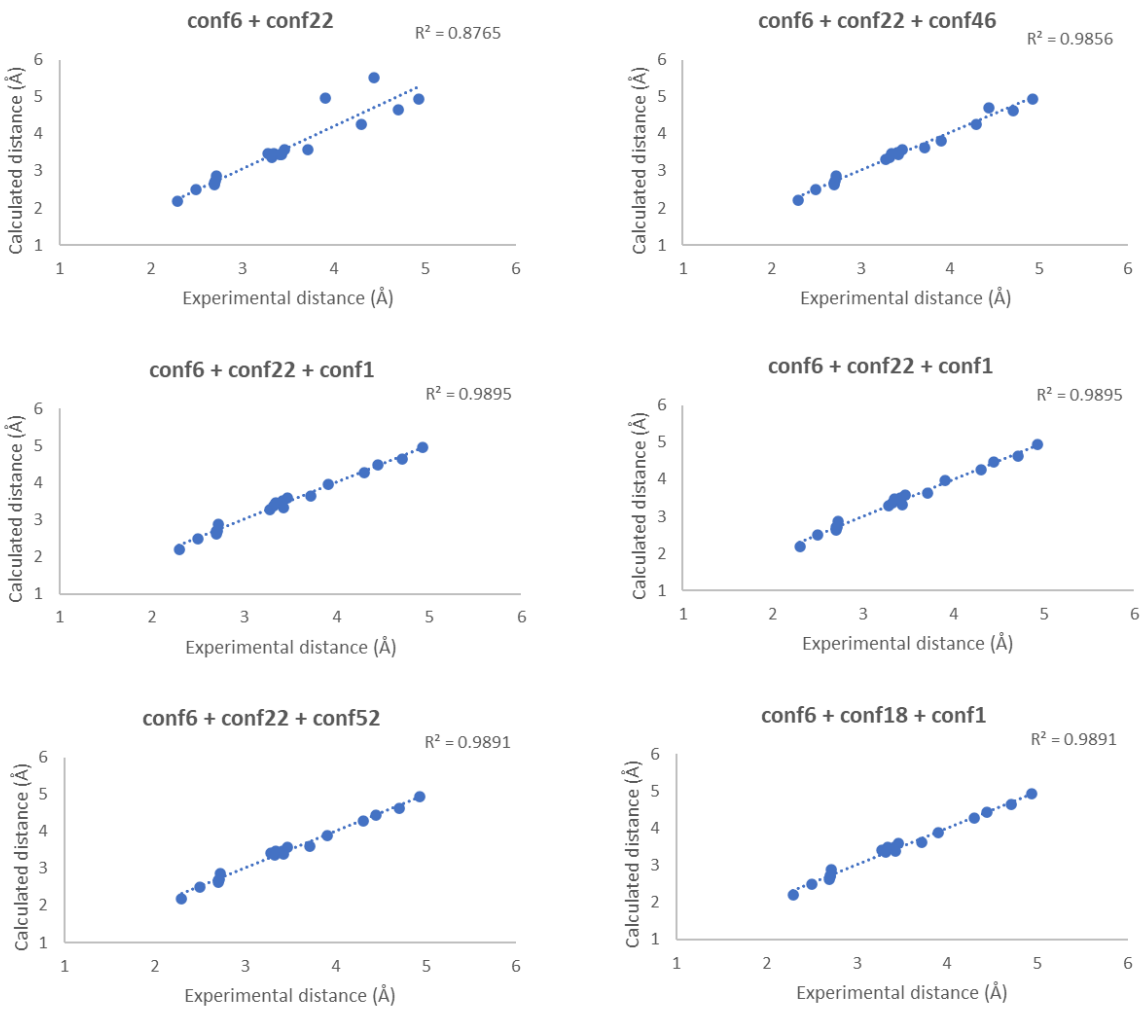


Figure 5.37 – Back-calculated against observed distances provided in Table S21.

5.2.4. Analysis through Residual Dipolar Couplings in PAN gel

The ^{1}H-NMR-spectrum of Pacritinib in polyacrylonitrile (PAN) in DMSO-d_6 is shown in Figure 5.38. The cross-linked PAN results in two broad signals at 2.0 ppm (CH_2-groups) and 3.1 ppm (CHCN-groups). Signals of unknown impurities at 6.0 and 6.3 ppm are present. As DMSO is quite hygroscopic, small amounts of water diffused into the gel.

As demonstrated by Navarro et al.[91] one of the advantages of PAN gel is that when completely compressed, only anisotropic signals should be visible; this feature should facilitate the analysis of spectra. Hence, this approach was tested: the sample was compressed to test whether or not isotropic signals from the solution disappear.

Unfortunately, a strong degree of compression led to a sample difficult to shim resulting in a very poor spectra quality and very large errors on measurements. Different degrees of alignment were then tested to find better-quality spectra. Quadrupolar splittings of deuterium of the solvent were measured to assess the degree of alignment; there were tested strong, medium and weak degrees in which $\Delta\nu_Q$ were 30.7 Hz, 26.2 Hz, 14.8 Hz, 7.45 Hz, 6.24 Hz and 4.7 Hz.

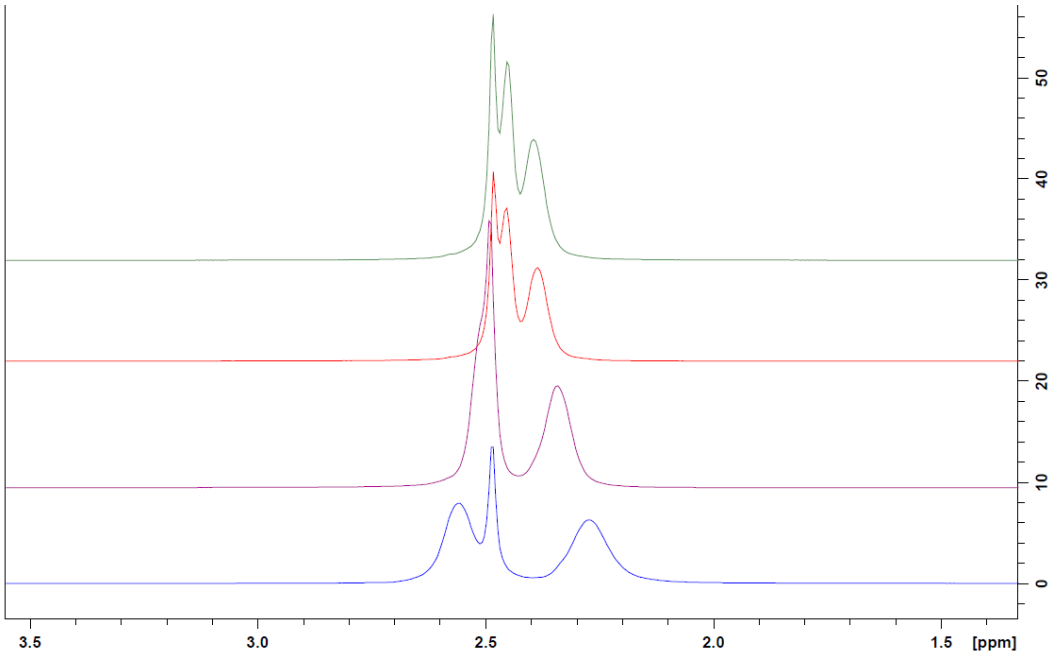


Figure 5.38 – ^2H NMR spectra of Lorlatinib in PAN/DMSO-d_6. Different alignment degrees were testes. Top-down quadrupolar splittings of the solvent signal are as follows: 6.24 Hz (green trace), 7.45 Hz (red trace), 18.2 Hz (purple trace), 30.7 Hz (blue trace).

F1-coupled HSQC of the sample stretched until a quadrupolar splitting of the solvent equal to 30.7 Hz was acquired. In Figure 5.39 superposition of isotropic and anisotropic spectra is shown. Signals coming from the gel presented very low intensities and distortions. For this reason, total dipolar couplings couldn't be extracted from the sample when it was totally compressed.

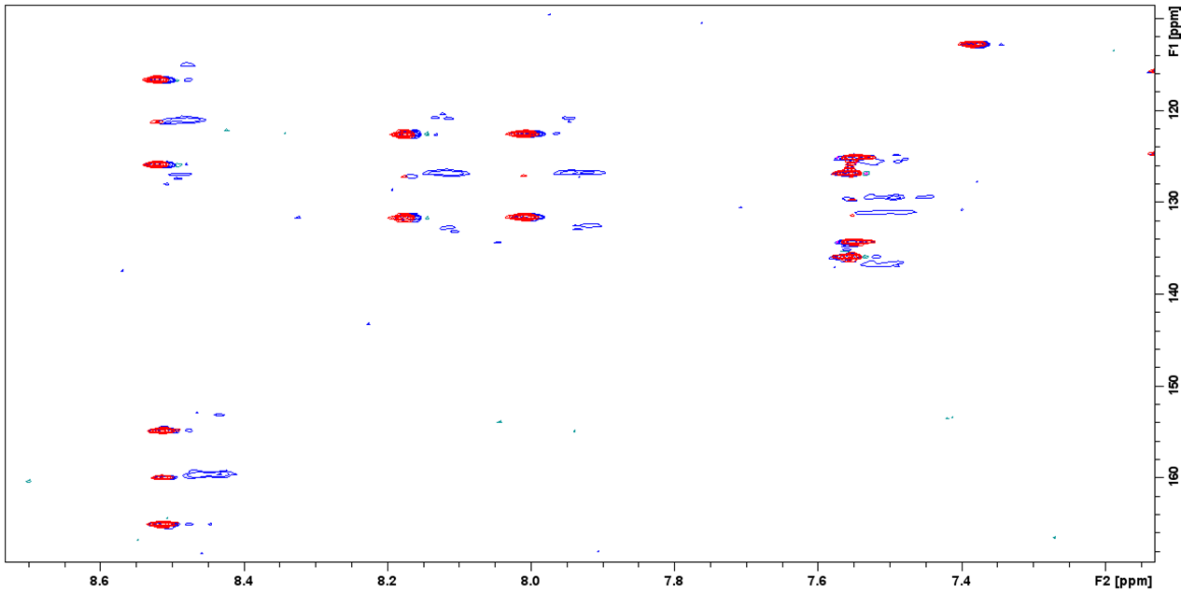


Figure 5.39 – Superposition of F1-coupled HSQC acquired in solution (red peaks) and gel (blue peaks).

Among the several degrees of alignment tested, the one which allowed to obtain the clearest spectra was chosen, i.e. the one in which ^{2}H quadrupolar splitting of the solvent was 7.45 Hz.

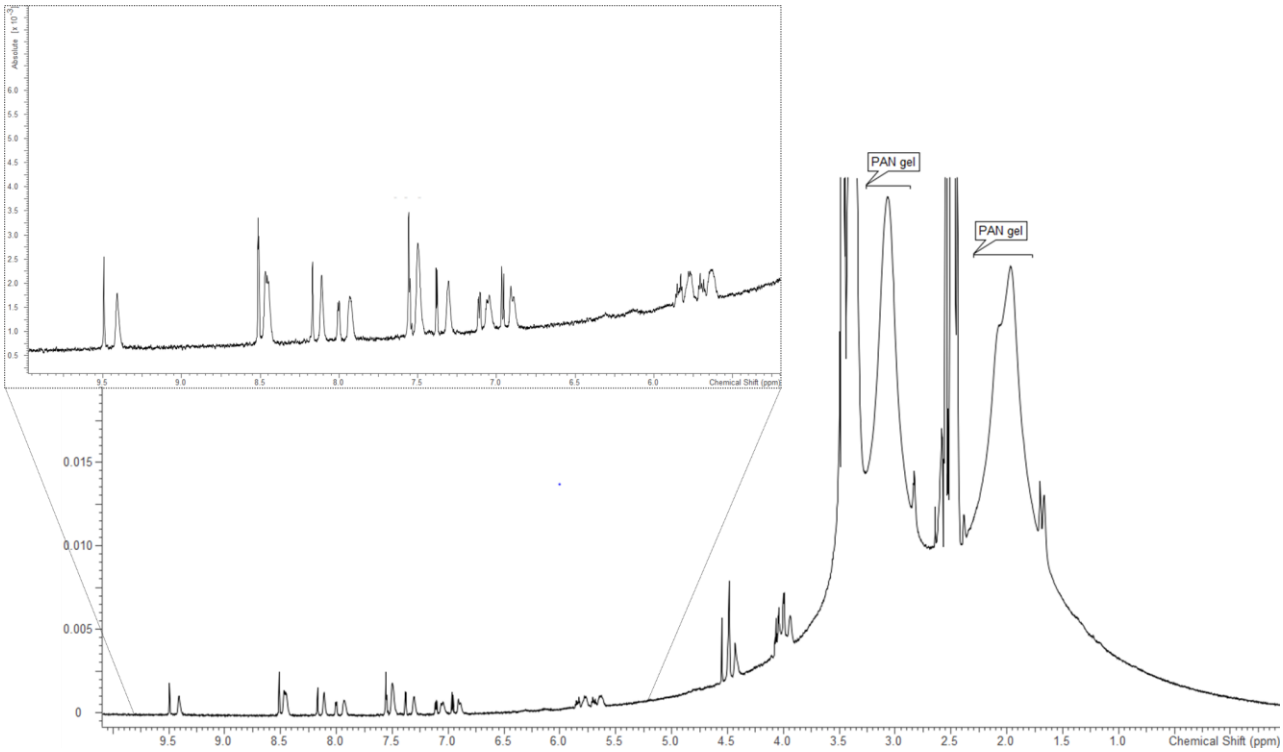


Figure 5.40 – ^{1}H spectrum of Pacritinib in PAN/DMSO-d_6. It is dominated by large broad

signals at 2 and 3.1 ppm coming from the polymer PAN; only in the enlargement the weak signals of the analyte are visible. Notably, peak multiplicity was still detectable for most of the peaks since polymeric gels usually cause signals' broadening.

Short-range and long-range Residual Dipolar Couplings were measured using this weakly aligned sample.

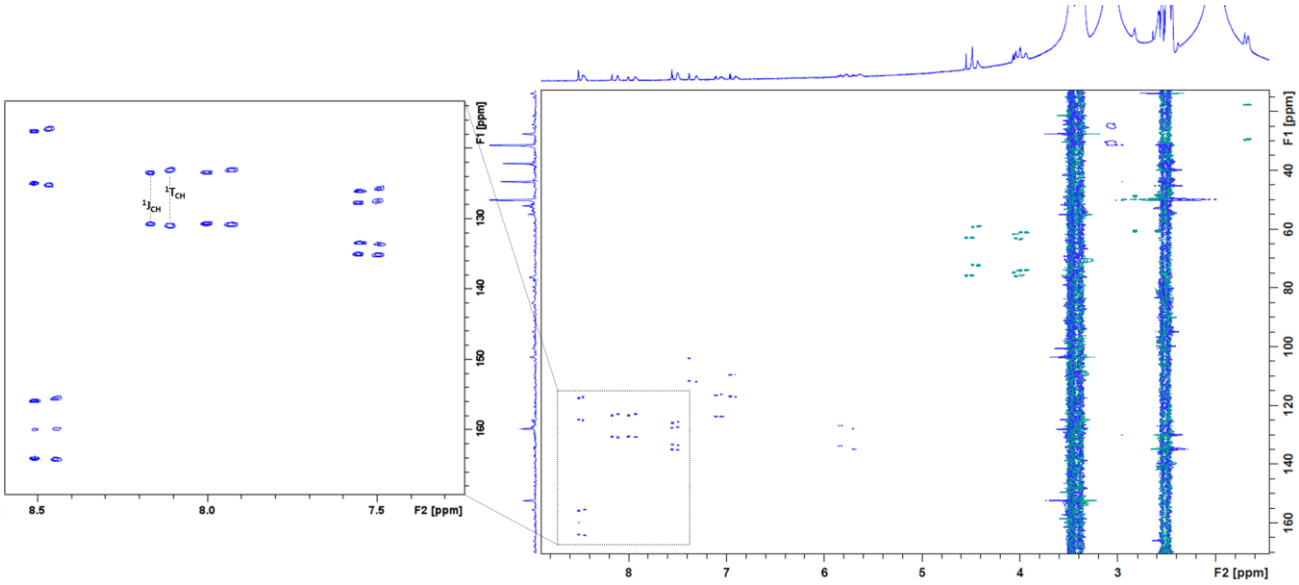


Figure 5.41 – ^{1}H-^{13}C F1-coupled HSQC spectrum of Pacritinib in PAN/DMSO-d_6. Alignment degree was weak ($\Delta\nu_Q$ = 7.45 Hz). Blue peaks represent CH and CH_3 groups, while green peaks refer to CH_2 groups. The inset shows the zoom in of the aromatic region to highlight the high-quality of the spectrum; all signals coming from isotropic and anisotropic environments are clearly distinguishable. $^{1}J_{CH}$ and $^{1}T_{CH}$ of C3-H9 coupling are drawn. The experiment was acquired corresponding to a spectral window of 5400 (F2) and 634 Hz (F1) respectively. Each dataset was processed by applying a zero-filling of 2048 in F1 and a function of the type π/2-shifted cos2 (QSINE, SSB: 2) in both dimensions.

A total of 17 residual dipolar couplings were calculated from scalar and total couplings extracted from ^{1}H-^{13}C F1-coupled HSQC spectrum. Total coupling regarding C1-H12 was not well resolved (peak was not symmetric), hence it was not included into the analysis.

^{13}C	^{1}H	δ ^{13}C (ppm)	δ ^{1}H (ppm)	$^{1}J_{CH}$ (Hz)	$^{1}T_{CH}$ (Hz)	$^{1}D_{CH}$ (Hz)	error
36	42	159.06	8.51	178.936	190.569	11.632	0.137
27	33	120.34	8.52	162.474	174.177	11.703	0.227
3	9	126.24	8.18	159.561	173.112	13.551	0.344
5	10	126.24	8.01	159.525	169.902	10.378	0.289
6	11	130.54	7.55	160.994	174.307	13.313	0.295
35	41	107.02	7.39	168.387	179.966	11.579	0.176
25	31	119.37	7.11	159.033	166.810	7.776	0.086
30	66	112.5	6.96	158.806	170.955	12.150	0.182
17	24	129.57	5.84	152.224	156.999	4.775	0.129
19	23	130.8	5.69	152.062	157.438	5.376	0.247
8	12,13	68.67	4.55	286.1218	280.3618	-2.880	0.065
63	64,65	65.01	4.48	280.8722	294.6733	6.901	0.142
43	45,46	67.79	4.05	289.0620	287.2718	-0.895	0.416
20	21,22	68.98	4.04	284.0121	283.5709	-0.221	0.184
15	16,18	66.85	3.99	287.0745	277.7405	-4.667	0.095
44	47,48	54.24	2.79	259.9545	274.5320	7.288	0.193
51,52	56,57,58,59	22.99	1.68	266.1715	254.9783	-5.597	0.097

Table S22 – Scalar ($^{1}J_{CH}$), total ($^{1}T_{CH}$) and residual dipolar couplings ($^{1}D_{CH}$) of Pacritinib in PAN/DMSO-d_6. $^{1}D_{CH}$ of methylene protons are divided by 2 for use in StereoFitter.

PSYCHEDELIC pulse sequence was used to measure proton-proton Redisual Dipolar Couplings. Acquisition in isotropic solution was successful leading to 10 $^{n}J_{HH}$, whereas acquisition in anisotropic environment did not lead to any anisotropic signals (see Figure 5.42). Therefore, it was not possible to obtain any long-range $^{1}D_{CH}$. A possible explanation may be that T_2 (spin-spin) relaxation in polymeric gel is too fast causing the loss of signals.

