BLANK PAGE

UNIVERSITA' DEGLI STUDI

DI PARMA

Dottorato di Ricerca In

Progettazione e Sintesi di

Composti Biologicamente Attivi

Ciclo XXVI

Modelling studies of bitter taste, glucocorticoid and VEGFR2 receptors and their ligands

Relatore

Prof. GABRIELE COSTANTINO

Dottoranda

ANNA MARIA CAPELLI

CONTENTS

ABSTRACTS 6

LIST OF PAPERS 10

LIST OF UNIVERSITY SEMINARS 12

INTRODUCTION 14

COMPUTATIONAL METHODS 18

- Homology modelling 18
- Chemoinformatics analyses 19
- Docking and Scoring methods 21
- Steered Molecular Dynamics 23

MODELLING STUDIES OF BITTER TASTE RECEPTORS AND THEIR LIGANDS

- Introduction 30
- Methods 34
- Results and discussion 37
- Conclusions and future perspective 54
- Abbreviations 56
- References 56
- Appendix 1 61
- Appendix 2 61

ELUCIDATION OF GLUCOCORTICOSTEROIDS UNBINDING PATHWAYS FROM THE GLUCOCORTICOID RECEPTOR

- Introduction 64
- Methods 66

Results 70

Discussion 82

Conclusions 83

Abbreviations 86

References 86

MODELLING STUDIES OF VEGFR2 AND ITS LIGANDS

Introduction 92

Methods 99

Results 102

Discussion 116

Conclusions 117

Abbreviations 119

References 119

Appendix 1 121

Appendix 2 122

CONCLUSIONS 128

ACKNOWLEDGEMENTS 132

POSTER 1: MODELLING STUDIES OF BITTER TASTE RECEPTORS AND THEIR LIGANDS 133

POSTER 2: INSIGHT ON THE BINDING MODE OF GLUCOCORTICOID LIGANDS 134

ABSTRACTS

Modelling studies of bitter taste receptors and their ligands

The human bitter taste receptor gene family (TAS2R) belongs to the Frizzled/Taste2 subfamily of the G-protein coupled receptors (GPCR) superfamily. TAS2R are expressed on the tongue in bitter taste receptor cells co-expressing specific signal transduction components like Gα gustducin and are able to detect stimuli of only one taste quality. Recently these receptors have been identified in isolated human airway smooth muscle cells (ASM). Besides, in the literature it has been reported that bitter taste receptor agonists such as saccharin, chloroquine and denatonium can evoke increased intracellular calcium in ASM eliciting relaxation of isolated ASM and dilation of airways greater than that showed by β-adrenergic receptor agonists. Furthermore, inhaled bitter taste compounds can decrease airway obstruction in a mouse model of asthma. In light of these evidences this novel pathway looks attractive and might be exploited to identify novel compounds to treat asthma and chronic obstructive pulmonary disease (COPD). During the 1st year of my PhD course, chemoinformatics analyses of known bitter tastants aimed at finding common structural features and pharmacophores were performed. Besides, the generation of hTAS2R10, hTAS2R14 and hTAS2R31 receptor models was carried out together with docking studies of a few known bitter agonists in the putative binding sites of these receptor subtypes. Results of these experiments were validated in light of site-directed mutagenesis experiments available in the literature and provided useful suggestion about additional site-directed mutagenesis studies and chimera constructs to further validate these models. Results of these studies were presented as a poster at the Research Gordon Conference on "Computer Aided Drug Design", held at Mount Snow Resort, West Dover, VT, in 2011.

Elucidation of glucocorticosteroids unbinding pathways from the glucocorticoid receptor

Glucocorticoids are endogenous steroid hormones that regulate essential biological functions including metabolism, growth, and apoptosis. Glucocorticoids and structurally related drugs represent the most effective anti-

inflammatory agents to treat several inflammatory conditions. However, the clinical use of such drugs is hampered by severe side effects. Therefore, the development of novel glucocorticoid receptor (GRs) modulators with increased therapeutic index is impelling. Herein, using steered molecular dynamics (SMD) simulations a detailed picture of the unbinding process of three clinically relevant GR modulators from GR ligand binding domain is provided. The SMD protocol described can be used to prioritize the synthesis of structural analogues on the basis of their potential of mean forces (PMFs) and calculated unbinding energies. Moreover, these results are instrumental to explain at atomic resolution the reduced ability of dexamethasone to activate the naturally occurring mutant I747M-GR, which is implicated in rare familial glucocorticoid resistance, clinically characterized by glucocorticoids insensitivity. Results of these simulations were presented as a poster at the ACS Meeting held in Philadelphia in 2012 and published in the Journal of Medicinal Chemistry in September 2013.