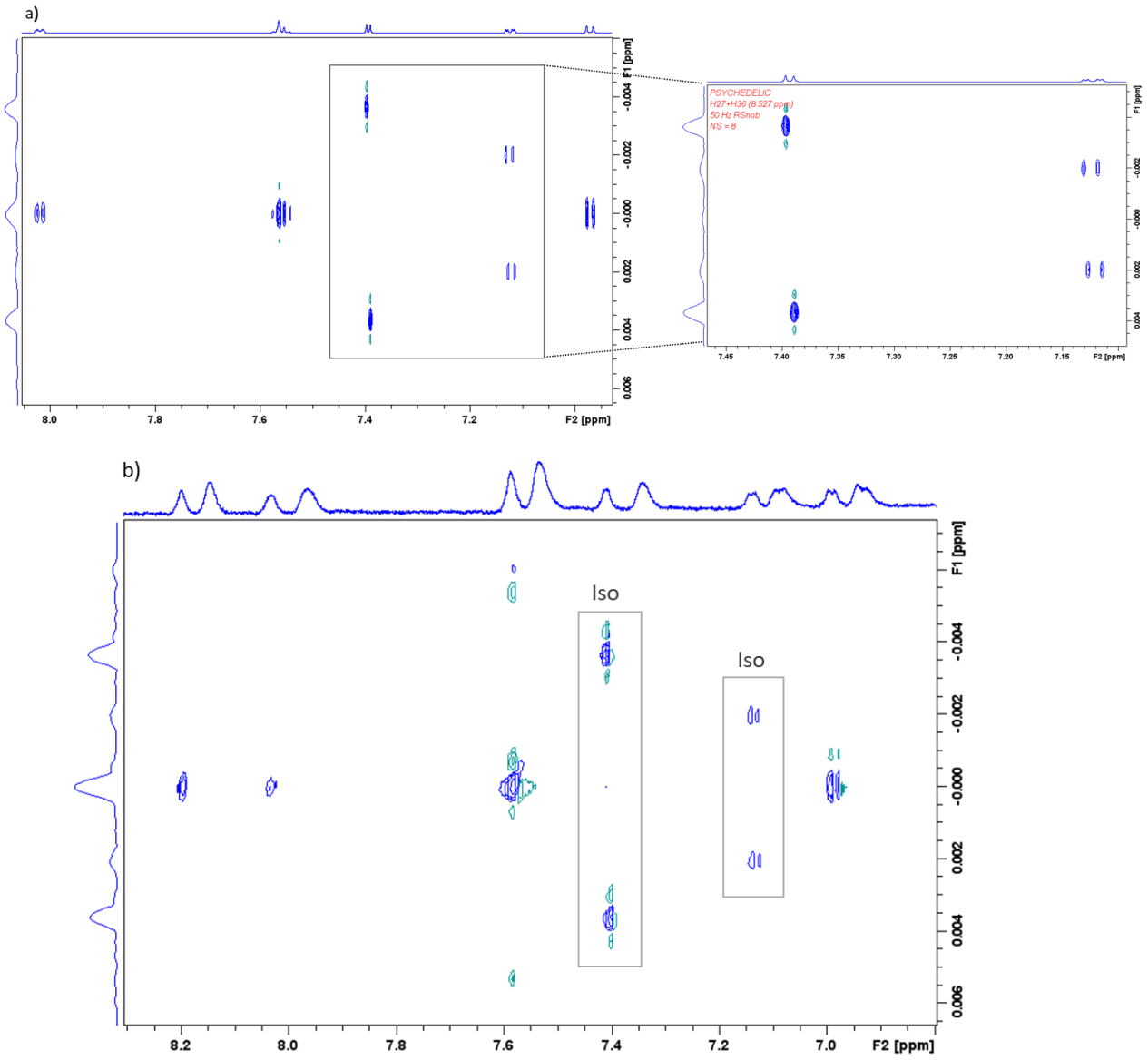


Figure 5.42 – a) PSYCHEDELIC spectrum of Pacritinib in DMSO-d_6 (isotropic). b) PSYCHEDELIC spectrum of Pacritinib in PAN/DMSO-d_6 (anisotropic). Protons H27 and H36 were selectively inverted using RSnob selective pulse with 50 Hz bandwidth. In the anisotropic spectrum only signals coming from the residual solution of the analyte surrounding the swollen gel are present; any anisotropic coupling was detected.

Heteronuclear long-range residual dipolar couplings were measured through HSQC-HECADE spectrum (Figure 5.43). 21 scalar couplings ($^{n}J_{CH}$) were extracted in solution, whereas only 5 total couplings were measured in polymeric gel. Moreover, the experimental error of 4 out of 5 residual dipolar couplings was close to or even bigger than the measured value. Therefore, these measurements were not used for StereoFitter analysis.

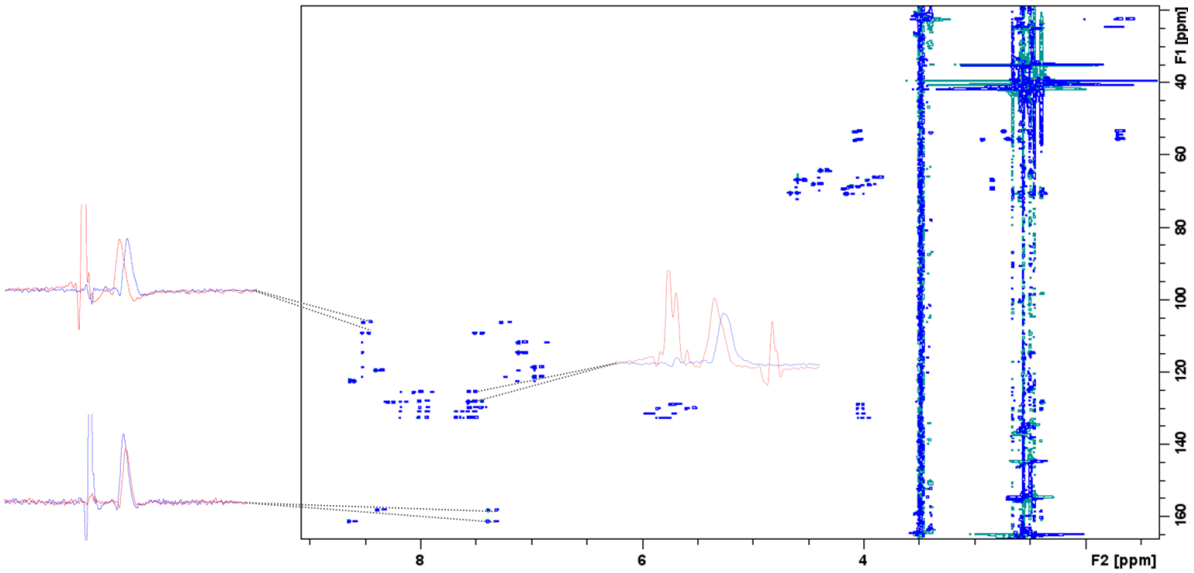


Figure 5.43 – ^{1}H-^{13}C HSQC-HECADE spectrum of Pacritinib in PAN/DMSO-d_6. F2 rows are depicted to show that some signals are broad or asymmetric leading to large error in experimental measurements.

^{13}C	^{1}H	δ ^{13}C (ppm)	δ ^{1}H (ppm)	$^{n}J_{CH}$ (Hz)	$^{n}T_{CH}$ (Hz)	$^{n}D_{CH}$ (Hz)	error
C36	H35	159.19	7.39	0.799	0.916	0.117	0.500
C3	H1	127	7.5	1.502	1.964	0.462	0.811
C1	H5	129.39	7.93	2.495	0.396	-2.099	1.102
C35	H36	107.02	8.51	2.518	3.081	0.563	0.369
C25	H30	120	6.96	0.059	0.492	0.433	0.247

Table S23 – Long-range ($^{n}D_{CH}$) couplings obtained from HSQC-HECADE spectrum shown in Figure 5.43.

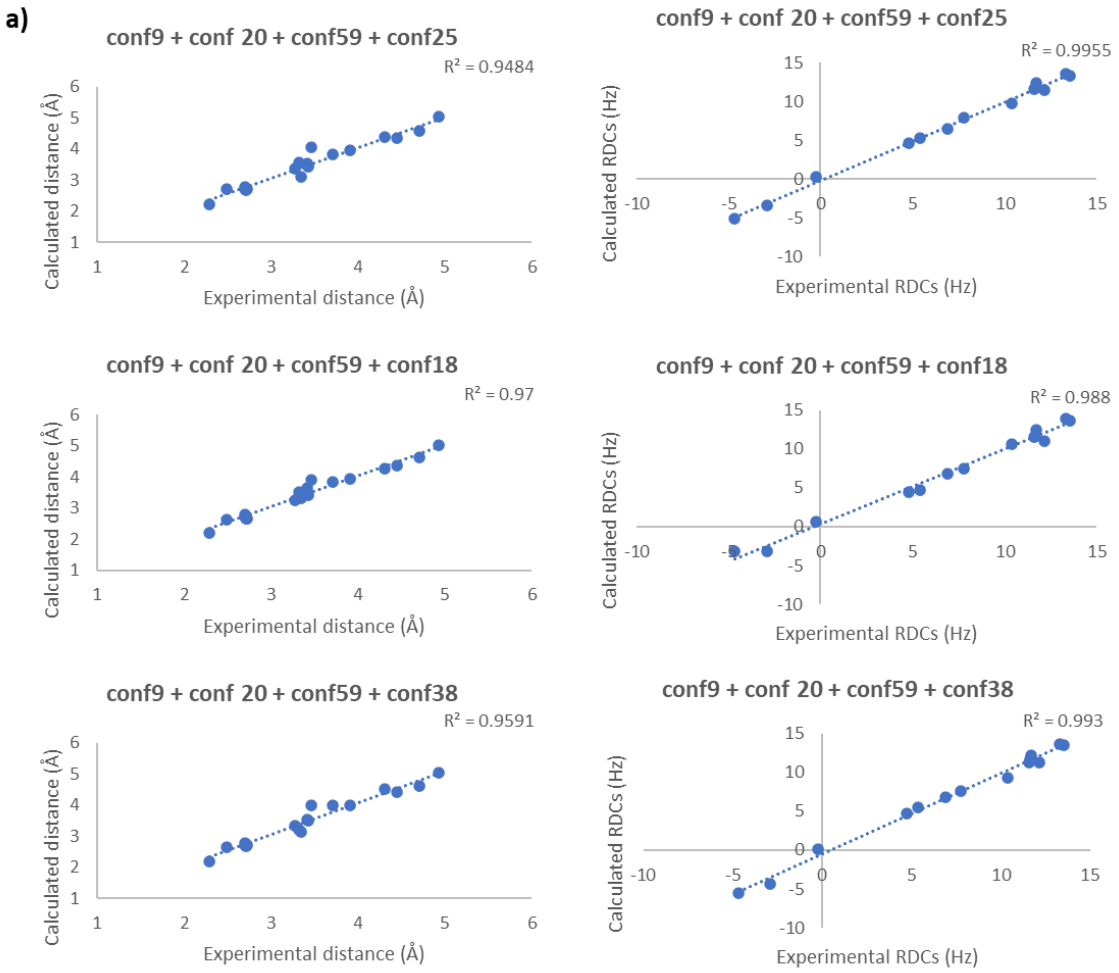
A total of 14 out of 17 (extracted) RDCs were used for the multiparametric fitting. The influence of the flexible sidechain was not taken into account; therefore, residual dipolar couplings belonging to the sidechain with pyrrolidine group were not used in StereoFitter analysis.

	Solution	1	2	3
		Conformer 9: 33.32%	Conformer 9: 17.73%	Conformer 9: 31.08%
		Conformer 20: 25.80%	Conformer 20: 24.23%	Conformer 20: 23.28%
		Conformer 59: 26.17%	Conformer 59: 35.43%	Conformer 59: 29.21%
		Conformer 25: 14.71%	Conformer 18: 22.61%	Conformer 38: 16.43%
	AIC	11.159	12.577	12.843
	χ^2	5.159	6.577	6.843
	Relative probability	1	0.492	0.431
	NOE distances (Å)			
Atoms (^{1}H-^{1}H)	**Experimental**	**Calculated**	**Calculated**	**Calculated**
9,33	2.495	2.695	2.64	2.658
9,41	4.71	4.594	4.638	4.624
9,24	3.906	3.944	3.954	4.003
9,23	3.412	3.542	3.642	3.52
9,(64,65)	3.463	4.071	3.903	3.981
9,(16,18)	4.302	4.381	4.27	4.507
10,41	2.295	2.219	2.187	2.193
(64,65),66	4.931	5.03	5.019	5.029
24,33	4.444	4.349	4.377	4.428
(12,13),24	3.278	3.354	3.259	3.318
24,(64,65)	3.428	3.421	3.402	3.511
(16,18),24	2.698	2.697	2.725	2.692
(21,22),24	2.7	2.791	2.801	2.781
23,33	3.715	3.831	3.836	3.974
(12,13),23	3.324	3.564	3.502	3.195
23,(64,65)	3.341	3.104	3.331	3.122
(21,22),23	2.708	2.667	2.658	2.672
(16,18),23	2.718	2.712	2.656	2.718
	RDCs (Hz)			
Atoms (^{13}C-^{1}H)	**Experimental**	**Calculated**	**Calculated**	**Calculated**
36,42	11.633	11.776	11.706	11.701
27,33	11.703	12.419	12.483	12.239
3,9	13.551	13.329	13.612	13.533
5,10	10.378	9.82	10.653	9.373
6,11	13.313	13.519	13.909	13.686
35,41	11.579	11.657	11.622	11.289
25,31	7.776	7.991	7.53	7.605
30,66	12.149	11.496	10.975	11.36
17,24	4.775	4.656	4.415	4.762
19,23	5.376	5.266	4.668	5.485
(8,12) (8,13)	-2.88	-3.317	-3.19	-4.376
(63,64) (63,65)	6.901	6.438	6.858	6.774
(20,21) (20,22)	-0.221	0.343	0.603	0.076
(15,16) (15,18)	-4.667	-5.143	-3.194	-5.485

Table S24 – StereoFitter results of multiparametric (distances, RDCs) fitting of Pacritinib weakly aligned in PAN/DMSO-d_6.

It is important to note that all the solutions shown in Table S24 share three conformations (conformers 9, 20 and 59 which represent the most populated conformations in two solutions), while they differ for the fourth conformation (conformers 35, 18 and 38). Examining the correlation plots of back-calculated and experimental it can be noted that the contribution of conformer 18 in the second solution led to a slightly better R^2 (0.970 vs 0.948) regarding interproton distances' plot with respect to the first solution. Nevertheless, all three models are good representations of observed data regarding CH orientations (see Figure 5.41 - a).

Notably, among the fourth conformation of the solutions, the macrocyclic core of conformers 18 and 38 was superimposable except for dihedral C8-O14-C15-C17, while it was not superimposable in conformers 18 and 25.



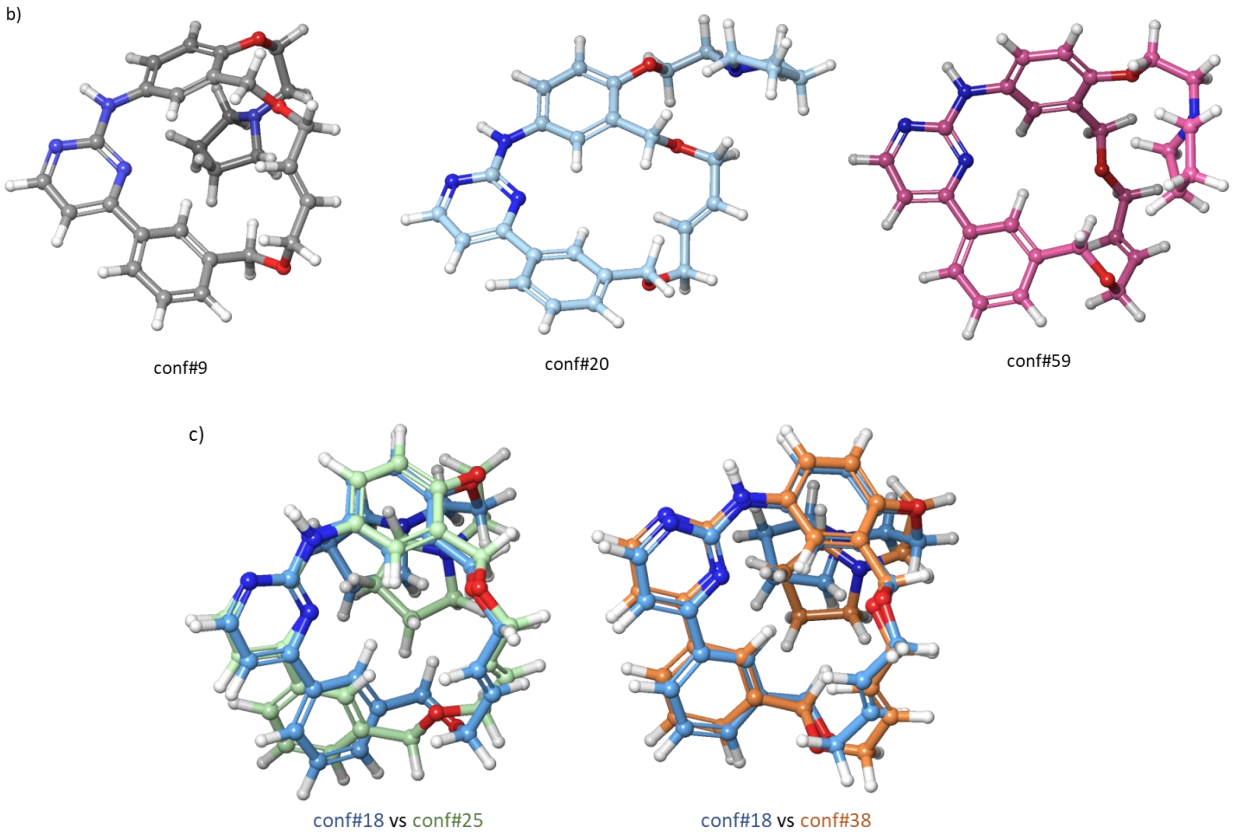


Figure 5.44 – Solutions of multiparametric fitting of Pacritinib restraints in DMSO-d_6. a) Plots of calculated vs observed interproton distances and residual dipolar couplings of StereoFitter results. b) Conformers derived from solutions shown in Table S24. c) Superimposition of conformers 18 vs 25 and conformers 18 vs 38.

Overall, contrary to the analysis based on interproton distances, multiparametric fitting allowed to obtain three likely solutions. It wasn't possible to perform the conformational analysis of Pacritinib in $CDCl_3$ or other apolar solvents because of low solubility problems.

5.3 Conformational analysis of Simeprevir

Macrocyclic compound Simeprevir ((1*R*,4*R*,6*S*,7*Z*,15*R*,17*R*)-*N*-cyclopropylsulfonyl-17-[7-methoxy-8-methyl-2-(4-propan-2-yl-1,3-thiazol-2-yl)quinolin-4-yl]oxy-13-methyl-2,14-dioxo-3,13-diazatricyclo[13.3.0.0^{4,6}]octadec-7-ene-4-carboxamide) is a selective inhibitor of hepatitis C virus NS3/4A protease. It was approved in 2013 in combination with peginterferon-alpha and ribavirin for treatment of hepatitis C genotype 1 and 4.[92]

5.3.1 Conformational search

Conformational space was explored using SPMC (Stationary Phase Monte Carlo) and mixed-torsional/Low MODe (LMOD)[93] methods. The latter is a hybrid method which exploits the efficiency of LMOD in exploring local minima and the effectiveness of Monte Carlo conformational search in sampling the global potential energy surface PES)[94]. Both searches were carried out by means of enhanced torsion sampling and RMDS cutoff of 0.5 Å. All conformations within 10 kcal/mol from the global minimum were saved. Conformations obtained through SPMC only differ for torsionals of quinoline group; the macrocyclic core wasn't well sampled. For this reason, conformers found using LMOD method were used for the following analysis. The geometry of 613 conformers were optimized through quantum mechanics (DFT-B3LYP theory with 6-31G+* basis set) in order to discard conformation with SCF energy greater than 5 kcal/mol. Refine conformations using DFT usually allows to reduce the number of conformations due to deeper basins and higher accuracy on prediction of energies which allow for lower energy threshold. A total of 212 conformers were selected for NMR-based conformational analysis. They were analyzed and divided into 10 clusters according to RMSD values comparing the macrocyclic core; the 14-membered ring presented some degree of flexibility as indicated by RMSD between 0.54 and 1.19 Å. All conformers within each cluster were superimposable as regards the macrocyclic core, while they differ in the orientations and conformations of quinoline moiety and acyl-sulfonamide sidechain.

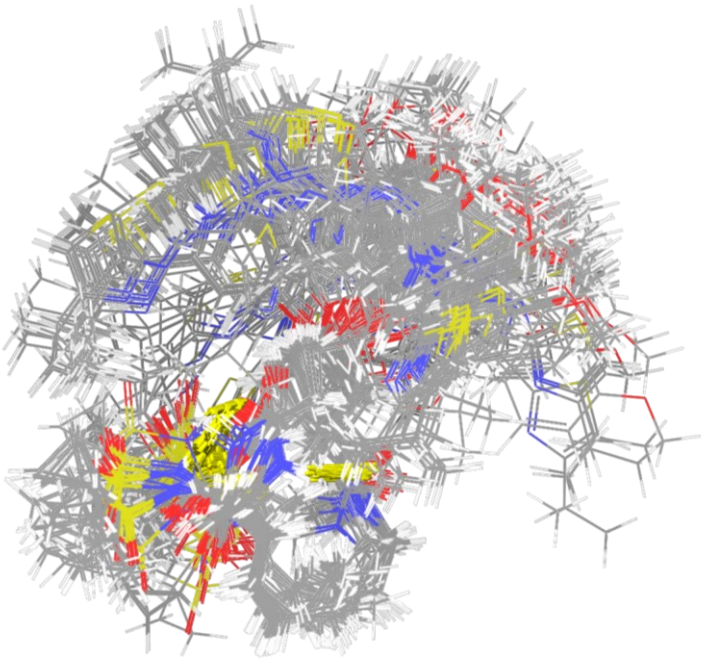


Figure 5.45 – Superposition of 212 conformations obtained through mixed-torsional/LMOD search.

5.3.2 Intramolecular hydrogen bonds (IMHBs) assessment

Variable Temperatures (VT) NMR experiments were recorded in order to assess the presence of intramolecular hydrogen bonds (IMHBs) of acyl sulfonamide proton NH-11a (resonance peak at 11.45 ppm in DMSO-d_6 and 11.09 ppm in $CDCl_3$) and amide proton NH-10a (8.69 ppm in DMSO-d and 6.42 in $CDCl_3$).