Modelling studies of VEGFR2 and its inhibitors

Vascular Endothelial Growth Factor Receptor 2 (VEGFR2) is a type III receptor tyrosine kinase (RTK) in the PDGFR family. VEGF preferably binds this receptor and triggers important signaling pathway implicated in the process of vasculogenesis and angiogenesis. VEGFR2 is a clinically validated target to treat Renal Cell Carcinoma (RCC). Recently the crystal structures of sunitinib and sorafenib, two clinically relevant TKIs, bound to the kinase domain of VEGFR2 have been solved together with the enzyme juxtamembrane (JM) domain, highlighting the critical role played by this enzyme portion to form the inactivated state of VEGFR2 kinase which binds these ligands. The construct utilized to generate these VEGFR2 crystal structures was also used to measure the potencies, time dependencies and selectivity of a wide panel of TKI, including sunitinib and sorafenib. Besides, kinetic parameters (k_{off}) were detected for both these inhibitors and the new construct. While sunitinib is a prototypical type IV inhibitor, showing short dissociative half-life from VEGFR2, sorafenib is an exemplar of type II inhibitor, characterized by long dissociative half-life from VEGFR2 receptor. Herein, a preliminary detailed atomistic description of the unbinding process of sorafenib and sunitinib from VEGFR2 was unveiled by using SMD simulations. While sunitinib exits the ATP binding site from the cavity entrance, sorafenib moves towards the JM domain, causing a change in the orientation of the αC-helix. While no rupture

point could be identified in the force profile of sunitinib, analysis of the average force profile of sorafenib reveals that the rupture point involves the water-assisted breakage of sorafenib interaction with the hinge amino acid C919 and subsequently the disruption of the hydrogen bond of the ligand carbonyl urea moiety with D1046 (DFG-out motif). At the same time the hydrogen bond interactions established between the NH ligand urea moiety and E885 in the αC-helix likely delay ligand exit. Besides, the phenyl ring decorated with chlorine and trifluoromethyl groups makes hydrophobic interactions with some hydrophobic residues in αC-helix, which would further facilitate the ligand unbinding process. More robust conclusions will be drawn once the simulations of more ligand frames have been performed.

LIST OF PAPERS

This thesis is based on the papers listed below, which are referred to in the text by the corresponding Roman numbers (I-IV):

I **Elucidation of bitter taste receptor agonists binding modes**. Anna Maria Capelli, Gabriele Costantino. Research Gordon Conference on Computer Aided Drug Design. The Impact of Computational Sciences along the Drug Discovery Process. July 17-22, 2011 Mount Snow Resort, West Dover, VT. Chair: Gerhard Klebe, Vice Chair: Martin Stahl.

II **Insight on the binding mode of glucocorticoid ligands.** Anna Maria Capelli, Agostino Bruno, Gabriele Costantino. Abstracts of Papers, 244th ACS National Meeting & Exposition, Philadelphia, PA, United States, August 19-23, 2012 (2012), COMP-409

III **Unbinding pathways from glucocorticoid receptor shed light on reduced sensitivity of glucocorticoid ligands to a naturally occurring, clinically relevant, mutant receptor.** Anna Maria Capelli, Agostino Bruno, Antonio Entrena Guadix, Gabriele Costantino. Journal of Medicinal Chemistry 2013, 56, 7003-7014.