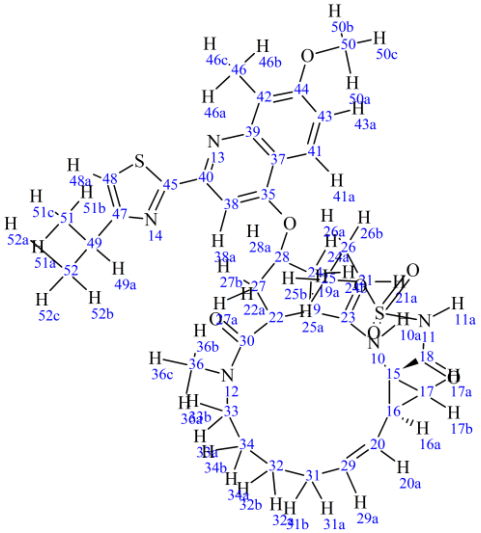
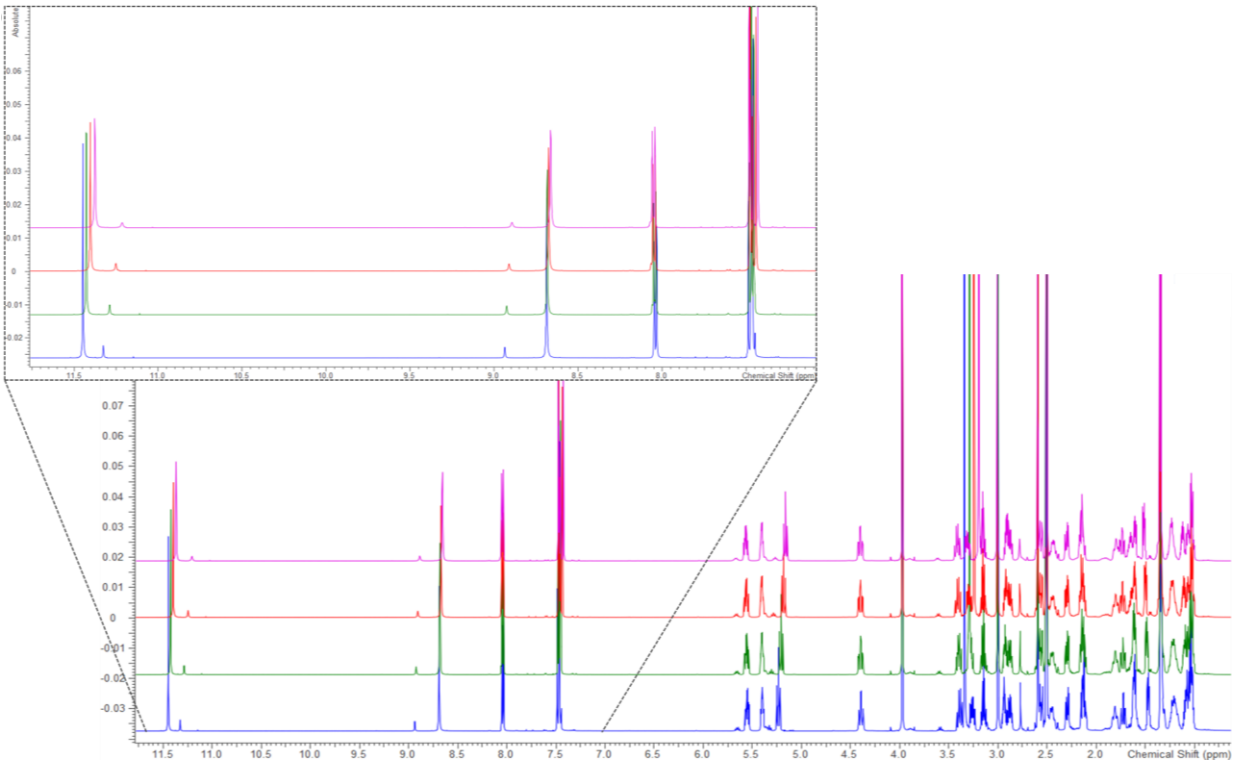


Figure 5.46 – 2D structure of Simeprevir. Atom numbering according to ACD/Spectrus.

A



B

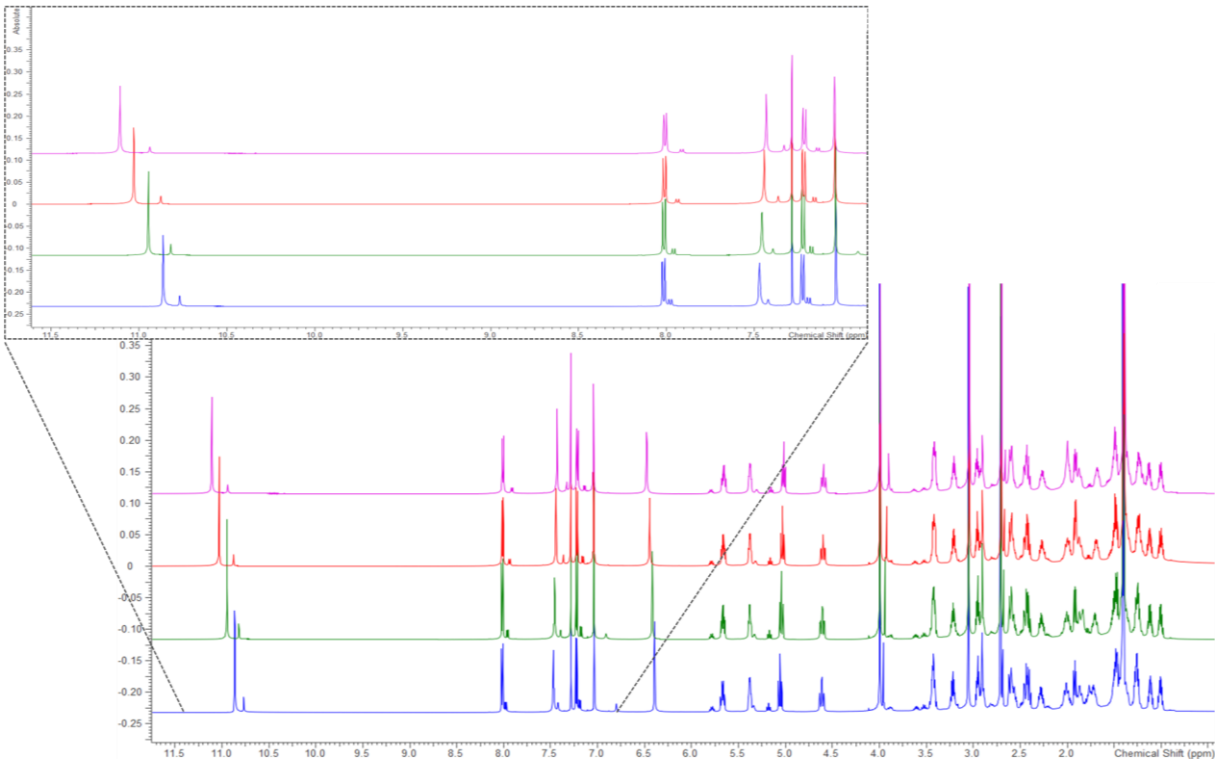


Figure 5.47 – ^{1}H NMR VT (variable temperature) spectra of Simeprevir DMSO-d_6 (A) and $CDCl_3$ (B). Enlargements of both spectra are shown to better highlight the temperature dependence of NH chemical shifts. Temperatures tested in DMSO-d_6 were 298 K (blue line),

308 K (green line), 318 K (red line), 328 (pink line). Temperature tested in $CDCl_3$ were 273 K (pink line), 283 K (red line), 293 K (green line), 303 K (blue line).

A

	chemical shifts in DMSO-d_6	
T (K)	NH-11	NH-10
298	11.4475	8.6884
308	11.4273	8.6848
318	11.4035	8.6767
328	11.3763	8.6642

B

	chemical shifts $CDCl_3$	
T (K)	NH-11	NH-10
273	11.0925	6.4760
283	11.0127	6.4440
293	10.9316	6.4169
303	10.8472	6.3940

C

Proton	T coeff. (DMSO-d_6)	T coeff. ($CDCl_3$)
NH-11	-2.37	-8.18
NH-10	-0.8	-2.73

Table S25 – Resonance frequencies of NH-10 and NH-11 in DMSO-d_6 (A) and $CDCl_3$ (B). Calculated temperature coefficients for both analytes in DMSO-d_6 and $CDCl_3$ (C).

Temperature coefficient (ppb/K) were calculated to determine whether amide protons are involved in IMHBs or they are shielded from the solvent. In DMSO-d_6, temperature coefficients of both protons presented an absolute value below 3 ppb/K; therefore, since in DMSO-d_6 hydrogen bonds network can exist between the analyte and the solvent, it can be stated that amide proton and acyl sulfonamide proton of Simeprevir in polar environment are involved in intramolecular hydrogen bond or they are sterically shielded by the solvent.

Regarding, instead, the temperature dependence of amide proton in $CDCl_3$, resonance of acyl sulfonamide proton (NH-11) shifted a lot upon increasing temperature. In fact, its temperature coefficient is higher than 3, suggesting that proton NH-11 is totally exposed to the solvent.[95] On the other hand, temperature coefficient of amide proton NH-10 is low (0.27), indicating that NH-10 is involved in an intramolecular hydrogen bond.

5.3.3 Conformational analysis in DMSO-d_6

5.3.3.1 NMR assignments

NMR assignments were performed through acquisition and analysis of ^{1}H, COSY, TOCSY, ^{1}H-^{13}C HSQC, ROESY, ^{1}H-^{13}C HMBC, ^{1}H-^{15}N HMBC spectra recorded on

a 600 MHz Bruker AVANCE III HD spectrometer equipped with cryogenic TCI probe.

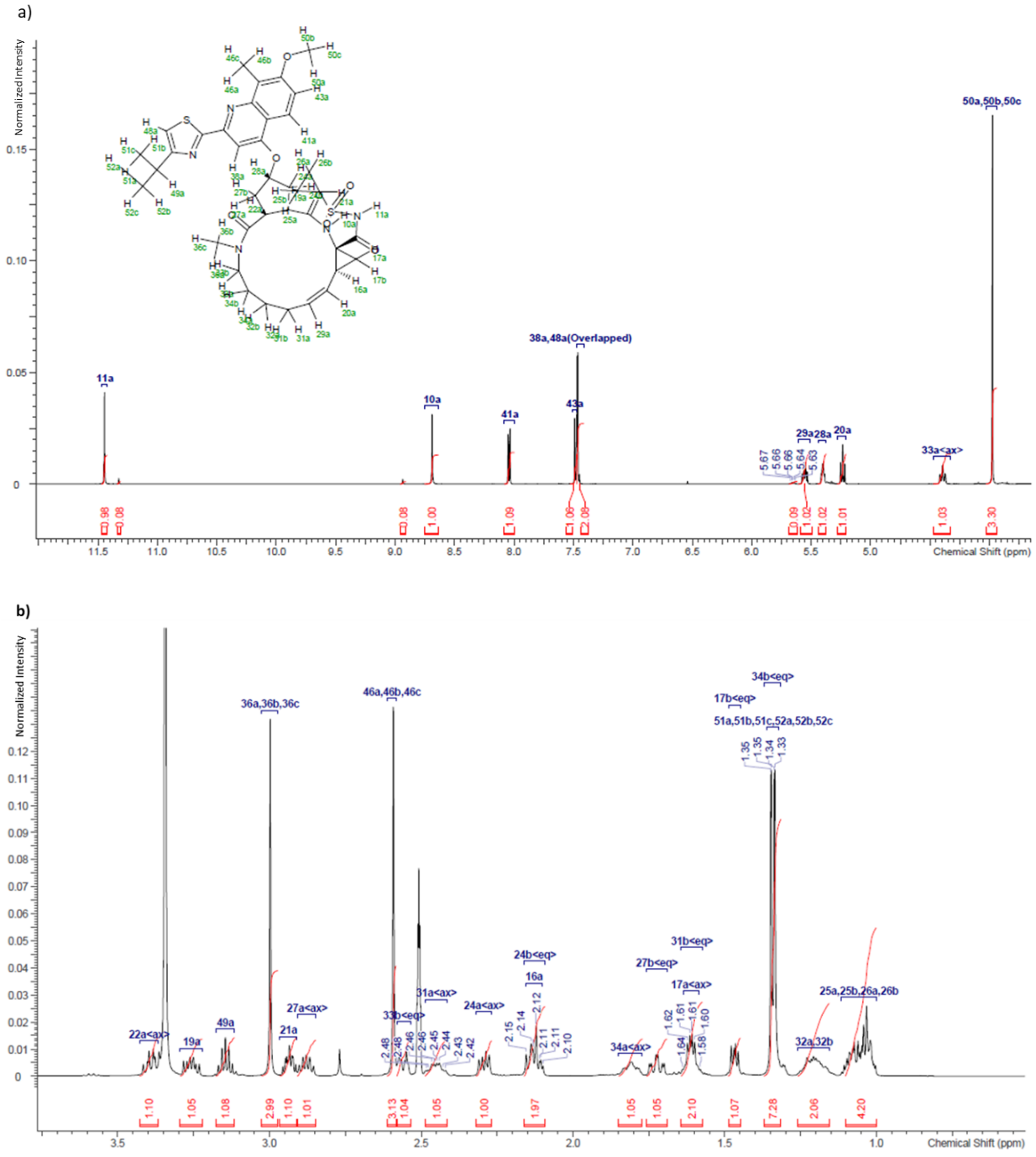


Figure 5.48 – ^{1}H NMR spectrum of Simeprevir in DMSO-d_6. a) Enlargement of ^{1}H spectrum between 12 and 3.8 ppm. b) Enlargement of ^{1}H spectrum between 3.8 and 0 ppm.

All methylene protons of the macrocycle except H32a,b have distinct resonance frequencies because they are diastereotopic protons, hence they are non-equivalent

protons. Moreover, even if the aliphatic region of the spectrum is quite crowded, only few signals such as H16a-H24a and H34b-Me51,52 overlap.

Simeprevir in DMSO-d_6 exhibited two sets of peak due to the cis/trans isomerization of the amide N10-C23 which is slow on the NMR timescale. Integrating peaks arising from the major (trans) and from the minor (cis) forms, 8% resulted from cis configuration (ratio 1:12.5 between trans and cis isomers).

No.	Atom# (^{1}H) ACD	Atom# (^{1}H) Maestro	Atom # (^{13}C)	Exp. Shift ^{1}H (ppm)	Exp. Shift ^{13}C (ppm)
1	11a	67	11	11.45	
2	10a	56	10	8.69	
3	41a	84	41	8.04	120.68
4	43a	85	43	7.49	113.19
5	48a	85	48	7.46	115.98
6	38a	83	38	7.44	95.91
7	29a	71	29	5.56	131.8
8	28a	70	28	5.41	78.83
9	20a	58	20	5.24	125.61
10	33a<ax>	77	33	4.4	43.49
11	50a, 50b, 50c	91,92,93	50	3.98	56.55
12	22	60	22	3.4	46.67
13	19	57	19	3.26	47.99
14	49a	90	49	3.15	30.84
15	36a, 36b, 36c	80,81,82	36	3	33.8
16	21a	59	21	2.94	30.83
17	27a<ax>	69	27	2.88	2.88
18	46a, 46b, 46c	86,87,88	46	2.6	10.33
19	33b<eq>	76	33	2.56	43.48
20	31a<ax>	72	31	2.45	27.4
21	24a<ax>	61	24	2.29	35.53
22	24b<eq>	62	24	2.13	35.53
23	16a	53	16	2.13	32.6
24	34a<ax>	78	34	1.81	25.95
25	27b<eq>	68	27	1.72	34.95
26	31b<eq>	73	31	1.61	27.34
27	17a<ax>	54	17	1.61	22.39
28	17b<eq>	55	17	1.47	22.39
29	51a, 51b, 51c	94,95,96	51	1.34	22.76
30	52a, 52b, 52c	97,98,99	52	1.34	1.34
31	34b<eq>	79	34	1.34	25.89
32	32a, 32b	74,75	32	1.21	24.17
33	25a, 25b	63,64	25	1.11	5.75
34	26a, 26b	65,66	26	1.11	5.75

Table S26 – NMR assignments of Simeprevir in DMSO-d_6. Atom numbering according to Maestro was also used in StereoFitter.

5.3.3.2 Analysis by means of interproton distances

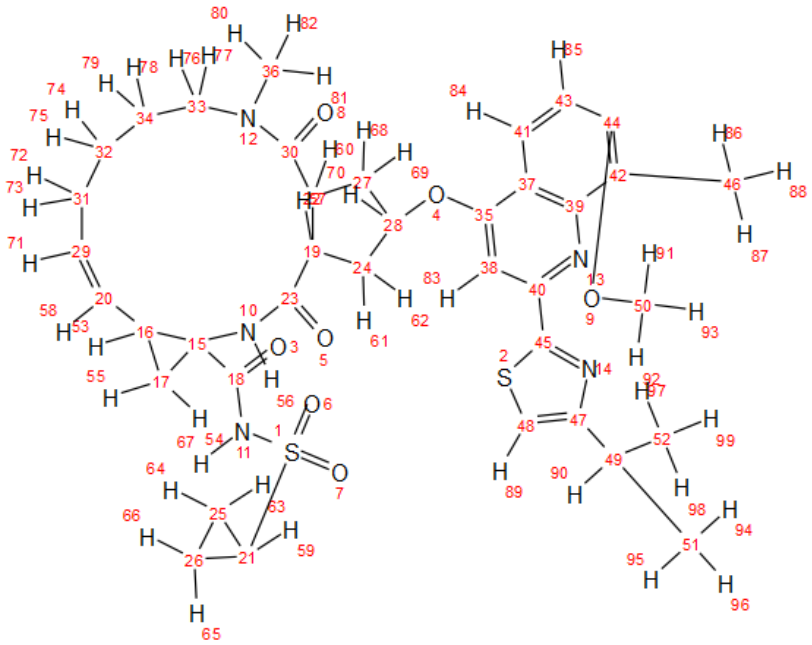


Figure 5.49 – Atom numbering used for interatomic distances' analysis, StereoFitter and Maestro software.

Proton-proton distances were measured analysing 1D-EASY-ROESY spectra derived from selective excitation of 20 protons. 10 mixing times for each resonance frequency were used to calculate build-up curves by applying PANIC approach. Only linear build-up rates (R^2> 0.95) were used for the analysis. The distance between two geminal diastereotopic protons of methylene group 33 was chosen as reference distance.

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	Atom # i	Atom # j	σ_{ij}	R^2	distance r_{ij} (Å)	average distance rij (Å)
H41	H19	8.04	3.26	84	57	0.0060	0.972	3.978	**3.727**
H41	H46a,b,c	8.04	2.60	84	86, 87, 88	0.0011	0.987	5.278	
H41	H24a	8.04	2.29	84	61	0.0031	0.950	4.441	-
H41	H27b	8.04	1.72	84	68	0.0053	0.959	4.061	**3.560**
H29	H16	5.55	2.13	71	53	0.0051	0.970	**4.087**	
H29	H31b	5.55	1.61	71	73	0.0626	0.980	**2.691**	
H28	H38	5.40	7.46	70	83	0.3575	0.996	**2.013**	

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	Atom # i	Atom # j	σ_{ij}	R^2	distance r_{ij} (Å)	average distance rij (Å)
H28	H27a	5.40	2.88	70	69	0.0931	0.989	2.519	**2.379**
H28	24b	5.40	2.13	70	62	0.0990	0.990	2.493	
H20	H11	5.23	11.45	58	67	0.0037	0.993	**4.312**	
H20	H21a	5.23	2.94	58	59	0.0067	0.991	**3.906**	
H20	H16	5.23	2.13	58	53	0.0249	0.994	**3.138**	
H20	H17b	5.23	1.47	58	55	0.0522	0.989	2.774	**2.623**
H33a	H33b	4.39	2.56	77	76	0.7476	0.998	1.780	ref
H33a	H31a	4.39	2.45	77	72	0.1025	0.994	2.479	**2.434**
H33a	H34b	4.39	1.34	77	79	0.0568	0.972	**2.735**	
H33a	H32a,b	4.39	1.20	77	74, 75	0.0258	0.969	**3.120**	
H50a,b,c	H48	3.98	7.46	91, 92, 93	85	0.2301	0.999	2.166	
H50a,b,c	H46a,b,c	3.98	2.56	91, 92, 93	86, 87, 88	0.0179	0.999	3.316	
H19a	H11	3.24	11.45	57	67	0.0021	0.967	**4.739**	
H19a	H10	3.24	8.69	57	56	0.1502	0.989	**2.326**	
H19a	H41	3.24	8.04	57	84	0.0135	0.99	3.475	-
H19a	H24a	3.24	2.29	57	61	0.0816	0.991	2.575	**2.420**
H19a	H24b	3.24	2.13	57	62	0.0195	0.993	**3.269**	
H19a	H27b	3.24	1.80	57	68	0.0533	0.991	2.764	**2.619**
H22	H28	3.43	5.40	60	70	0.0146	0.948	3.430	
H22	H36a,b,c	3.43	3.00	60	80, 81, 82	0.1252	0.997	2.398	**2.291**
H22	27a	3.43	2.88	60	69	0.0666	0.973	2.664	**2.444**
H22	24b	3.43	2.12	60	62	0.0578	0.963	2.727	
H36a,b,c	H22	3.01	3.40	80, 81, 82	60	0.2189	0.998	2.184	-
H36a,b,c	H27a	3.01	2.88	80, 81, 82	69	0.0438	0.992	2.856	**2.683**
H36a,b,c	H33b	3.01	2.56	80, 81, 82	76	0.0781	0.986	2.594	**2.493**
H36a,b,c	H34a	3.01	1.81	80, 81, 82	78	0.0803	0.985	2.582	**2.436**
H27a	H38	2.86	7.46	69	83	0.0456	0.995	**2.837**	
H27a	H28	2.86	5.40	69	70	0.1882	0.999	2.240	-
H27a	H22	2.86	3.39	69	60	0.1962	0.989	2.225	-
H27a	H36a,b,c	2.86	2.99	69	80, 81, 82	0.095	0.997	2.510	-
H27a	H27b	2.86	1.72	69	68	0.7657	0.999	1.773	
H33b	H33a	2.56	4.39	76	77	0.7375	0.998	1.784	
H33b	H36a,b,c	2.56	2.99	76	80, 81, 82	0.1266	0.999	2.393	-
H33b	H34a	2.56	1.81	76	78	0.0700	0.992	2.642	**2.589**
H31a	H29	2.45	5.56	72	71	0.0411	0.995	**2.887**	
H31a	H33a	2.45	4.40	72	77	0.1279	0.999	2.389	-
H31a	H16	2.45	2.13	72	53	0.4297	0.996	**1.952**	
H31a	H31b	2.45	1.60	72	73	0.7469	0.998	1.780	
H31a	H32a,b	2.45	1.18	72	74, 75	0.0595	0.964	**2.714**	
H24a	H11	2.29	11.45	61	67	0.0052	0.9866	4.074	
H24a	H10	2.29	8.69	61	56	0.0175	0.9801	3.328	
H24a	H41	2.29	8.04	61	84	0.0168	0.9784	3.351	
H24a	H38	2.29	7.46	61	83	0.0931	0.994	2.519	
H24a	H28	2.29	5.40	61	70	0.0342	0.981	2.976	
H24a	H19	2.29	3.26	61	57	0.1764	0.997	2.264	-

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	Atom # i	Atom # j	σ_{ij}	R^2	distance r_{ij} (Å)	average distance r_{ij} (Å)
H24a	H24b	2.29	2.13	61	62	0.7577	0.998	1.776	
H34a	H36a,b,c	1.81	3.00	78	80, 81, 82	0.1647	0.998	2.290	-
H34a	H33b	1.81	2.56	78	76	0.0893	0.993	2.536	-
H34a	H34b	1.81	1.34	78	79	0.6713	0.996	1.812	
H34a	H32a,b	1.81	1.20	78	74, 75	0.2053	0.998	**2.208**	
H27b	H41	1.72	8.04	69	84	0.029	0.975	3.059	-
H27b	H38	1.72	7.46	69	83	0.0216	0.993	**3.213**	
H27b	H19	1.72	3.26	69	57	0.1037	0.996	2.474	-
H27b	H27a	1.72	2.88	69	68	0.7237	0.998	1.790	
H17b	H11	1.47	11.45	55	67	0.0127	0.994	**3.511**	
H17b	H20	1.47	5.24	55	58	0.1042	0.993	2.472	-
H17b	H17a	1.47	1.61	55	54	0.6087	0.9989	1.842	

Table S27 – Interproton distances of Simeprevir in DMSO-d_6.