IV **Insight on the binding mode of glucocorticoid ligands.** Anna Maria Capelli, Gabriele Costantino. Lecture delivered at the Computationally-Driven Drug Discovery Meeting, IIT Genova, 4^{th} – 6^{th} February 2013

LIST OF UNIVERSITY SEMINARS

During my PhD course the following presentations were delivered at the "Dipartimento Farmaceutico" of the University of Parma:

I "Elucidation of bitter taste agonists binding mode and identification of novel chemotypes"

PhD course kick-off. March, 13th 2011

II "Elucidation of bitter taste agonists binding mode and identification of novel chemotypes"

Summary of the activities performed during the 1st year of PhD course. December, 5th 2011

III "Overview of the human sweet receptor and its ligands"

Mid-year 2012 seminar. June, 28th 2012

IV "Insight on the binding mode of glucocorticoid ligands"

Summary of the activities performed during the 2nd year of PhD course. January, 10th 2013

V "Overview of Steered-MD methodology and applications"

Mid-year 2013 seminar. July, 18th 2013

INTRODUCTION

Nowadays computational chemistry techniques are extensively used to accelerate drug discovery. Several methods have been developed and the use of the most appropriate ones strictly depends on the information available both on the target under investigation and its ligands. In the early phase of drug discovery generally the information available on the target and its ligands is low. In the past high-throughput random screening (HTS) of large repositories had been extensively utilized in big pharmaceutical companies to find new chemical starting points (hits). However, this massive screening approach has failed to keep up with its initial promises due to overall poor hit rate associated with large costs. In light of that, nowadays it is preferable to utilize computational methods to build and then screen rationally-designed focused sets exploiting all the information available on the target and its ligands. In order to accomplish this task, data-mining of large databases encoding chemical structures as well as their biological profiles is generally performed in order to fish out ligands of the target of interest or for closely-related members of the target family. Then, the derivatives identified are analyzed with the use of chemoinformatics methods in order to find common fragments, shapes, pharmacophores and biological profiles (patterns). Subsequently, these patterns can be used to drive the selection of commercial database compounds exhibiting similar patterns and described in the same way. According to the similarity principle, compounds showing similar patterns should also have similar biological activity. Further to common patterns, focused sets are built by taking into account physico-chemical properties cut-offs, rules (e.g. Lipinski's rule of 5) or mathematical equations (e.g. Hansch's blood brain barrier relationship) that models compound ADMET properties and prospectively in vivo PK profiles suitable for either oral or topic administration. These rules are generally derived by chemoinformatics analyses of compound datasets and their biological response (e.g. blood-brain barrier permeation). When no ligands are known for the target of interest (e.g. a new protein kinase), focused screening is performed by using the datasets built for the same target family (protein kinases) using for example pharmacophoric information of known ATP competitive inhibitors or ATP binding site.

In the early phase of drug discovery generally the crystal structure of the target under investigation is not known. Even though significant experimental improvements in the crystallization approaches have been done, in particular in

the field of membrane receptors (CRF1, GLP), there are still many targets for which the structure determination is elusive. When the experimental 3D structure of the target is not known, homology models might be generated exploiting for example the crystal structures of similar targets as templates, conserved amino acids involved in target activation or reaction mechanism, hydrophobicity moments of helices, site-directed mutagenesis data and chimera constructs. Then, potential binding sites are identified and ligands accommodated (docked) in these cavities. The binding modes obtained are validated by using site-directed mutagenesis or synthesis of close structural analogues for screening (structure-activity relationship generation).

The 1st chapter of this thesis describes the use of chemoinformatics methods to identify common patterns shared by a set of known bitter taste ligands. Besides, this chapter reports the generation and validation of the homology models of three bitter taste receptor subtypes relatively more abundant in the air smooth muscle cells. Furthermore, these models were used to dock a few selective bitter tastants in their putative binding sites in order to identify their potential binding modes and subsequently design experiments to validate them.