Interproton distances' analysis by means of StereoFitter resulted in various combinations of ensemble models comprising two or three conformers; at least 10 possible solutions with similar statistical values (χ^2, AIC) were provided by StereoFitter (data not shown). Nevertheless, information derived from interatomic distances, combined with chemical shifts information derived from quantum mechanic simulations, were found to be useful for assigning diastereotopic protons. Four groups in which diastereotopic protons were permutated were created thanks to this information.

Specifically, some assignments and inferences drawn from interproton distances and from deshielded/shielded chemical shifts predictions are as follows:

- Resonant frequency of H33a is higher (downfield) with respect to H33b. In all conformers the most deshielded proton was the one lying on the plane of the adjacent carbonyl group; it is proton 77 in all theoretical conformations. Hence, H33a was assigned to 77 while H33b to 76.
- It was evident that H29 was closer to H31b (more shielded resonance frequency) than H31a. In some conformations H29 was close to 72 while it was close to 73 in others. This deduction was also supported by QM chemical shift predictions. Therefore, protons 72 and 73 were permutated.
- The distance (2.379 Å) between H27a and H28 allowed to assign H27a to atom 69 in all conformers.

- Analysing the distance between methyl group CH_3-36 and H34a, the latter corresponded to atom 78 in some structures, to atom 79 in others.

Four input files regarding the four groups with permutated protons (as indicated in the above inferences) were created and StereoFitter was run. As mentioned above, 10 solutions for each group were printed out. However, assignments used within the group with the best-fitting (AIC) value, group 3, were chosen to be employed for the subsequent multiparametric analysis.

5.3.3.3 Analysis through Residual Dipolar Couplings

Simeprevir was aligned using polyacrylonitrile gel in DMSO-d_6. A solution of the analyte was added to a dried stick of gel and, after 48 hours for equilibration, the sample was compressed. A quadrupolar splitting of the deuterium of the solvent was achieved; it was 20 Hz.

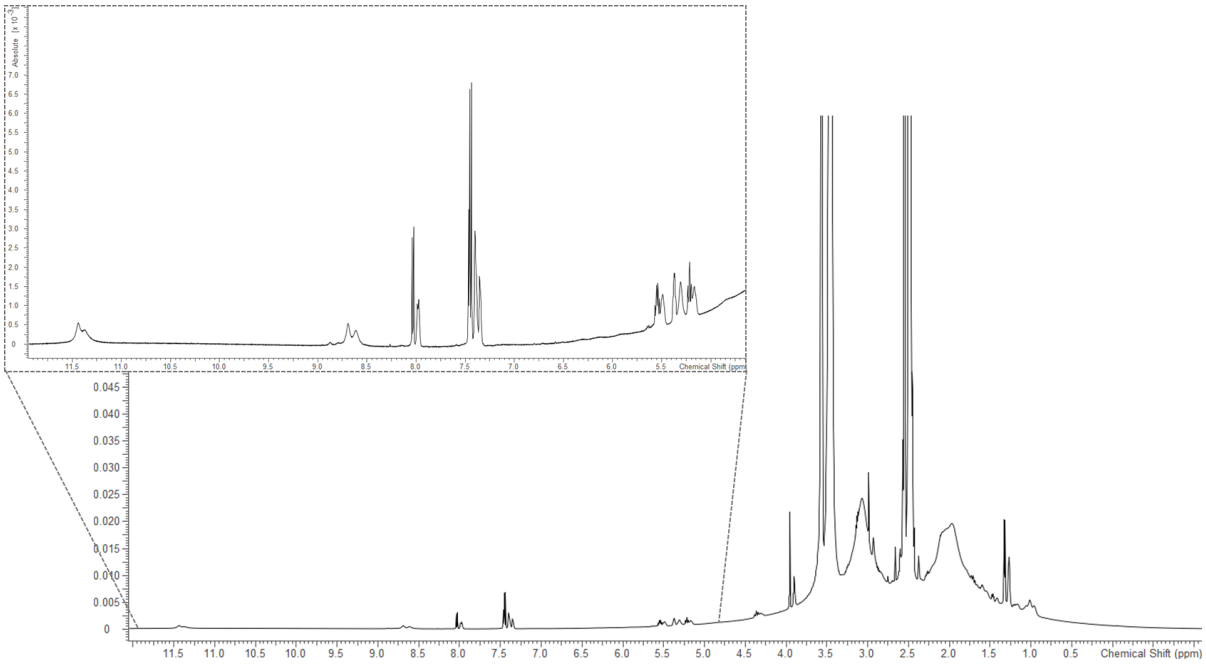


Figure 5.50 – ^1H spectrum of Pacritinib in PAN/DMSO-d_6. Broad signals at 2 and 3.1 ppm coming from the polymer PAN are predominant; in the enlargement the weak signals of the analyte (both isotropic and anisotropic) are visible.

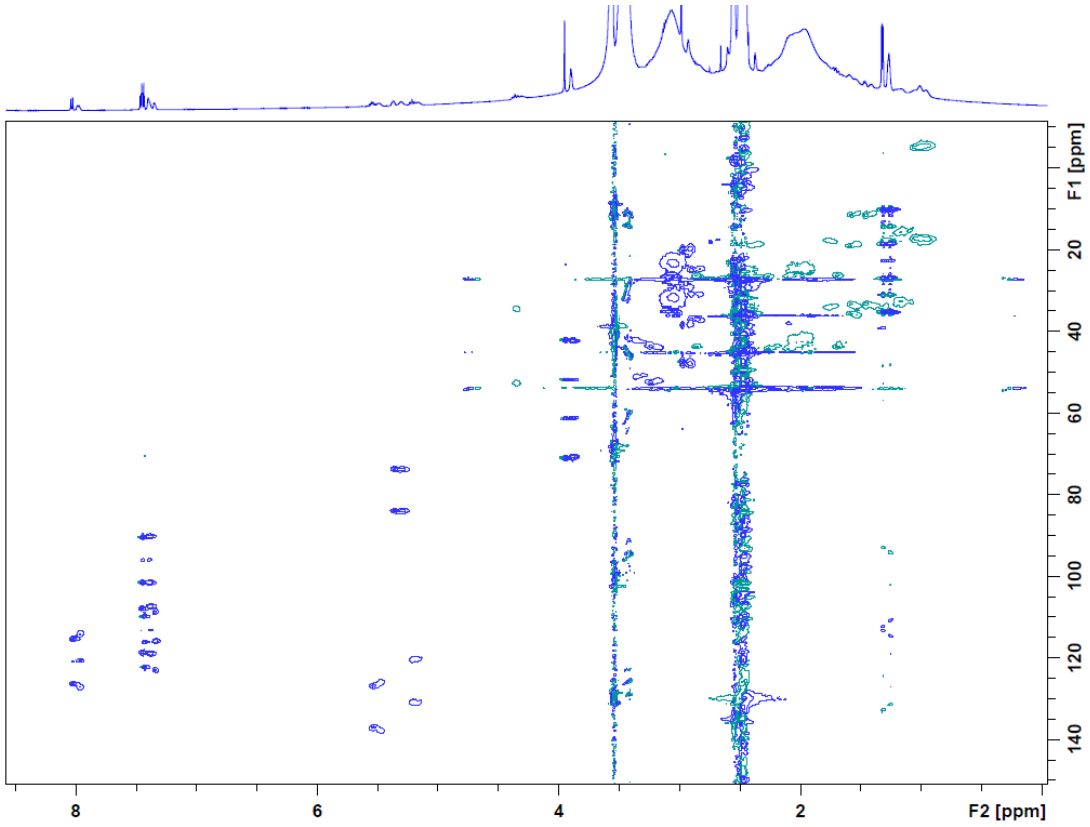


Figure 5.51 – F1-coupled-HSQC of Simperevir in DMSO-d_6/PAN. Signals belonging to polymeric gel PAN are clearly visible at 2 and 3.1 ppm and overlap with some methylene and methyl group's signals of Simeprevir.

A total of 14 out of 26 residual dipolar couplings were extracted from F1-coupled-HSQC. The very intense signals of the polymeric gel in the aliphatic region did not allow to measure a lot of total couplings regarding methylene groups because of overlapping. $^{1}J_{CH}$-resolved-HSQC spectra were also acquired in order to circumvent the problem but only 2 anisotropic couplings regarding two separated diastereotopic protons could be extracted. Moreover, they presented very large experimental errors; in fact, it was not possible to measure them with high accuracy since peaks were broad and asymmetric. Hence, they were not used for StereoFitter analysis.

^{13}C	^{1}H	δ ^{13}C (ppm)	δ ^{1}H (ppm)	$^{1}J_{CH}$ (Hz)	$^{1}T_{CH}$ (Hz)	$^{1}D_{CH}$ (Hz)	error
41	84	120.68	8.04	164.4482	195.9313	31.483	2.794
43	85	113.19	7.49	160.8238	173.9505	13.127	2.852
48	89	115.98	7.46	186.6167	214.2447	27.628	2.395
38	83	95.91	7.44	165.9729	169.0346	3.062	2.697

^{13}C	^{1}H	δ ^{13}C (ppm)	δ ^{1}H (ppm)	$^{1}J_{CH}$ (Hz)	$^{1}T_{CH}$ (Hz)	$^{1}D_{CH}$ (Hz)	error
29	71	131.8	5.56	153.0709	153.9241	0.853	1.172
28	70	78.83	5.41	153.7944	152.5751	-1.219	1.851
20	58	125.61	5.24	157.9970	160.5807	2.584	2.601
50	91,92,93	56.55	3.98	434.2981	430.1824	-1.3719	1.335
19	57	47.99	3.26	134.1185	125.4268	-8.6917	3.256
49	90	30.84	3.15	128.1854	128.1118	-0.0736	1.049
36	80,81,82	33.8	3	413.0035	424.4382	3.8116	1.673
21	59	30.83	2.94	180.3363	189.1497	8.8134	2.937
34a	78,79	27.34	1.6	249.8059	240.1031	-4.8514	2.839
31b	72,73	22.39	1.47	252.6246	260.6778	4.0266	3.122

Table S28 – Scalar, total and residual dipolar couplings of Simeprevir in PAN/DMSO-d_6 obtained from F1-coupled-HSQC spectrum.

A total of 14 residual dipolar couplings and 28 interproton distances were employed in multiparametric fitting run in StereoFitter. Two best-fitting solution could be identified; both solutions share two conformations, conformers 66 and 68, but differ for the major conformer. Notably, conformers 12 and 38 belong to the same cluster (VI cluster) and both their macrocyclic core and quinoline moiety were superimposable except for the methoxy group. This feature was consistent with back-calculated residual dipolar couplings. In fact, analysing SteroFitter results (Table S29) it was evident that the major difference in back-calculated residual dipolar couplings was methoxy group's $^{1}D_{CH}$.

	Solution	1	2
		Conformer 12: 52.51%	Conformer 38: 52.44%
		Conformer 68: 33.52%	Conformer 68: 34.82%
		Conformer 66: 13.97%	Conformer 66: 12.74%
	AIC	36.527	37.365
	χ^2	32.527	33.365
	Relative probability	1	0.658
	NOE distances (Å)		
Atoms (^{1}H-^{1}H)	Experimental	Calculated	Calculated
57,84	3.727	4.732	4.719
68,84	3.56	3.584	3.584
53,71	4.087	3.734	3.733
71,73	2.691	2.736	2.727
70,83	2.013	2.261	2.265
69,70	2.379	2.404	2.403
58,67	4.312	5.076	5.072
53,58	3.138	3.014	3.018
55,58	2.623	2.647	2.641
72,77	2.434	2.881	2.864
77,79	2.735	2.841	2.834
(74,75),77	3.12	2.849	2.851
57,67	4.739	4.217	4.222
56,57	2.326	2.233	2.232
57,68	2.619	2.699	2.699
60,(80,81,82)	2.291	2.623	2.622
60,69	2.444	2.451	2.45
69,(80,81,82)	2.683	2.841	2.846
76,(80,81,82)	2.493	2.877	2.876
78,(80,81,82)	2.436	3.571	3.556
69,83	2.837	3.182	3.225
76,78	2.589	2.374	2.376
71,72	2.887	2.566	2.569
53,72	1.952	2.758	2.743
72,(74,75)	2.714	2.744	2.745
(74,75),78	2.208	2.446	2.446
68,83	3.213	4.406	4.426
55,67	3.511	3.918	3.912
	RDCs (Hz)		
Atoms (^{13}C-^{1}H)	Experimental	Calculated	Calculated
41,84	31.483	31.897	31.613
43,85	13.127	13.419	13.301
48,89	27.628	27.004	27.524
38,83	3.062	1.713	1.703
29,71	0.853	1.633	1.727
28,70	-1.219	-1.407	-1.702
20,58	2.584	2.915	2.866
(50,91) (50,92) (50,93)	-1.372	-1.92	0.347
19,57	-8.692	-7.506	-7.771
49,90	-0.074	-0.71	-0.842
(36,80) (36,81) (36,82)	3.812	4.205	4.385
21,59	8.813	9.571	8.953
(34,78) (34,79)	-4.851	-5.22	-5.162
(31,72) (31,73)	4.027	2.416	2.475

Table S29 – SteroFitter results of multiparametric fitting of Simeprevir in DMSO-d_6.

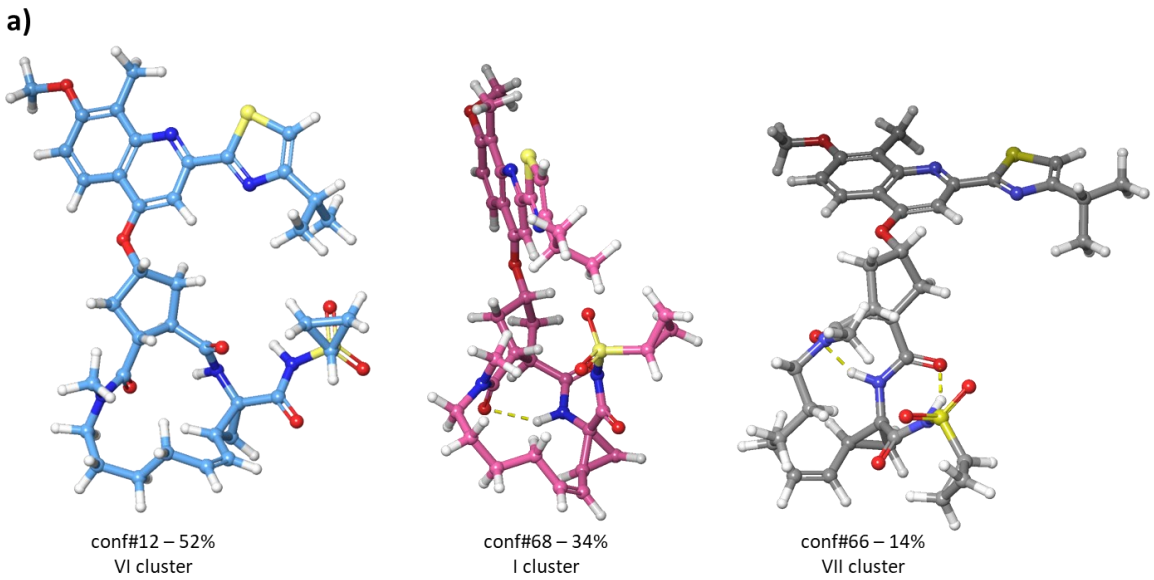


Figure 5.52 – Simeprevir conformers resulting from multiparametric fitting of NMR experimental data. Population percentages and clusters are shown.

By superimposing the 3 conformations with the crystal structure (PDB: 3KEE)[96], conformer 68 showed the lowest RMSD (0.6 Å) with respect to the crystalline form (Figure 5.53). Hence, one of the conformations found by means of NMR-based conformational analysis in solution is similar to the crystallized one. Even if some atoms of the macrocyclic core presented some deviations from the crystal, differences in the acylsulfonamide motif and translations in the thiazole of the quinoline moiety, conformer 68 is almost superimposable to the crystal structure.

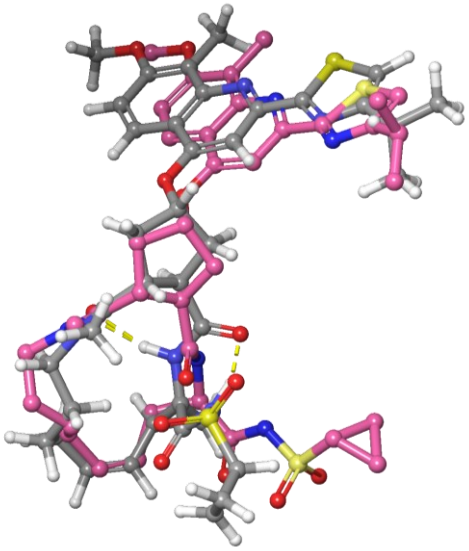


Figure 5.53 – Superposition between conformer 68 (grey) vs crystal structure (pink).

5.3.4 Conformational analysis of Simeprevir in $CDCl_3$

5.3.4.1 NMR Assignment

NMR assignments were carried out through the analysis of ^{1}H, COSY, TOCSY, ^{1}H-^{13}C HSQC, ROESY, ^{1}H-^{13}C HMBC, ^{1}H-^{15}N HMBC spectra recorded on a 600 MHz Bruker AVANCE III HD spectrometer equipped with cryogenic TCI probe.