In the lead optimization phase of drug discovery several information are generally available both on the target and its ligands. In this section of drug discovery, medicinal chemistry efforts are mainly focused on achieving the best balance between the pharmacodynamics (PD), the pharmacokinetics (PK) and the selectivity/toxicity profiles of the lead compounds in order to minimize undesirable side effects. In this context the optimization of ligand efficacy at the target of interest is the most relevant as it might result in longer duration of action in vivo with a low dose. Binding of a compound to a target protein is a dynamic process involving molecular recognition and complex stability measured as the residence time of the drug molecule on its target. For in vivo situations the duration of ligand efficacy depends on the rate of receptor-ligand association (k_{on}) and on the dissociation rate constant (k_{off}). The former is mainly relevant for drug dose definition; the latter reflects mean complex lifetime and impacts compound PD and PK profiles. Then, the k_{off} parameter can be translated into the dissociative half-life ($t_{1/2}$ dissociative) for the receptor-ligand complex, which is a direct measure of the residence time[1]. Hence, the availability of a robust computational method able to successfully rank compounds according to their *in silico* estimated kinetic dissociation constants for a specific target would be highly desirable to prioritize compound synthesis and accelerate the lead optimization process. In this respect, the use of

molecular dynamics simulations on long time scale fulfills these purposes. Besides, analyses of the ligand unbinding trajectories obtained with these methods can provide atomistic details that could be exploited in the design of novel derivatives.

In the 2nd chapter of this PhD thesis are reported the steered molecular dynamics (SMD) simulations performed with Amber software and CINECA Consortium infrastructure to gather detailed information about the unbinding pathways of prototypical glucorticosteroids from the glucocorticoid receptor and to rationalize their duration of action at this target. Then, the protocol successfully developed and validated can be used to properly rank newly designed structural analogues of these agonists on the basis of their PMFs calculated with the aforementioned approach.

Finally, in the 3rd chapter of this PhD thesis SMD technique has been utilized to identify potential unbinding pathways of two clinically relevant kinase inhibitors (sunitinib, sorafenib) from the prototypical receptor tyrosine kinase Vascular Endothelium Growth Factor Receptor 2 (VEGFR2).

References

(1) Tummino, P. J., Copeland, R. A. Residence Time of Receptor−Ligand Complexes and Its Effect on Biological Function. *Biochemistry* **2008**, *47*, 5481–5492.

COMPUTATIONAL METHODS

Several computational approaches have been used to accomplish the studies described in this thesis. In the paragraphs below is reported an overview of the main techniques utilized together with highlights of their strengths and weaknesses.

(i) Homology modeling

Structure-Based Drug Design is a computational approach which exploits the 3D coordinates of the target and its ligands to gather information about their binding modes and to drive the selection and/or the design of novel derivatives. Either X-ray crystallography or NMR technique can be used to solve the structure of a protein with bound ligands providing detailed information of the binding site location as well as that of its electrostatic properties and shape. When a ligand is not known, both these techniques can be used to solve the structure of the target in its "apo" conformation, which can be exploited to guide the design of experiments aimed at finding potential binding sites (e.g. site-directed mutagenesis, chimera constructs, and computational simulations). Despite the significant technological development of both X-ray crystallography and NMR technique, there are still target families for which target-ligand structural information is not known. This is often the case when an exploratory research project initiates on a novel target. In this situation, homology modeling can be exploited to generate its 3D structure at atomic resolution by using the experimentally known structure of a template. This computational approach is based on two observations: 1) the structure of a protein is uniquely defined by its sequence, hence knowledge of the sequence should be enough to generate the structure; 2) during the evolution, the structure changes slower than its sequence; therefore, similar sequences give rise to similar structures[1,2]. In light of that, two sequences are likely to adopt the same structure provided that the length of the two sequences and the percentage of their amino acid identity fall within a "safe" region[3].

Homology modeling is a multistep process which requires: 1) the identification of the template and generation of a preliminary sequence alignment; 2) alignment correction; 3) generation of the protein backbone; 4) modeling of the loops; 5) modeling of the side chains; 6) optimization of the model; 7)

validation of the model. Beside sequence similarity, knowledge of the amino acids involved in target activation or reaction mechanism, hydrophobicity moments of helices for GCPRs, site-directed mutagenesis data and chimera constructs can be exploited to refine the homology models. Due to the complexity of the whole process which requires accurate evaluation and selection of the most appropriate solution among similar choices, the availability of computational tools for automatic model generation is highly desirable. Despite the efforts to elaborate reliable and robust software for high-throughput model generation (e.g. Modeler, PRIME), worldwide evaluations of protein structure prediction tools (CASP) have highlighted that human expertise is still invaluable, in particular when the sequence alignment is close to the twilight zone[4]. Finally, when no template with a known structure is available, fold recognition or *ad initio* folding techniques can still be tried. In this PhD thesis, homology modeling was utilized to generate the 3D structures of bitter taste receptors for subsequent docking experiments.

(ii) Chemoinformatics analyses

"Chemoinformatics is the mixing of those information resources to transform data into information and information into knowledge for the intended purpose of making better decisions faster in the area of drug lead identification and optimization"[5]. The core of this discipline which combines chemistry, computer and information science, consist in the chemical structure and its representation, which comprises the definition and calculation of molecular descriptors (e.g. structural, pharmacophoric and affinity fingerprints). Then, chemical structures can be compared in pairs with the use of a metric (e.g. Tanimoto index) on the basis of either 1D (e.g. MW), 2D (e.g. structural fingerprints) or 3D (e.g. pharmacophoric) descriptors. Tanimoto coefficient[6,7] compares the matching of ON bits between two fingerprints to all the ON bits that could have been matched between them. Tanimoto coefficient which extends the Jaccard coefficient[8] (Jaccard), measures overall similarity between two molecules and is symmetric. On the contrary, the Twersky index[9] evaluates whether one fingerprint is a subset of the other i.e. if one molecule is a perfect substructure of a larger one it would achieve a perfect score. Hence, Twersky metric is asymmetric. Generally, Twersky indices are calculated in both directions and the best score is taken. With proper structural descriptors as well as similarity or distance measurements, compounds can be classified by using either clustering or partitioning. These analyses are generally performed on large number of

compounds, which require tools for storage, indexing, searching and retrieval of molecular structures.

Data-mining is a key approach of chemoinformatics. It consists in the identification of clear relationships (patterns) between chemical structures and activities or properties which are used to subdivide the compounds in groups. As a result of this analysis, the compounds in the same group share common patterns. In this process it is critical to select appropriate descriptors relevant to the biological phenomena under study. Common chemoinformatics patterns are represented by Markush structure, fingerprint, 3D-pharmacophore model, regression, decision tree classification, hierarchical and non-hierarchical clustering as well as self-organizing map.

Chemoinformatics has several applications in drug discovery. First of all, it can be used to select compounds from databases (in house collection or commercially available repositories) according to their similarity or diversity with respect to reference compounds, described in the same way. Besides, it can be used to generate virtual libraries based on their structural/pharmacophoric/shape diversity or similarity evaluated with respect to a reference set of compounds or among them. Similarly, chemoinformatics can be used to carry out ligand-based virtual screening of compounds collection following different criteria. This process allows for discarding unwanted derivatives prior to experimental determination of their biological profile. Virtual screening is also a great tool for the design of a combinatorial library with a given target.

Chemoinformatics is also utilized to carry out structure-activity relationship (SAR) analysis of high-throughput screening (HTS) compounds to prioritize hits before lead identification and to generate ADMET mathematical models exploiting heterogeneous sets of compounds. These "general" models are used to discard compounds prior to screening.

A hot chemoinformatics topic is represented by polypharmacology predictions. Some chemical similarity approaches use set-wise ligand chemical similarity as a proxy for the pharmacological similarities of their protein targets. Hence, the protein space is reorganized by using both known and statistically–predicted inferred relationships, which exploits the internal similarity of most ligands for a particular target and the observation that similar ligands often have similar protein binding patterns. Instead of evaluating overall similarity between ligand sets, the average nearest neighbor detected similarity arising from small ligand

sub-groups is measured[10]. Similarly, Keiser *et al.*[11] introduced a similarity ensemble approach (SEA) to link receptors by exploiting the statistical significance of similarity among high-scoring ligand pairs across receptor sets. Following the same principles, virtual profiling protocols may assist lead safety assessment and the prioritization of lead series, by favoring those predicted to have fewer off-target interactions. Further examples of chemoinformatics application to drug design are reported in the literature[12-14].