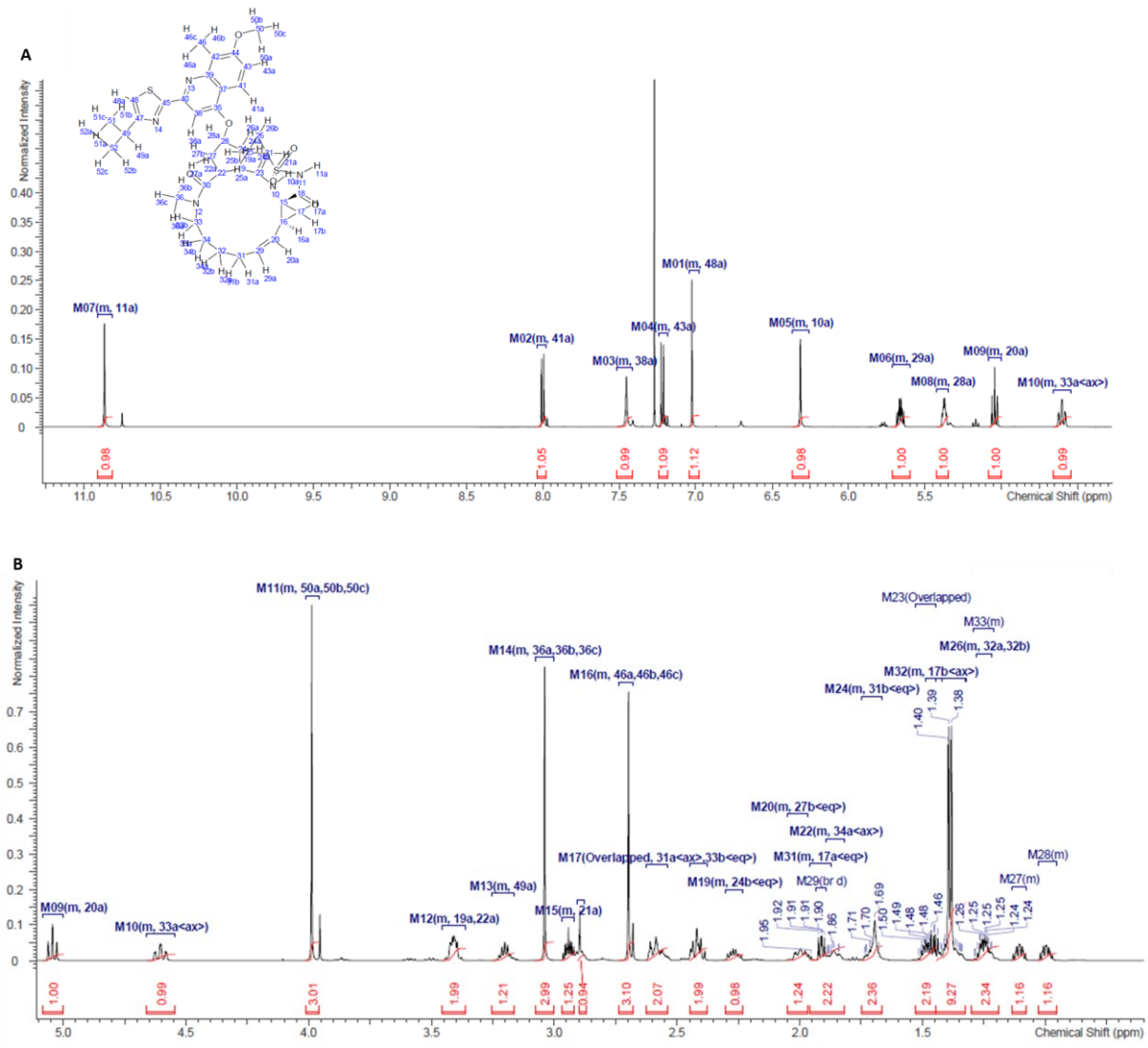


Figure 5.54 – ^{1}H NMR spectrum of Lorlatinib in $CDCl_3$.

^{1}H NMR spectrum of Simeprevir in $CDCl_3$ presented two sets of peaks, probably due to the cis/trans isomerization of NH-10 (0.16%). Moreover, many peaks of the aliphatic region overlap, leading to a loss of some information, e.g. some interproton distances.

5.3.4.2 Analysis by means of interproton distances

Interproton distances were measured through the acquisition of 1D-EASY-ROESY spectra; ten spectra for each single resonant peak varying the mixing time from 50 ms to 500 ms. T_1 inversion-recovery spectrum was recorded to calculate d1 delay; the aromatic proton H48 showed the largest T_1 relaxation (3.07 s). Therefore, a delay of 15 s was used. Spectra derived from 14 resonant frequencies were analysed and distance measured using geminal protons H33a,b as reference.

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	Atom # i	Atom # j	σ_{ij}	R^2	distance r_{ij} (Å)	average distance r_{ij} (Å)
H48	H49	7.04	3.21	89	90	0.0115	0.997	3.512	**3.476**
H48	H51, H52	7.04	1.4	89	94,95,96 97,98,99	0.0454	0.996	2.794	
H38	H28	7.427	5.39	83	70	0.1459	0.989	2.300	**2.238**
H38	H24a	7.427	2.4	83	61	0.0242	0.987	3.102	
H10	H11	6.31	10.89	56	67	0.0142	0.955	**3.391**	
H29	H31b	5.67	1.72	71	73	0.0644	0.994	**2.636**	
H29	H34b	5.67	1.38	71	79	0.0263	0.996	**3.060**	
H28	H41	5.386	8.04	70	84	0.0031	0.991	**4.369**	
H28	H38	5.386	7.47	70	83	0.2039	0.999	2.175	-
H28	H22	5.386	3.42	70	60	0.023	0.983	3.129	
H28	H27a	5.386	2.91	70	69	0.0759	0.996	2.564	
H28	H24a	5.386	2.44	70	61	0.0265	0.976	3.056	
H28	H24b	5.386	2.28	70	62	0.1056	0.981	2.427	
H20	H11	5.055	10.89	58	67	0.0049	0.962	**4.048**	
H20	H21a	5.055	2.95	58	59	0.0077	0.986	3.755	**3.869**
H20	H16	5.055	2.43	58	53	0.0213	0.993	**3.169**	
H20	H17a	5.055	1.92	58	55	0.0438	0.995	2.810	
H33a	H33b	4.61	2.559	77	76	0.6793	0.994	1.780	ref
H33a	H34b	4.61	1.4	77	79	0.054	0.96	**2.714**	
H49	H48	3.21	7.04	90	89	0.013	0.999	3.441	-
H36a,b,c	H33b	3	2.61	80, 81, 82	76	0.0667	0.996	**2.620**	
H36a,b,c	H34a	3	1.88	80, 81, 82	78	0.0774	0.994	2.556	**2.470**
H21	H11	2.95	10.89	59	67	0.0042	0.992	**4.153**	
H21	H20	2.95	5.05	59	58	0.0054	0.993	3.983	-
H27a	H38	2.895	7.46	69	83	0.0125	0.972	**3.463**	
H27a	H28	2.895	5.38	69	70	0.0547	0.991	2.708	-
H24b	H38	2.282	7.46	62	83	0.0133	0.969	3.428	
H24b	H10	2.282	6.37	62	56	0.0043	0.96	4.137	
H24b	H28	2.282	5.4	62	70	0.1593	0.995	2.266	
H24b	H24a	2.282	2.44	62	61	0.7289	0.998	1.759	
H27b	H41	2.01	8.01	68	84	0.0096	0.983	**3.619**	
H27b	H38	2.01	7.46	68	83	0.0102	0.962	**3.583**	

Proton i	Proton j	δ_i (ppm)	δ_j (ppm)	Atom # i	Atom # j	σ_{ij}	R^2	distance r_{ij} (Å)	average distance r_{ij} (Å)
H27b	H28	2.01	5.38	68	70	0.0136	0.95	3.415	
H34a	H36a,b,c	1.86	3	78	80,81,82	0.1175	0.995	2.384	-
H34a	H33b	1.86	2.61	78	76	0.066	0.991	**2.625**	
H34a	H32a,b	1.86	1.2	78	74,75	0.0785	0.986	**2.550**	

Table S30 – Interproton distances of Simeprevir in $CDCl_3$.

A total of 20 interatomic distances were submitted to StereoFitter calculation. The analysis resulted in multiple likely solutions with very low AIC, χ^2 and high relative probability (greater than 0.7%). AIC values were comparable between all solutions; hence a single solution could not be identified. However, by comparing different models it was noted that all solutions consist of three conformers belonging to the same three clusters (i.e. II, VI, X). Superposing conformers of different solutions but within the same cluster, they presented very small deviations, not considering the flexibility of the sidechain. Therefore, it was concluded that the first solution could be considered as representative of all the other conformations found in chloroform. All conformers (see Figure 5.55) of different solutions sharing the same cluster presented the same orientation of the quinoline moiety with representative conformers; they only differ from each other for the acylsulfonamide moiety, thiazole, methoxy and isopropyl groups' arrangement.

	Solution	1	2	3	4	5
		Conformer 32: 53.20%	Conformer 32: 53.15%	Conformer 7: 34.07%	Conformer 32: 53.06%	Conformer 32: 53.03%
		Conformer 22: 33.99%	Conformer 22: 34.04%	Conformer 32: 53.11%	Conformer 22: 37.77%	Conformer 22: 37.76%
		Conformer 181: 12.81%	Conformer 181: 12.81%	Conformer 181: 12.82%	Conformer 115: 9.17%	Conformer 115: 9.20%
	AIC	9.517	9.522	9.525	9.850	9.852
	χ^2	5.517	5.522	5.525	5.850	5.852
	Relative probability	1	0.997	0.996	0.846	0.845
	NOE distances (Å)					
Atoms (^{13}C-^{1}H)	Experimental	Calculated	Calculated	Calculated	Calculated	Calculated
89,90	3.476	3.67	3.671	3.671	3.36	3.36
70,83	2.238	2.215	2.215	2.214	2.217	2.217
56,67	3.391	3.309	3.309	3.309	3.336	3.334
71,73	2.636	2.624	2.624	2.624	2.632	2.632
71,79	3.06	3.292	3.291	3.291	3.252	3.252
70,84	4.369	4.408	4.41	4.41	4.407	4.407
69,70	2.636	2.571	2.57	2.57	2.589	2.589
58,67	4.048	4.259	4.259	4.259	4.251	4.251
58,59	3.869	3.694	3.694	3.695	3.688	3.688
53,58	3.169	3.085	3.085	3.085	3.082	3.082
55,58	2.81	2.64	2.639	2.639	2.652	2.652
77,78	2.714	2.627	2.627	2.627	2.636	2.636
76,(80,81,82)	2.62	2.834	2.833	2.833	2.815	2.815
79,(80,81,82)	2.47	2.988	2.989	2.989	3.059	3.059
59,67	4.153	3.892	3.892	3.892	3.894	3.894
69,83	3.463	3.354	3.354	3.353	3.38	3.382
68,84	3.619	3.797	3.799	3.802	3.753	3.753
68,83	3.583	3.757	3.757	3.756	4.155	4.157
76,79	2.625	2.623	2.623	2.623	2.645	2.645
(74,75),79	2.55	2.746	2.746	2.746	2.758	2.758

Table S31 – StereoFitter results of interproton distances' fitting of Simeprevir in $CDCl_3$. Results presenting a relative probability greater than 0.84 are shown.

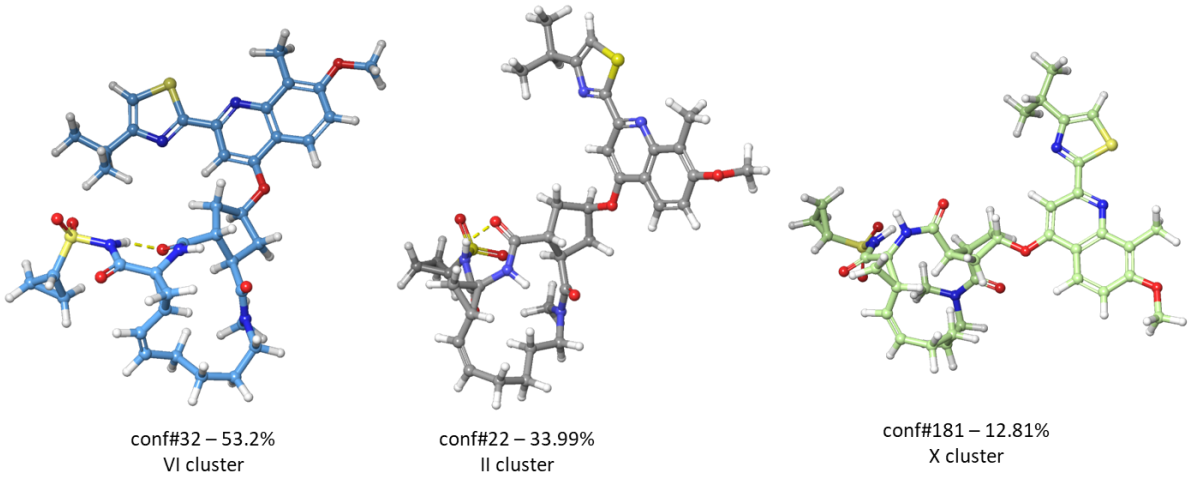


Figure 5.55 – Conformational ensemble of Simeprevir in $CDCl_3$ derived from NMR experimental data (solution #1). These conformers are representative for all the other solution shown in Table S31.

6. Conclusions

This research project aimed at setting-up a protocol for the NMR-based conformational analysis of free ligands in solution. In this thesis it was shown that conformational ensemble can be obtained from routinely available NMR spectra combined with less conventional NMR experiments such as the measurements of anisotropic parameters requiring weak alignment of the analyte.

Interactions derived from alignment such as Residual Dipolar Couplings (RDCs) provide global spatial arrangement and complement the more common NMR local information resulting from interproton distances and J couplings.

Three pharmacologically active macrocyclic compounds were examined: Lorlatinib, Pacritinib and Simeprevir. They present high affinity for their targets; hence it can be assumed that they could be already prearranged into the bioactive conformation in solution.

Internuclear distances were obtained through the acquisition of 1D NOE or 1D-EASY-ROESY spectra analysed through the normalization approach called PANIC using 10 different mixing times. In some cases, distances derived from build-up rates of PANIC methods were compared with the ones measured from 2D-EASY-ROESY or 1D-EASY-ROESY using a single mixing time.

As concerns RDCs values, they were extracted from various spectra such as HD-J-HSQC, F1-coupled-HSQC, $^{1}J_{CH}$-resolved-HSQC, HSQC-HECADE.

In addition to NMR data, quantum mechanical calculations of chemical shifts were integrated into the conformational analysis as far as they provide useful information such as the assignment of diastereotopic protons.

All this experimental information were fitted to back-calculated data obtained from theoretical models using StereoFitter. Hence, for each particular property, the individual contributions of each conformation were calculated. The goodness-of-fit was evaluated through χ^2 and AIC values.

As regard to Lorlatinib, 14 conformations were employed in the analysis performed in polar DMSO-d_6 environment. Interproton distances did not provide a unique and unambiguous outcome. It resulted in more than one solution having a high relative

probability (near to 1) and similar AIC numbers (between 3.48 and 4.48). Nevertheless, distances analysis was extremely effective in assigning diastereotopic protons helping in discriminating the correct conformation resulted from QM chemical shifts predictions. Multiparametric fitting (i.e. interproton distances and RDCs-based analysis) led to two likely solutions that were very similar sharing the same major conformer in agreement with QM predictions. As a final step best-fitting conformer was compared to the crystal structure exhibiting a high degree of superposition ($RMSD_{heavy\ atoms}$ = 0.0348 Å).

Conformational analysis of Lorlatinib was also performed in an apolar solvent ($CDCl_3$) to study if a shift in solvent polarity might induce a change in the conformational arrangement of the molecule. Internuclear distances-based and residual dipolar couplings-based analyses using PMMA gel led to several conformational models if considered separately. It was only when the information coming from both constraints were integrated that a unique solution was obtained. Two conformations resulted to be the same of the ones found in DMSO-d_6 while a third different conformer was added to the ensemble. It was conformer 9 presenting a particular conformation in which the carbonyl group might be less exposed to the solvent because of the pyridine ring shielding.

As concerns conformational analysis of Pacritinib, 63 different conformations for the macrocyclic core were utilized. Local-character information provided by distances was not discriminant leading to 10 high-probable solutions. The analyte was then aligned using PAN gel in DMSO-d_6 allowing the acquisition of very well resolved $^1D_{CH}$. Nevertheless, some attempts were made to measure homonuclear and heteronuclear long-range residual dipolar couplings, but they weren't successful. $^nD_{CH}$ were extracted from HSQC-HECADE, but, since the experimental errors were greater than or close to the absolute values of RDC, they were not included into the analysis. On the other hand, signals coming from $^nD_{HH}$ were not detected. Anyway, the number of $^1D_{CH}$ was sufficient for determining the alignment tensor to obtain calculated RDCs. Multiparametric fitting resulted in three probable solutions sharing three conformations out of four. Conformational analysis of Pacritinib wasn't carried out in apolar solvents because of solubility issues. Analyses by means of other polar solvents such as D_2O and methanol-d_3 will be performed in the near future using commercially available or in-house developed liquid crystals. D_2O represents the ideal solvent for

this analysis to mimic body fluid environment. In addition, further investigations on NMR acquisition of long-range Residual Dipolar Couplings ($^{n}D_{CH,}$ $^{n}D_{HH}$) recorded by means of polymeric gels are needed in order to improve the quality of acquired spectra since in literature there are few case studies of long-range Residual Dipolar Couplings measured in this kind of alignment tensor. This type of angular constraints would add more information to the fitting analysis.

Regarding the analysis of Simeprevir, it was conducted in both polar and apolar environments employing 212 conformations divided into 10 clusters according to the RMSD of macrocyclic core. Similarly to Pacritinib interproton distances-based analysis failed in providing a single solution. However, distances coupled with QM chemical calculations of chemical shift were found to be useful for the assignments of diastereotopic methylene protons. RDCs-based analysis combined with interproton distances resulted in three principal conformations: one, in particular, was nearly superimposable to the crystal structure of Simeprevir complexed with hepatitis C NS3/4A.

The analysis of Simeprevir in chloroform-d was carried out by means of interproton distances only. In fact, no anisotropic measurement was possible since the analyte appeared not to diffuse into the various PPMA gels that were tested. However, information provided by distances seemed to be sufficient in this case. Single-parametric fitting led to multiple solutions composed of conformations belonging to the same three clusters (i.e. II, VI, X). Results showed that in all the best-fitting solutions conformers shared a similar orientation of the quinoline moiety within the same cluster. Therefore, it was concluded that conformers of the first solution could be considered as representative of all the other solutions.

Overall, throughout this research project a general workflow for the NMR-based conformational analysis has been defined. Depending on the molecule under investigation, it can be considered more appropriate to focus on the analysis of certain NMR constraints or predictions rather than others. In general, the use of both local (i.e. interproton distances, J-couplings and chemical shift) and global (i.e. residual dipolar couplings) information allows a clearer identification of conformational ensembles in solution avoiding ambiguous results. In some circumstances the identification of specific fingerprints (i.e. considering whether or not some resonance

frequencies can be regarded as diagnostic peaks since three-dimensional conformations affect the chemical shifts) or geometrical constraints (dihedral angles, interproton distances or angular descriptors) plays a crucial role in establishing an effective and robust workflow for conformational assessment.

References

1. Garbisch EW, Vol J, Edgar Garbisch BW, et al. *Conformations. V. Conformational Analysis of 2-Bromocyclohexanone by Nuclear Magnetic Resonance Spectroscopy*. Vol 38.; 1962. https://pubs.acs.org/sharingguidelines

2. Blackledge M. Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Prog Nucl Magn Reson Spectrosc*. 2005;46(1):23-61. doi:10.1016/j.pnmrs.2004.11.002

3. Bryce DL, Grishaev A, Bax A. Measurement of Ribose Carbon Chemical Shift Tensors for A-form RNA by Liquid Crystal NMR Spectroscopy. *J Am Chem Soc*. 2005;127(20):7387-7396. doi:10.1021/ja051039c

4. Azurmendi HF, Bush A. *Conformational Studies of Blood Group A and Blood Group B Oligosaccharides Using NMR Residual Dipolar Couplings*. Vol 337.; 2002. www.elsevier.com/locate/carres

5. Reinscheid UM, Farjon J, Radzom M, et al. Effect of the solvent on the conformation of a depsipeptide: NMR-derived solution structure of hormaomycin in DMSO from residual dipolar couplings in a novel DMSO-compatible alignment medium. *ChemBioChem*. 2006;7(2):287-296. doi:10.1002/cbic.200500277

6. Klages J, Neubauer C, Coles M, Kessler H, Luy B. Structure refinement of cyclosporin A in chloroform by using RDCs measured in a stretched PDMS-gel. *ChemBioChem*. 2005;6(9):1672-1678. doi:10.1002/cbic.200500146

7. Fredersdorf M, Kurz M, Bauer A, et al. Conformational Analysis of an Antibacterial Cyclodepsipeptide Active against Mycobacterium tuberculosis by a Combined ROE and RDC Analysis. *Chemistry - A European Journal*. 2017;23(24):5729-5735. doi:10.1002/chem.201605143

8. Lancefield CS, Slawin AMZ, Westwood NJ, Lebl T. The use of residual dipolar coupling for conformational analysis of structurally related natural products. *Magnetic Resonance in Chemistry*. 2015;53(6):467-475. doi:10.1002/mrc.4213

9. Huben K, Jewgiński M, Pabis A, Paluch P, Luy B, Jankowski S. The structure of cyclolinopeptide a in chloroform refined by RDC measurements. *Journal of Peptide Science*. 2014;20(11):901-907. doi:10.1002/psc.2683

10. Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol*. 2009;5(11):789-796. doi:10.1038/nchembio.232

11. Blundell CD, Packer MJ, Almond A. Quantification of free ligand conformational preferences by NMR and their relationship to the bioactive conformation. *Bioorg Med Chem*. 2013;21(17):4976-4987. doi:10.1016/j.bmc.2013.06.056

12. Doak BC, Over B, Giordanetto F, Kihlberg J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem Biol*. 2014;21(9):1115-1142. doi:10.1016/j.chembiol.2014.08.013

13. Garcia Jimenez D, Poongavanam V, Kihlberg J. Macrocycles in Drug Discovery–Learning from the Past for the Future. *J Med Chem*. 2023;66(8):5377-5396. doi:10.1021/acs.jmedchem.3c00134

14. Karplus M. Contact Electron-Spin Coupling of Nuclear Magnetic Moments. *J Chem Phys*. 1959;30(1):11-15. doi:10.1063/1.1729860

15. Saupe A, Englert G. High-Resolution Nuclear Magnetic Resonance Spectra of Orientated Molecules. *Phys Rev Lett*. 1963;11(10):462-464. doi:10.1103/PhysRevLett.11.462

16. Meddour A., Canet I., Loewenstein A., Péchiné J.M., Courtieu J. *Observation of Enantiomers, Chiral by Virtue of Isotopic Substitution, through Deuterium NMR in a Polypeptide Liquid A Total of 120 Mg of PBLG*. Vol 116. Pergamon Press; 1994. https://pubs.acs.org/sharingguidelines