In this PhD thesis, chemoinformatics methods have been used to identify common fragments and pharmacophores of bitter tastants and then cluster them accordingly. Patterns identification was also tried with the use of Principal Component Analysis coupled to classical 1D and 2D descriptors. Besides, bitter tastants similarity was evaluated by using affinity fingerprints encoding their pharmacological profile for bitter taste receptors. This methodology described by Kumar *et al.*[15] provides a complementary description of the compounds independently from their chemical structures and can be useful to find biological patterns among chemically disparate set of ligands.

(iii) Docking and scoring methods

Molecular docking is a computational technique in which the 3D structure of a protein is held rigid while conformations and positions of the ligand are generated or sampled within the protein binding site ("lock and key" principle). Scoring is a computational method that estimates how well the docked ligand interacts with or complements the binding site. Hence, the scoring function is used both for the identification of the "correct" binding mode and for the prediction of ligand affinity. Docking and scoring methodology involves the following steps: 1) characterization of the target binding site (e.g. electrostatic and shape profiling); 2) positioning and conformational sampling of the ligands and (partially) of the protein binding site amino acids (docking); 3) quantitative assessment of the quality of the protein-ligand interactions (scoring). All these tasks are automatically performed by computational tools which have been mainly developed to perform high-throughput docking and to rapidly score millions of compounds (structure-based virtual screening). Ligand conformational sampling can be performed either outside or within the protein binding site. In the former, the ligand conformations are generally positioned in the protein cavity using a grid-based method; in the latter, the ligand core is initially placed in the pocket and then progressively grown following a build-up approach.

Considerable efforts have been dedicated by software developers to the generation of accurate scoring functions which fall under three categories: knowledge-based, empirical and (physical) force fields[16]. Knowledge-based scoring functions have been derived from statistical analyses of inter-atomic pairwise potentials identified in protein-ligand crystal structures while ligand affinity is not taken into account. Empirical scoring functions decompose or simplify protein-ligand interactions into a few scalable functions. They typically count the number of hydrophobic interactions, hydrogen bonds and rotatable bonds as a rough estimation of entropy loss. The weight of each component is identified by performing a linear regression so that to find the best fit between the calculated scores and the experimentally derived binding affinities of a ligand-protein training set. Finally, physical scoring functions rely on force fields originally developed to simulate the conformational, thermodynamic and kinetic behavior of small molecules or macromolecules. The interaction components of the force fields should provide more accurate predictions of protein-ligand binding affinities.

Large benchmarking sets covering diverse proteins and their ligands such as the "Directory of Useful Decoys" (DUD)[17] have been disclosed in order to evaluate docking screening calculations and compare different methods. Other sets have been generated by exploiting company targets and ligands with their associated biological data (GSK dataset)[18]. The availability of public datasets and heterogeneous software has stimulated the publication of several comparative studies, which in turn have stimulated debates about the use of appropriate quantitative assessment metrics. An example is represented by the evaluation performed by Warren *et al.*[18]. According to this exercise, docking programs are able to reproduce the crystal conformations of the ligands (correct binding modes) for 31-94% of the ligands for a particular protein target by using root mean square deviation (RMSD) as metric. Besides, the docking algorithms could not generate docked ligand conformations (poses) within 2.0 Å RMSD of the correct pose when applied to large binding sites, flexible ligands o systems with few non-hydrophobic interactions between the ligand and the protein. On the contrary, the scoring functions were less successful at identifying the docking solution closest to the crystal conformation. Based on the results of this exercise and others it can be reasonably concluded that docking engines can generate good docking poses even though the scoring functions limit their predictive ability.

When docking was used as a virtual screening tool, docking programs could identify molecules active against a target out of a population of decoy molecules. However, the evaluation performed by Warren *et al.*[18] has shown that the performance across targets was inconsistent. In addition, the application of knowledge about a protein target improved enrichment and consistency, but did not aid in the rapid identification of diverse chemotypes. Hence, structure-based virtual screening is successful by eliminating inactive compounds rather than by identification of active derivatives.