17. Schuetz A, Junker J, Leonov A, Lange OF, Molinski TF, Griesinger C. Stereochemistry of Sagittamide A from Residual Dipolar Coupling Enhanced NMR. *J Am Chem Soc*. 2007;129(49):15114-15115. doi:10.1021/ja075876l

18. Farès C, Hassfeld J, Menche D, Carlomagno T. Simultaneous determination of the conformation and relative configuration of archazolide A by using nuclear overhauser effects, J couplings, and residual dipolar couplings. *Angewandte Chemie - International Edition*. 2008;47(20):3722-3726. doi:10.1002/anie.200800225

19. Gil RR, Gayathri C, Tsarevsky N V., Matyjaszewski K. Stretched Poly(methyl methacrylate) Gel Aligns Small Organic Molecules in Chloroform. Stereochemical Analysis and Diastereotopic Proton NMR Assignment in Ludartin Using Residual Dipolar Couplings and ^{3}J Coupling Constant Analysis. *J Org Chem*. 2008;73(3):840-848. doi:10.1021/jo701871g

20. Efimov S V., Karataeva FK, Aganov A V., Berger S, Klochkov V V. Spatial structure of cyclosporin A and insight into its flexibility. *J Mol Struct*. 2013;1036:298-304. doi:10.1016/j.molstruc.2012.11.005

21. Tzvetkova P, Sternberg U, Gloge T, Navarro-Vázquez A, Luy B. Configuration determination by residual dipolar couplings: Accessing the full conformational space by molecular dynamics with tensorial constraints. *Chem Sci*. 2019;10(38):8774-8791. doi:10.1039/c9sc01084j

22. Kramer F, Deshmukh M V., Kessler H, Glaser SJ. Residual dipolar coupling constants: An elementary derivation of key equations. *Concepts Magn Reson Part A Bridg Educ Res*. 2004;21(1):10-21. doi:10.1002/cmr.a.20003

23. Thiele CM. Residual dipolar couplings (RDCs) in organic structure determination. *European J Org Chem*. 2008;(34):5673-5685. doi:10.1002/ejoc.200800686

24. Losonczi JA, Andrec M, Fischer MWF, Prestegard JH. *Order Matrix Analysis of Residual Dipolar Couplings Using Singular Value Decomposition*.; 1999. http://www.idealibrary.com

25. Thiele CM, Schmidts V, Böttcher B, et al. On the Treatment of Conformational Flexibility when Using Residual Dipolar Couplings for Structure Determination. *Angewandte Chemie*. 2009;121(36):6836-6840. doi:10.1002/ange.200902398

26. Zweckstetter M. NMR: prediction of molecular alignment from structure using the PALES software. *Nat Protoc*. 2008;3(4):679-690. doi:10.1038/nprot.2008.36

27. Valafar H, Prestegard JH. REDCAT: a residual dipolar coupling analysis tool. *Journal of Magnetic Resonance*. 2004;167(2):228-241. doi:10.1016/j.jmr.2003.12.012

28. Berger R, Fischer C, Klessinger M. Calculation of the Vibronic Fine Structure in Electronic Spectra at Higher Temperatures. 1. Benzene and Pyrazine. *J Phys Chem A*. 1998;102(36):7157-7167. doi:10.1021/jp981597w

29. Schmidts V. *PhD Thesis, Technische Universität, Darmstadt, 2013.*

30. Navarro-Vázquez A. MSpin-RDC. A program for the use of residual dipolar couplings for structure elucidation of small molecules. *Magnetic Resonance in Chemistry*. 2012;50:S73-S79. doi:10.1002/mrc.3905

31. Matsukawa S, Yasunaga H, Zhao C, Kuroki S, Kurosu H, Ando I. Diffusion processes in polymer gels as studied by pulsed field-gradient spin-echo NMR spectroscopy. *Prog Polym Sci*. 1999;24(7):995-1044. doi:10.1016/S0079-6700(99)00022-2

32. Deloche B, Samulski ET. Short-range nematic-like orientational order in strained elastomers: a deuterium magnetic resonance study. *Macromolecules*. 1981;14(3):575-581. doi:10.1021/ma50004a024

33. Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S. Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. *J Biomol NMR*. 2000;18(4):303-309. doi:10.1023/A:1026703605147

34. Naumann C, Bubb WA, Chapman BE, Kuchel PW. Tunable-Alignment Chiral System Based on Gelatin for NMR Spectroscopy. *J Am Chem Soc*. 2007;129(17):5340-5341. doi:10.1021/ja070348v

35. Luy B, Kobzar K, Kessler H. An Easy and Scalable Method for the Partial Alignment of Organic Molecules for Measuring Residual Dipolar Couplings. *Angewandte Chemie International Edition*. 2004;43(9):1092-1094. doi:10.1002/anie.200352860

36. Luy B, Kobzar K, Knör S, Furrer J, Heckmann D, Kessler H. Orientational Properties of Stretched Polystyrene Gels in Organic Solvents and the Suppression of Their Residual ^{1}H NMR Signals. *J Am Chem Soc*. 2005;127(17):6459-6465. doi:10.1021/ja043344o

37. Kaden P, Freudenberger JC, Luy B. Noncovalently and covalently cross-linked polyurethane gels as alignment media and the suppression of residual polymer signals using diffusion-filtered spectroscopy. *Magnetic Resonance in Chemistry*. 2012;50(S1). doi:10.1002/mrc.3887

38. Freudenberger JC, Knör S, Kobzar K, et al. Stretched Poly(vinyl acetate) Gels as NMR Alignment Media for the Measurement of Residual Dipolar Couplings in Polar Organic Solvents. *Angewandte Chemie International Edition*. 2005;44(3):423-426. doi:10.1002/anie.200461241

39. Gil-Silva LF, Santamaría-Fernández R, Navarro-Vázquez A, Gil RR. Collection of NMR Scalar and Residual Dipolar Couplings Using a Single Experiment. *Chemistry – A European Journal*. 2016;22(2):472-476. doi:10.1002/chem.201503449

40. Kummerlöwe G, Auernheimer J, Lendlein A, Luy B. Stretched Poly(acrylonitrile) as a Scalable Alignment Medium for DMSO. *J Am Chem Soc*. 2007;129(19):6080-6081. doi:10.1021/ja071248s

41. Ramirez BE, Bax A. *Modulation of the Alignment Tensor of Macromolecules Dissolved in a Dilute Liquid Crystalline Medium*. Vol 101. Academic Press; 1997.

42. Lorieau JL, Maltsev AS, Louis JM, Bax A. Modulating alignment of membrane proteins in liquid-crystalline and oriented gel media by changing the size and charge of phospholipid bicelles. *J Biomol NMR*. 2013;55(4):369-377. doi:10.1007/s10858-013-9720-3

43. Zweckstetter M, Hummer G, Bax A. Prediction of Charge-Induced Molecular Alignment of Biomolecules Dissolved in Dilute Liquid-Crystalline Phases. *Biophys J*. 2004;86(6):3444-3460. doi:10.1529/biophysj.103.035790

44. Trempe JF, Morin FG, Xia Z, Marchessault RH, Gehring K. Characterization of polyacrylamide-stabilized Pf1 phage liquid crystals for protein NMR spectroscopy. *J Biomol NMR*. 2002;22(1):83-87. doi:10.1023/A:1013832422428

45. Aroulanda C, Lesot P. Molecular enantiodiscrimination by NMR spectroscopy in chiral oriented systems: Concept, tools, and applications. *Chirality*. 2022;34(2):182-244. doi:10.1002/chir.23386

46. Lesot P, Lafon O, Aroulanda C, Dong RY. 2H NMR Studies on Two-Homopolypeptide Lyotropic Enantiodiscriminating Mesophases: Experimental Quantification of Solute–Fiber Affinities. *Chemistry – A European Journal*. 2008;14(13):4082-4092. doi:10.1002/chem.200701543

47. Solgadi A, Jean L, Lasne MC, Rouden J, Courtieu J, Meddour A. NMR in chiral polypeptide liquid crystals: the problem of amines. *Tetrahedron Asymmetry*. 2007;18(13):1511-1516. doi:10.1016/j.tetasy.2007.06.018

48. Jeziorowski S, Thiele CM. Poly-γ-p-Biphenylmethyl-Glutamate as Enantiodifferentiating Alignment Medium for NMR Spectroscopy with Temperature-Tunable Properties. *Chemistry - A European Journal*. 2018;24(58):15631-15637. doi:10.1002/chem.201802921

49. Pake GE. Nuclear resonance absorption in hydrated crystals: Fine structure of the proton line. *J Chem Phys*. 1948;16(4):327-336. doi:10.1063/1.1746878

50. Bielecki A, Pines A. *Pake Patterns from Zero to High Field*. Vol 74.; 1987.

51. Thiele CM, Bermel W. Speeding up the measurement of one-bond scalar (1J) and residual dipolar couplings (1D) by using non-uniform sampling (NUS). *Journal of Magnetic Resonance*. 2012;216:134-143. doi:10.1016/j.jmr.2012.01.008

52. Fehér K, Berger S, Kövér KE. Accurate determination of small one-bond heteronuclear residual dipolar couplings by F1 coupled HSQC modified with a G- BIRD(r) module. *Journal of Magnetic Resonance*. 2003;163(2):340-346. doi:10.1016/S1090-7807(03)00113-7

53. Castañar L, Garcia M, Hellemann E, Nolis P, Gil RR, Parella T. One-Shot Determination of Residual Dipolar Couplings: Application to the Structural Discrimination of Small

Molecules Containing Multiple Stereocenters. *Journal of Organic Chemistry*. 2016;81(22):11126-11131. doi:10.1021/acs.joc.6b02103

54. Kozminski W, Nanz D. *HECADE: HMQC- and HSQC-Based 2D NMR Experiments for Accurate and Sensitive Determination of Heteronuclear Coupling Constants from E.COSY-Type Cross Peaks*. Vol 124.; 1997.

55. Sinnaeve D, Foroozandeh M, Nilsson M, Morris GA. A general method for extracting individual coupling constants from crowded 1H NMR spectra. *Angewandte Chemie - International Edition*. 2016;55(3):1090-1093. doi:10.1002/anie.201508691

56. Parella T, Espinosa JF. Long-range proton-carbon coupling constants: NMR methods and applications. *Prog Nucl Magn Reson Spectrosc*. 2013;73:17-55. doi:10.1016/j.pnmrs.2013.07.001

57. Drain LE. A Direct Method of Measuring Nuclear Spin-Lattice Relaxation Times. *Proceedings of the Physical Society Section A*. 1949;62(5):301-306. doi:10.1088/0370-1298/62/5/306

58. Kessler H, Oschkinat H, Griesinger C, Bermel W. Transformation of homonuclear two-dimensional NMR techniques into one-dimensional techniques using Gaussian pulses. *Journal of Magnetic Resonance (1969)*. 1986;70(1):106-133. doi:10.1016/0022-2364(86)90366-5

59. Thrippleton MJ, Keeler J. Elimination of Zero-Quantum Interference in Two-Dimensional NMR Spectra. *Angewandte Chemie International Edition*. 2003;42(33):3938-3941. doi:10.1002/anie.200351947

60. Thiele CM, Petzold K, Schleucher J. EASY ROESY: Reliable Cross-Peak Integration in Adiabatic Symmetrized ROESY. *Chemistry – A European Journal*. 2009;15(3):585-588. doi:10.1002/chem.200802027

61. David Neuhaus, Michael P. Williamson. *The Nuclear Overhauser Effect in Structural and Conformational Analysis, 2nd Edition*. John Wiley and Sons Ltd; 2000.

62. Hu H, Krishnamurthy K. Revisiting the initial rate approximation in kinetic NOE measurements. *Journal of Magnetic Resonance*. 2006;182(1):173-177. doi:10.1016/j.jmr.2006.06.009

63. bruker.com/it/products-and-solutions/mr/nmr-software/topspin.

64. NMR Workbook Suite, version 2021-2023, Advanced Chemistry Development, Inc. (ACD/Labs), Toronto, ON, Canada, www.acdlabs.com.

65. Mohamadi F, Richards NGJ, Guida WC, et al. Macromodel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J Comput Chem*. 1990;11(4):440-467. doi:10.1002/jcc.540110405

66. Bochevarov AD, Harder E, Hughes TF, et al. Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int J Quantum Chem*. 2013;113(18):2110-2142. doi:10.1002/qua.24481

67. Gil R, Navarro-Vázquez A, Lopez I, Gil-Silva L. Stereofitter new multiparameter 3d structure elucidator.

68. Willcott MR. MestRe Nova. *J Am Chem Soc*. 2009;131(36):13180-13180. doi:10.1021/ja906709t

69. Cicero DO, Barbato G, Bazzo R. *NMR Analysis of Molecular Flexibility in Solution: A New Method for the Study of Complex Distributions of Rapidly Exchanging Conformations. Application to a 13-Residue Peptide with an 8-Residue Loop*. Vol 117.; 1995. https://pubs.acs.org/sharingguidelines

70. Akaike H. A New Look at the Statistical Model Identification. *IEEE Trans Automat Contr*. 1974;(6).

71. fda.gov/drugs/resources-information-approved-drugs/fda-approves-lorlatinib-metastatic-alk-positive-nsclc.

72. ema.europa.eu/en/medicines/human/EPAR/lorviqua.

73. fda.gov/drugs/resources-information-approved-drugs/fda-approves-lorlatinib-metastatic-alk-positive-nsclc.

74. Johnson TW, Richardson PF, Bailey S, et al. Discovery of (10 R)-7-Amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17- tetrahydro- 2H -8,4-(metheno)pyrazolo[4,3- h][2,5,11]- benzoxadiazacyclotetradecine-3-carbonitrile (PF-06463922), a macrocyclic inhibitor of anaplastic lymphoma kinase (ALK) and c-ros oncogene 1 (ROS1) with preclinical brain exposure and broad-spectrum potency against ALK-resistant mutations. *J Med Chem*. 2014;57(11):4720-4744. doi:10.1021/jm500261q

75. Peng C, Atilaw Y, Wang J, et al. Conformation of the Macrocyclic Drug Lorlatinib in Polar and Nonpolar Environments: A MD Simulation and NMR Study. *ACS Omega*. 2019;4(26):22245-22250. doi:10.1021/acsomega.9b03797

76. Chang G, Guida WC, Still WC. An internal-coordinate Monte Carlo method for searching conformational space. *J Am Chem Soc*. 1989;111(12):4379-4386. doi:10.1021/ja00194a035

77. Goodman JM, Still WC. Searching Conformation Space. *J Comput Chem*. 1991;12(1110).

78. Jorgensen WL, Tirado-Rives J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc*. 1988;110(6):1657-1666. doi:10.1021/ja00214a001

79. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc*. 1996;118(45).

80. Qiu D, Shenkin PS, Hollinger FP, Still WC. The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii. *J Phys Chem A*. 1997;101(16):3005-3014. doi:10.1021/jp961992r

81. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J Phys Chem*. 1994;98(45):11623-11627. doi:10.1021/j100096a001

82. Becke AD. Correlation energy of an inhomogeneous electron gas: A coordinate-space model. *J Chem Phys*. 1988;88(2):1053-1062. doi:10.1063/1.454274

83. Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B*. 1988;37(2):785-789. doi:10.1103/PhysRevB.37.785

84. Amovilli C, Barone V, Cammi R, et al. Recent Advances in the Description of Solvent Effects with the Polarizable Continuum Model. In: ; 1998:227-261. doi:10.1016/S0065-3276(08)60416-5

85. Martin NH, Allen NW, Brown JD, Kmiec DM, Vo L. An NMR shielding model for protons above the plane of a carbonyl group. *J Mol Graph Model*. 2003;22(2):127-131. doi:10.1016/S1093-3263(03)00152-9

86. McTigue M, Deng YL, Liu W, Brooun A, Stewart A. rcsb.org/structure/5AA9.

87. Verdier L, Sakhaii P, Zweckstetter M, Griesinger C. Measurement of long range H, C couplings in natural products in orienting media: A tool for structure elucidation of natural products. *Journal of Magnetic Resonance*. 2003;163(2):353-359. doi:10.1016/S1090-7807(03)00063-6

88. Rehman Z, Franks WT, Nguyen B, Schmidt HF, Scrivens G, Brown SP. Discovering the Solid-State Secrets of Lorlatinib by NMR Crystallography: To Hydrogen Bond or not to Hydrogen Bond. *J Pharm Sci*. 2023;112(7):1915-1928. doi:10.1016/j.xphs.2023.02.022

89. fda.gov/drugs/news-events-human-drugs/fda-approves-drug-adults-rare-form-bone-marrow-disorder.

90. Watts KS, Dalal P, Tebben AJ, Cheney DL, Shelley JC. Macrocycle Conformational Sampling with MacroModel. *J Chem Inf Model*. 2014;54(10):2680-2696. doi:10.1021/ci5001696

91. Carvalho DS, da Silva DGB, Hallwass F, Navarro-Vázquez A. Chemically cross-linked polyacrylonitrile. A DMSO compatible NMR alignment medium for measurement of residual dipolar couplings and residual chemical shift anisotropies. *Journal of Magnetic Resonance*. 2019;302:21-27. doi:10.1016/j.jmr.2019.03.005

92. accessdata.fda.gov/drugsatfda_docs/nda/2013/205123orig1s000toc.cfm.

93. Chen IJ, Foloppe N. Tackling the conformational sampling of larger flexible compounds and macrocycles in pharmacology and drug discovery. *Bioorg Med Chem*. 2013;21(24):7898-7920. doi:10.1016/j.bmc.2013.10.003

94. Parish C, Lombardi R, Sinclair K, et al. A comparison of the Low Mode and Monte Carlo conformational search methods. *J Mol Graph Model*. 2002;21(2):129-150. doi:10.1016/S1093-3263(02)00144-4

95. Stevens ES, Sugawara N, Bonora GM, Toniolo C. Conformational analysis of linear peptides. 3. Temperature dependence of NH chemical shifts in chloroform. *J Am Chem Soc*. 1980;102(23):7048-7050. doi:10.1021/ja00543a025

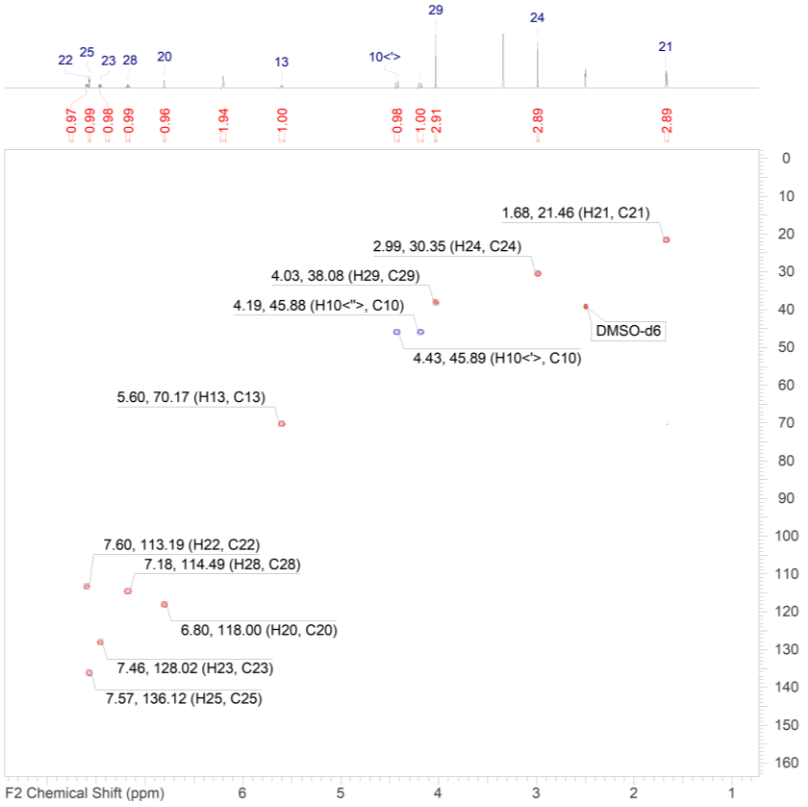
96. Lindberg JD, Nystrom S, Cummings MD. doi.org/10.2210/pdb3KEE/pdb.

List of Figures

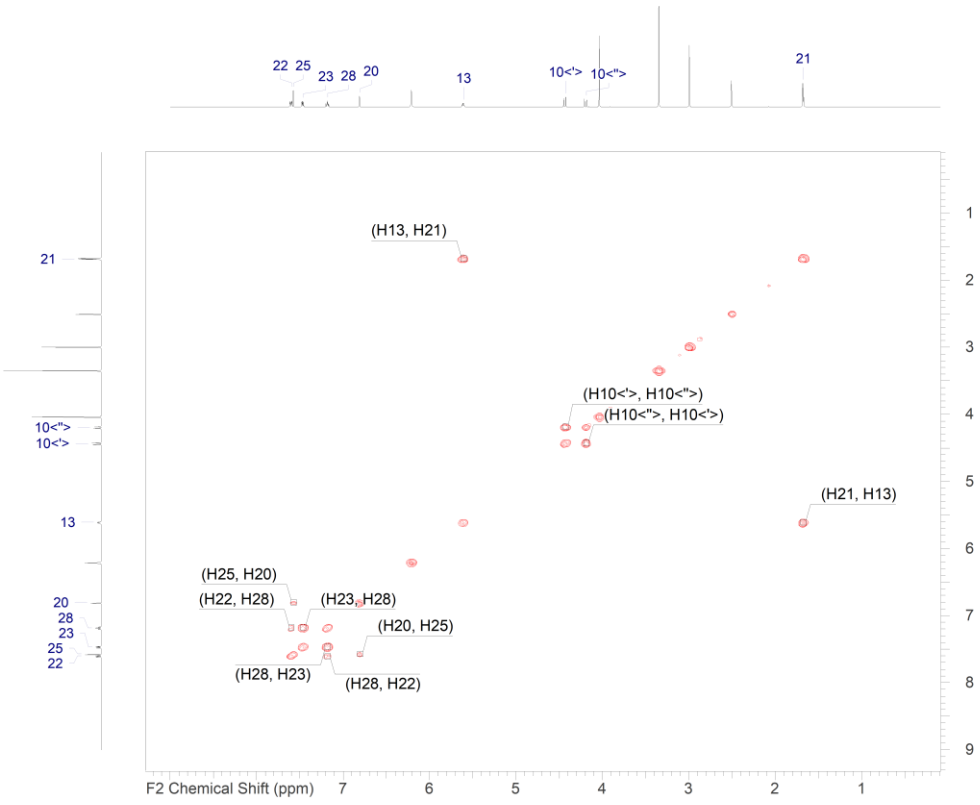
List of Tables

Appendix

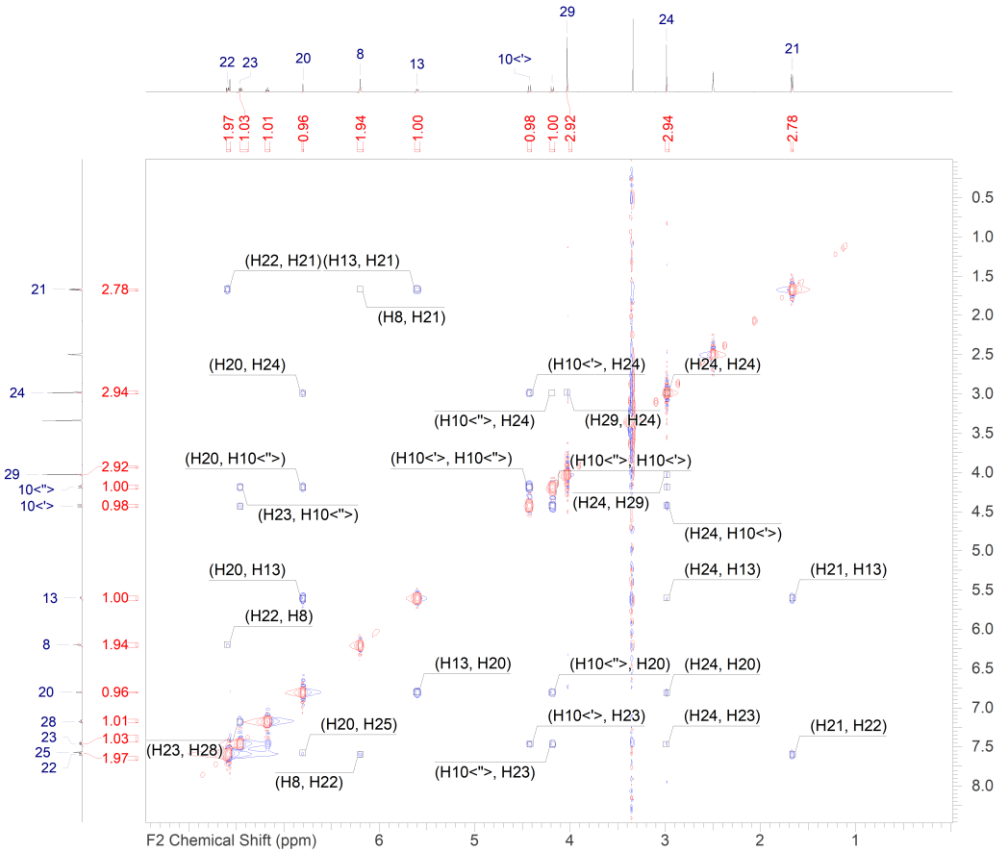
A1. Lorlatinib spectra in DMSO-d_6.



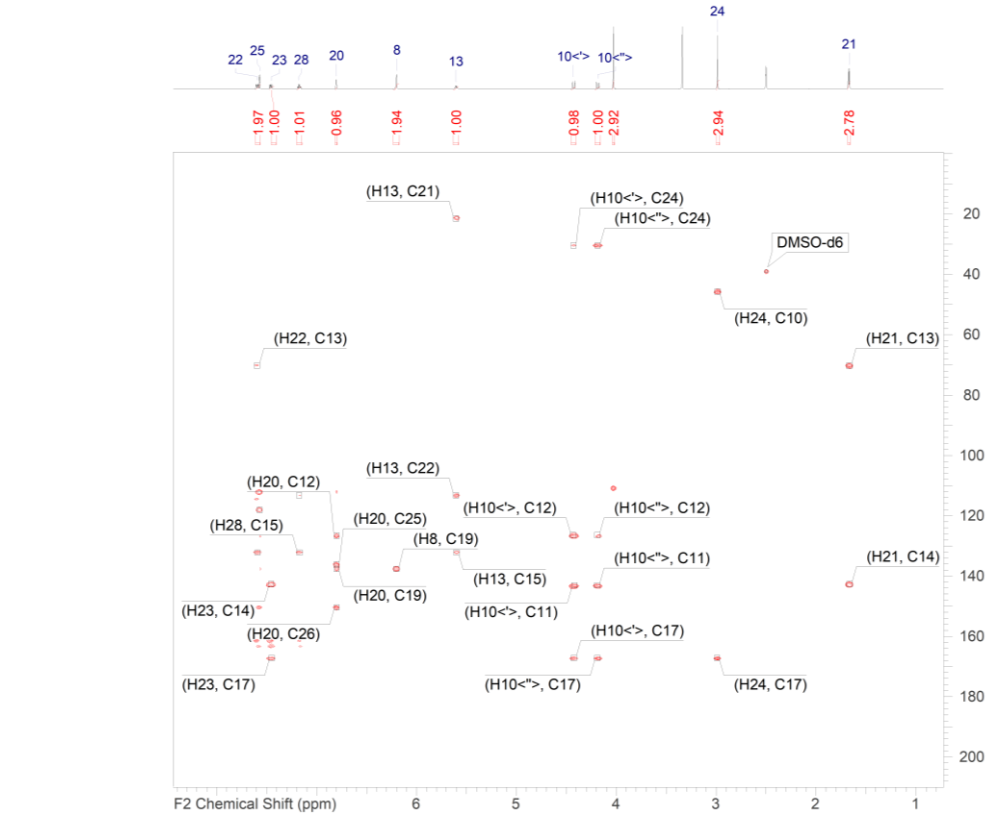
^{1}H-^{13}C HSQC-DEPT NMR spectrum of Lorlatinib in DMSO-d_6.



^{1}H-^{1}H COSY NMR spectrum of Lorlatinib in DMSO-d_6.

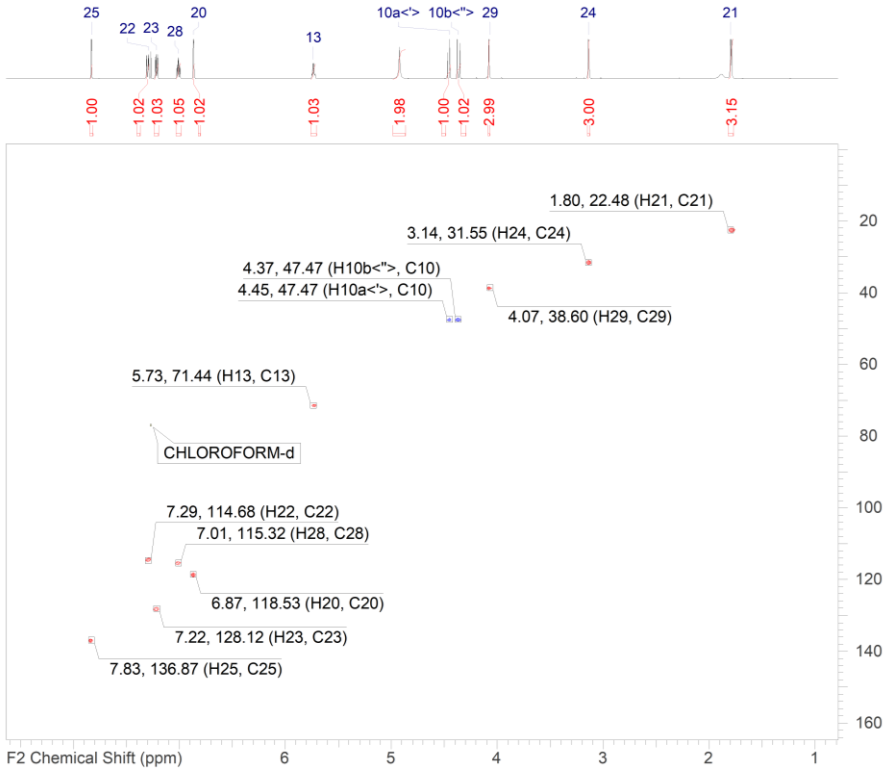


ROESY NMR spectrum of Lorlatinib in DMSO-d_6.

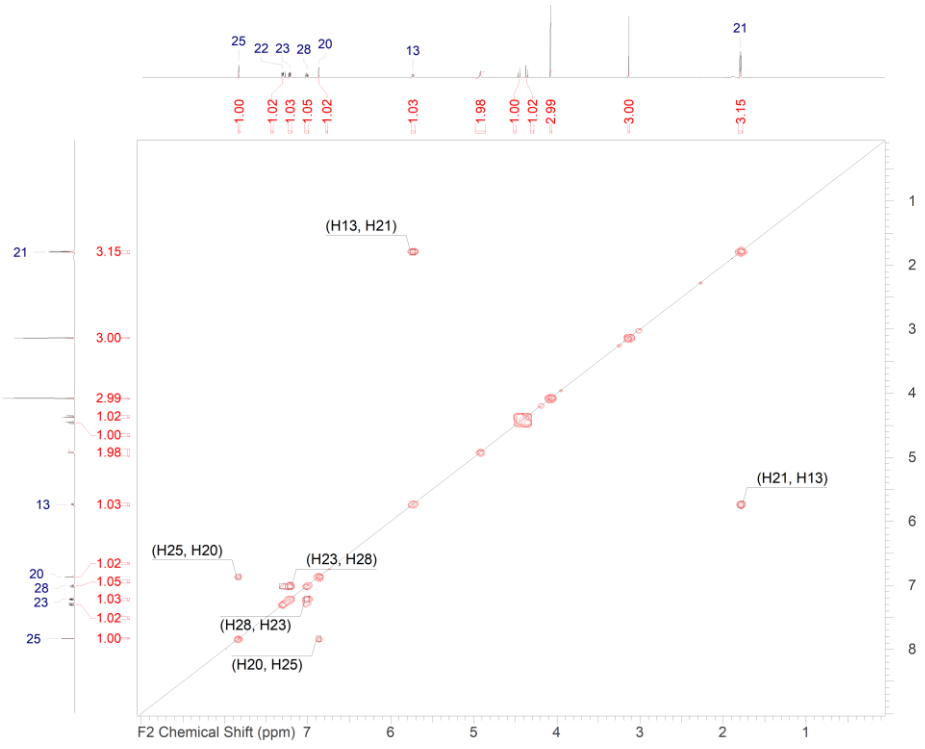


^{1}H-^{13}C HMBC NMR spectrum of Lorlatinib in DMSO-d_6.

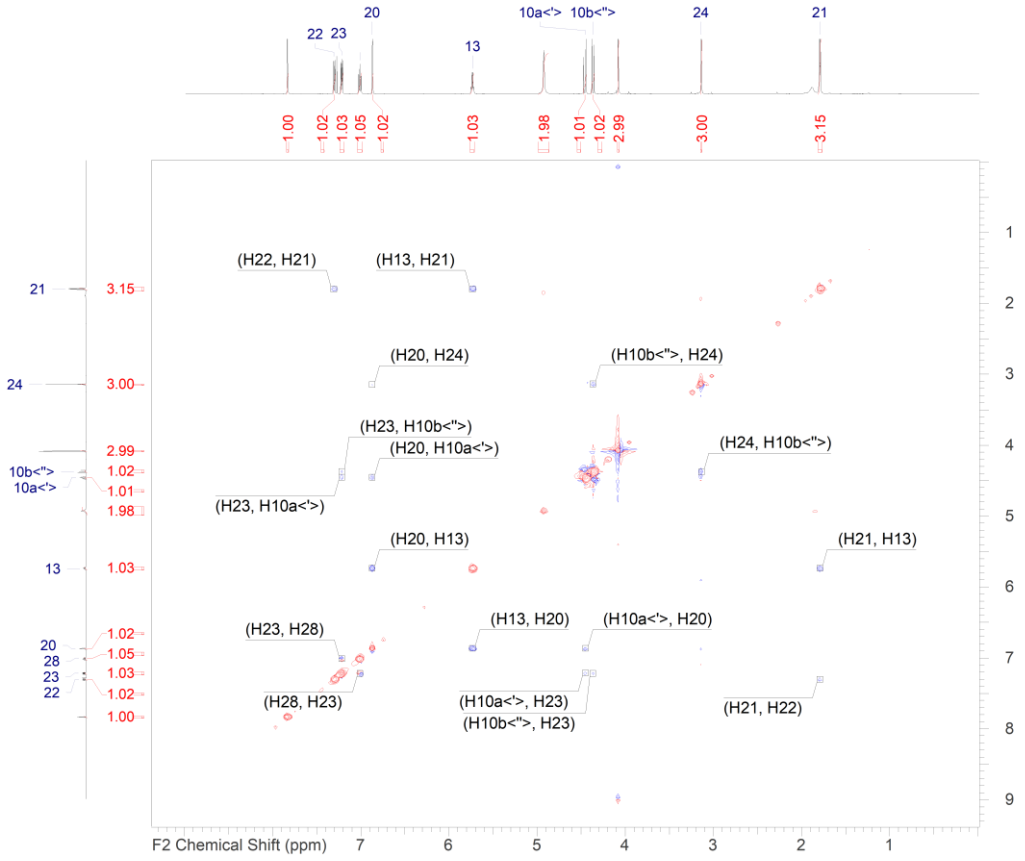
A2. NMR spectra of Lorlatinib in $CDCl_3$.



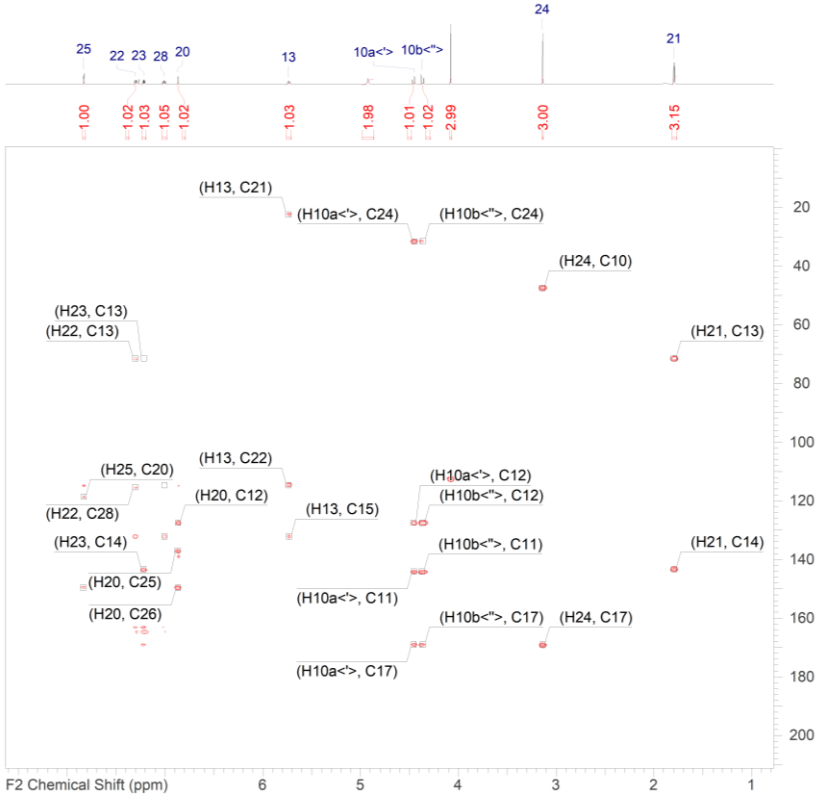
^{1}H-^{13}C HSQC-DEPT spectrum of Lorlatinib in $CDCl_3$.



^{1}H-^{1}H COSY spectrum of Lorlatinib in $CDCl_3$.

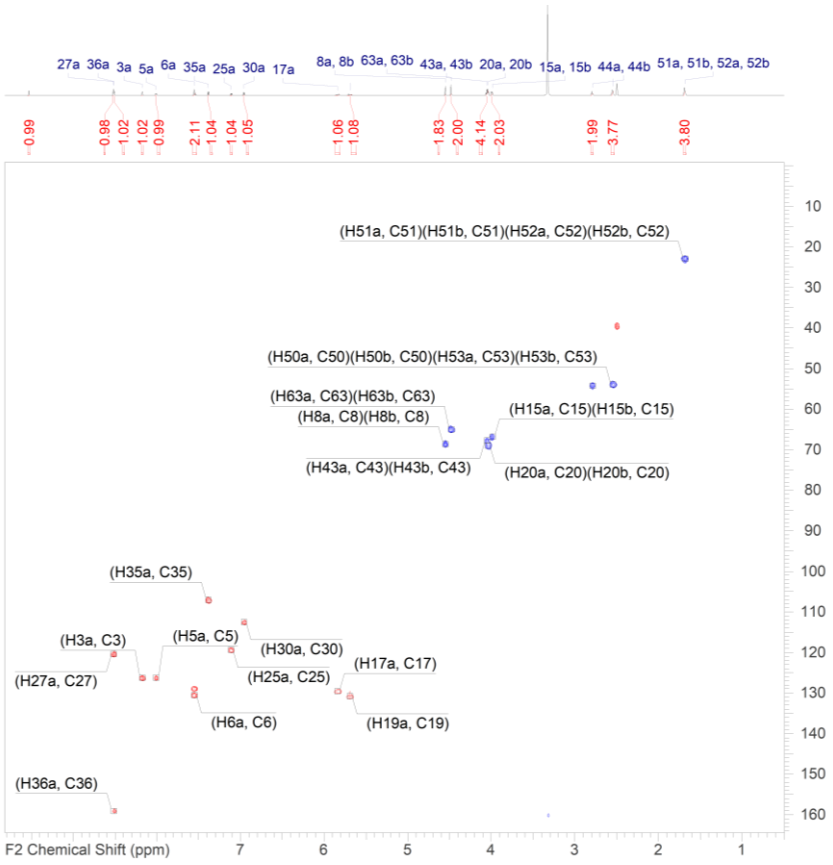


ROESY spectrum of Lorlatinib in $CDCl_3$.

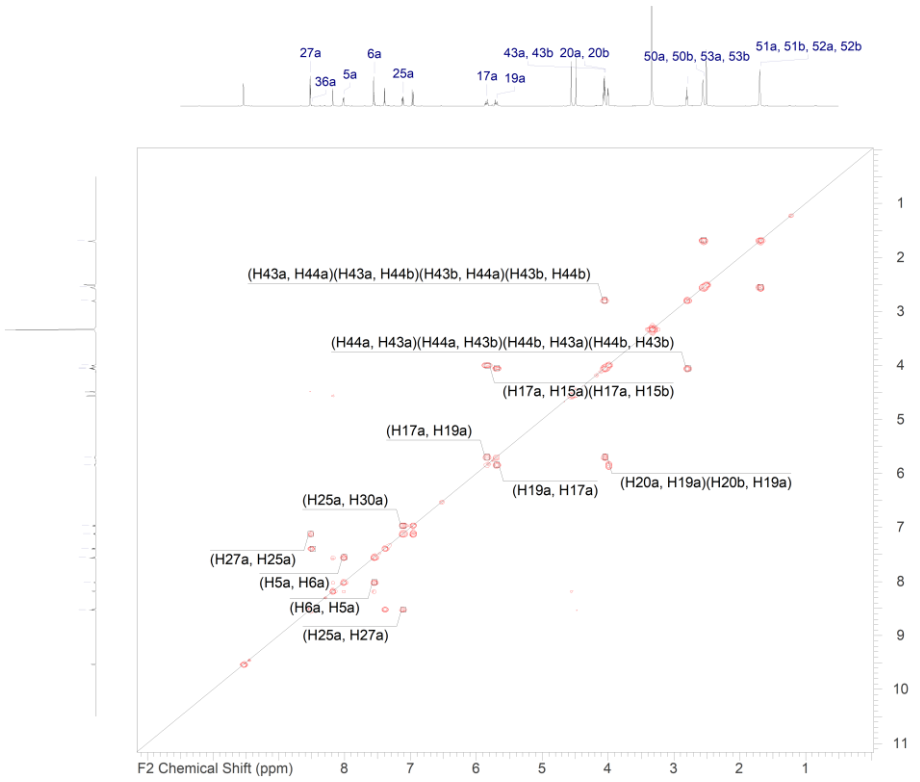


^{1}H-^{13}C HMBC spectrum of Lorlatinib in $CDCl_3$.

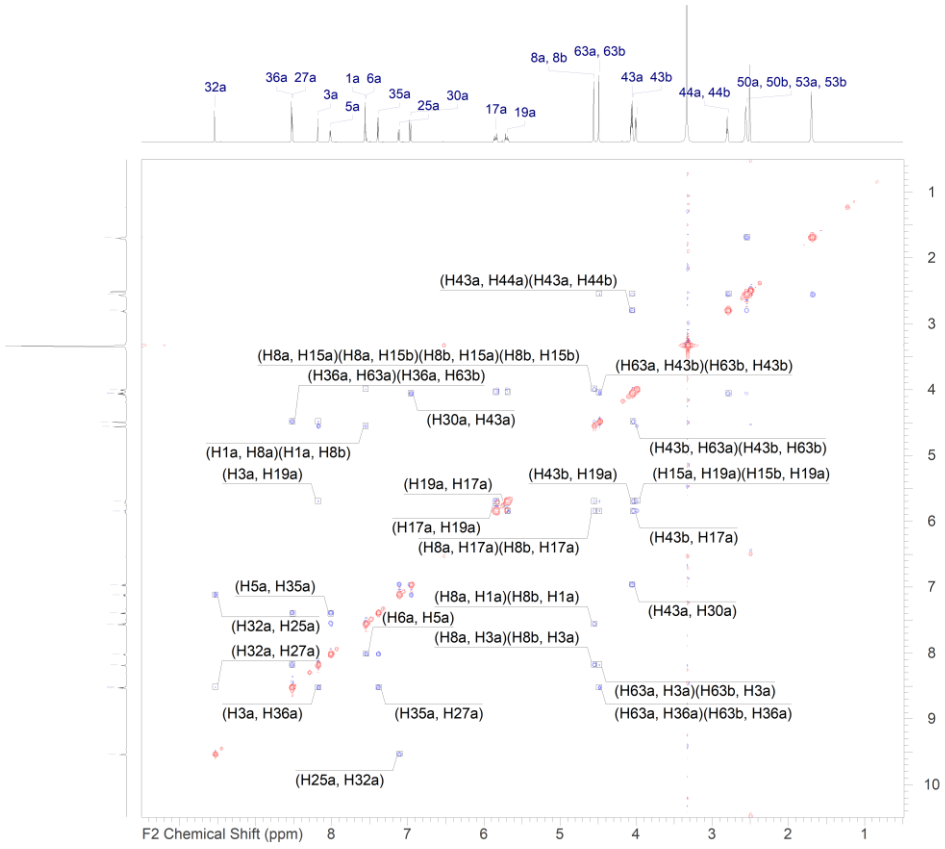
A3. NMR spectra of Pacritinib in DMSO-d_6.



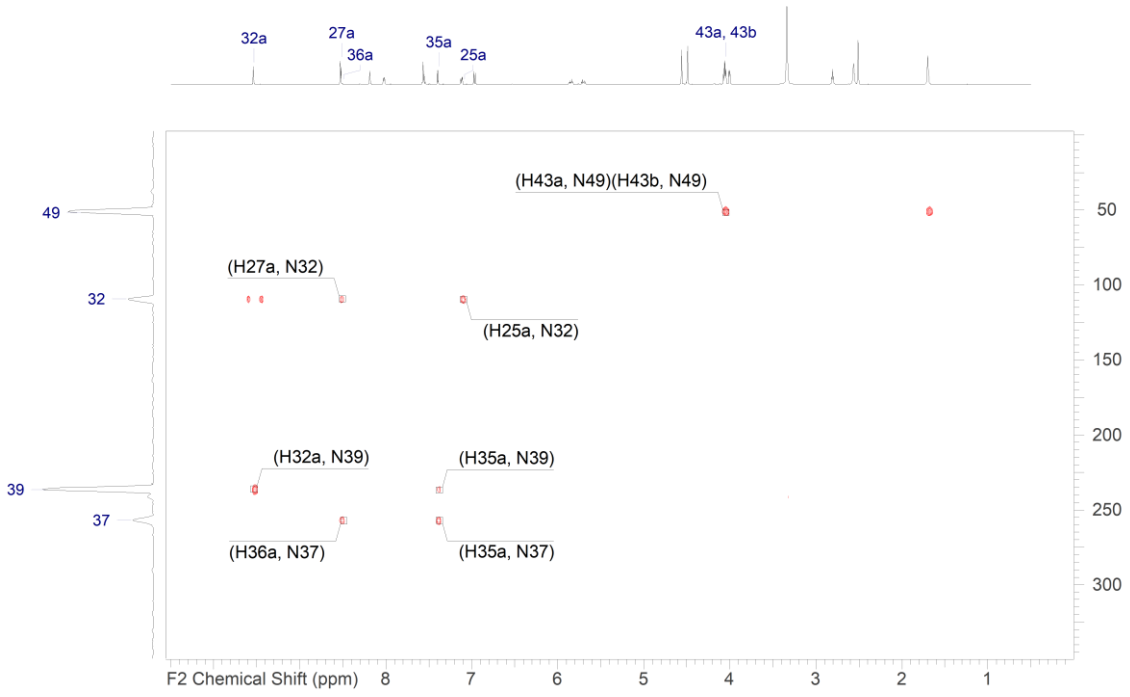
^{1}H-^{13}C HSQC-DEPT spectrum of Pacritinib DMSO-d_6.



^{1}H-^{1}H COSY spectrum of Pacritinib DMSO-d_6.



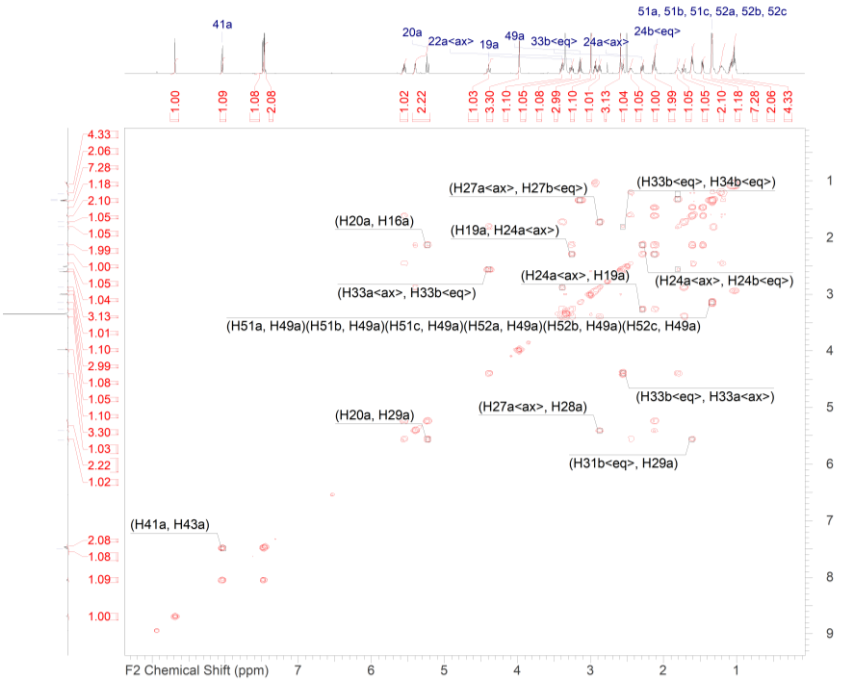
ROESY spectrum of Pacritinib DMSO-d_6.



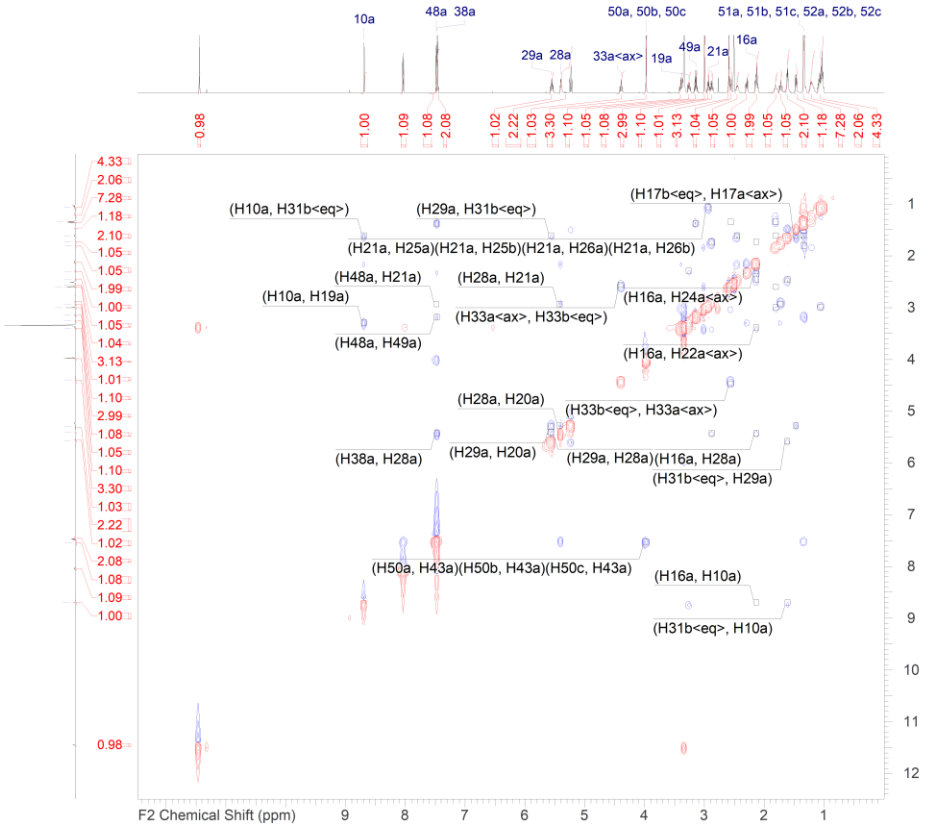
^{1}H-^{15}N HMBC spectrum of Pacritinib in DMSO-d_6.

A4. Spectra of Simeprevir in DMSO-d_6.

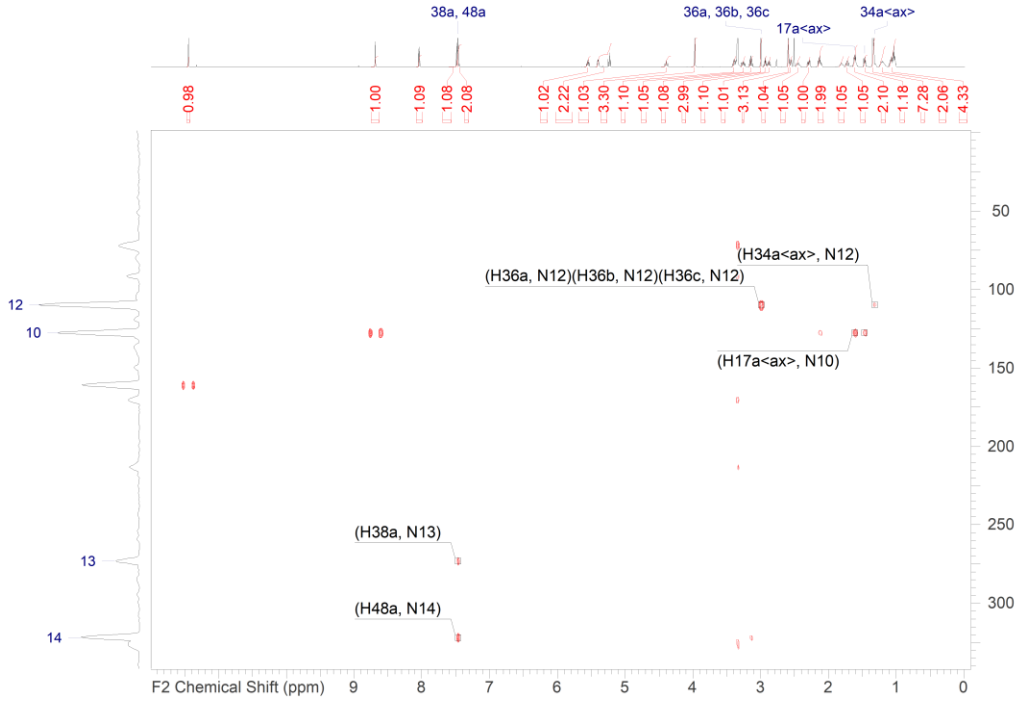
^{1}H-^{13}C HSQC spectrum of Simeprevir in DMSO-d_6.



^{1}H-^{1}H COSY spectrum of Simeprevir in DMSO-d_6.

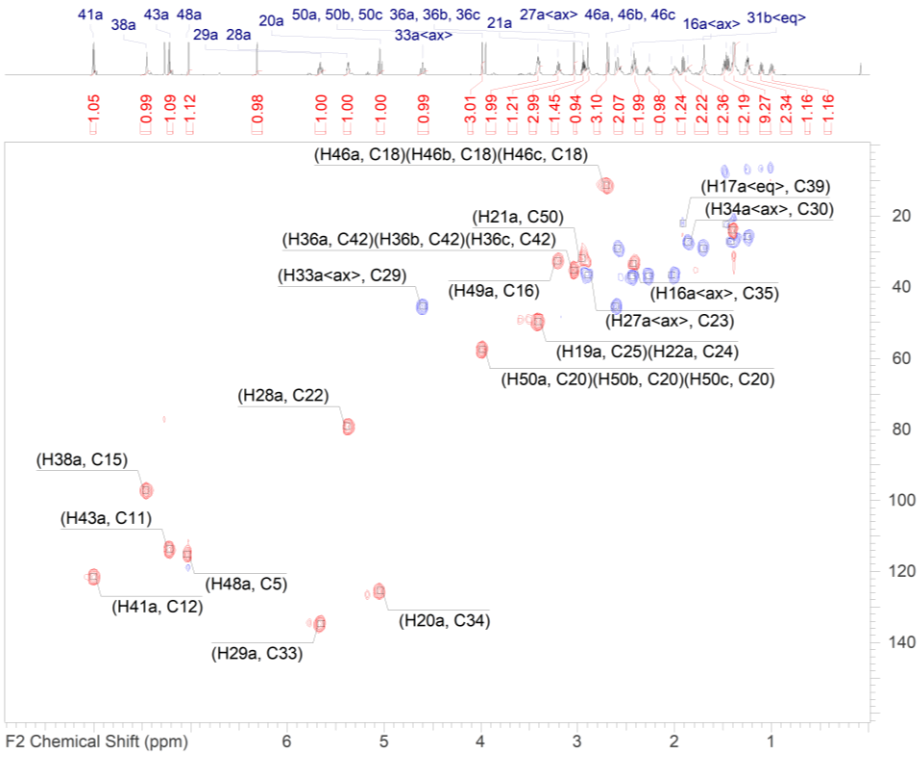


ROESY spectrum of Simeprevir in DMSO-d_6.

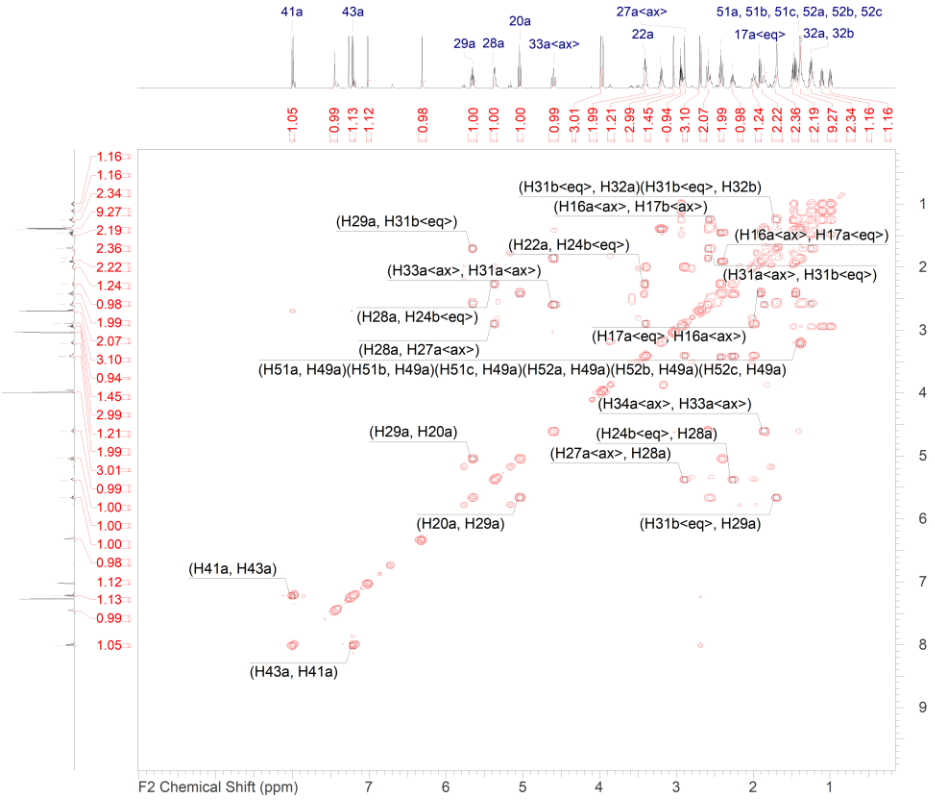


^{1}H-^{15}N HMBC spectrum of Simeprevir in DMSO-d_6.

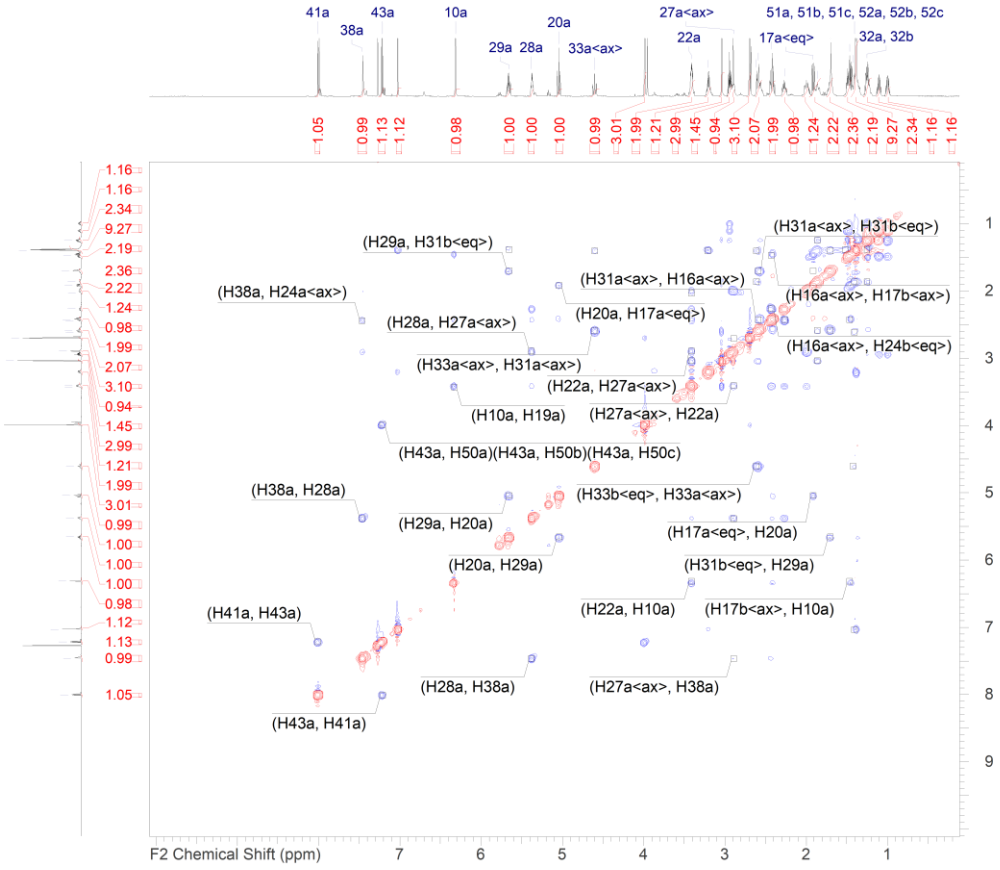
A5. NMR spectra of Simeprevir in $CDCl_3$.



^{1}H-^{13}C HSQC-DEPT spectrum of Simeprevir in $CDCl_3$.



^{1}H-^{1}H COSY spectrum of Simeprevir in $CDCl_3$.



ROESY spectrum of Simeprevir in $CDCl_3$.

Acknowledgements

I would like to acknowledge my tutors, Dott.ssa Valentina Mileo and Dr. Luca Venturi, and the head of the department, Dott.ssa Silvia Catinella, for welcoming me in Analytics and Early Formulations department's laboratories.

I would like to express my sincere gratitude to my supervisor, Dr. Luca Venturi for his guidance, valuable comments, motivation and for mentoring me throughout my PhD research project.

I am thankful to Prof. Marco Mor and his group for the opportunity of the PhD position and for introducing me to Dott.ssa Mileo's team. I would also like to acknowledge Prof. Mor, Prof. Alessio Lodola, Prof. Silvia Rivara and Dr. Laura Scalvini for teaching me lots of computational chemistry techniques.

I am truly grateful to Prof. Dr. Christina M. Thiele for giving me the opportunity to spend a research period in her group. Particularly, I would like to give my warmest thanks to Prof. Dr. Thiele, Dr. Volker Schmidts and Jan Rettig for sharing their extensive knowledge and expertise on RDCs and NMR pulse sequences. I am also thankful to Rimjhim Hossain, Lukas Laux, Michael Gölz, Patrick Maibach and Kilian Heckenberger for all the advice on LLC phases, for their support and enjoyable moments spent together.

I would like to thank all my colleagues, especially Dr. Elisa Moretti, Francesco Castagnini, Michael Vernizzi, Dr. Daniel Pecorari and Dr. Elena Picchi, for conveying their scientific knowledge during lab meetings and helping me in facing challenging moments.

Last, but not least, I am deeply grateful to all my friends and family for being extremely supportive and understanding. Their persistent encouragement throughout the course of my doctoral studies was of crucial importance in achieving my goals.