The most remarkable aspect of Warren *et al*[18].'s exercise is the observation about the lack of a strong correlation noticed for any scoring function protein target pair. In most cases, the reproduction of the binding mode did not improve rank-order or potency prediction performance even for congeneric series of compounds. Finally, no single docking program performed well across all proteins, which suggests that a few validation experiments should be performed prior to initiate a docking and scoring exercise on a specific target. This study demonstrated that significant improvements of the compound scoring are needed in order to routinely use it in lead optimization. Unfortunately, despite all the efforts spent to accomplish this task, scoring functions are still not accurate enough to be exploited for quantitative ligand potency prediction. Hence, the interest in Free Energy Perturbation (FEP)[19] methods has increased in recent years and these approaches are now perceived as a "follow-up" of docking and scoring for drug design purposes.

In this PhD study, docking and scoring were used to find binding modes of bitter tastants in their putative binding sites of the bitter taste receptors. Predicted binding modes were validated in light of the site-directed mutagenesis data reported in the literature.

(iv) Steered Molecular Dynamics

The target-ligand binding energy can be estimated with either docking methods or molecular dynamics (MD) simulations. However, docking scores have been mainly developed to rank huge number of structurally heterogeneous compounds (*in silico* screening). Hence, they do not provide an accurate estimation of ligand binding affinity[18]. At the same time the availability of cost-effective accelerator processors such as IBM Cell or Nvidia GPUs offers the opportunity to simulate dynamic rather than static models and to get quantitative description of the binding sites (e.g. milliseconds MD time scale simulation of β2AR-Nb80 with the agonist BI-167107 bound in the binding

site)[20]. There are several methods based on MD simulations such as linear interaction energy (LIE), molecular mechanics/Poisson-Boltzmann (e.g. MM/PB(GB)SA), free energy perturbation (FEP) or thermodynamic integration (TI). Despite the availability of high speed processors and sophisticated MD techniques, conformational sampling of a protein is still very complex. In order to simplify that, enhanced sampling methods have been elaborated. Among these techniques, Steered-MD (SMD) has been developed and successfully applied to study mechanical unfolding of biomolecules, transportation of ions and organic compounds through membrane channels and to probe unbinding pathways of a ligand from its receptor. Recently an example of SMD application to drug design purposes in combination with docking studies was reported in the literature[21], enlarging the applicability of this technique to a new interesting field. Steered-MD is based on the Jarzynski's equality[22] which states that for a canonical system equilibrium information of the Helmoltz free energy ΔF can be extracted form an ensemble of non-equilibrium measurements ($\beta=1/k_BT$):

$$e^{-\beta\Delta F} = < e^{-\beta W} >$$

This equality is independent from the rate and the path γ to go from state A to state B and it works for extremely slow changes of parameters (i.e. quasistatic systems). The Jarzynsky's equality is a real mathematical breakthrough as it opens the possibility of calculating free energies from non-equilibrium processes (non-equilibrium thermodynamic integration) for large biomolecular systems. The Potential of Mean Force (PMF) $\Phi(\xi)$ represents a free energy profile as a function of an internal coordinate $\xi(\mathbf{r})$. In order to apply Jarzynski's equality to the calculation of the PMF $\Phi(\xi)$, an external parameter λ correlated to ξ is introduced with the use of a guiding potential, i.e. a spring that constraints ξ to be near λ.

$$h(\mathbf{r};\lambda) = \frac{k}{2}[\xi(\mathbf{r}) - \lambda]^2$$

The region of ξ for which the PMF $\Phi(\xi)$ has to be calculated is covered by changing λ with a constant velocity $\lambda_t = \lambda_0 + vt$. A sufficiently large k force constant for the guiding potential (stiff spring) allows the reaction coordinate ξ to closely follow the constraint center λ. When k is large, most of the

contribution to the PMF comes from the region around λ, leading to the stiff spring approximation[23].

$$F(\lambda) \cong \Phi(\lambda)$$

One of the most critical aspects of this methodology is represented by the definition of the reaction coordinate. At this purpose, knowledge-based information could be exploited to define a privileged ligand-receptor way-out. As an alternative, Random Acceleration Molecular Dynamics (RAMD)[24] can be used to carry out molecular dynamics simulations with an additional randomly oriented acceleration applied to the center of mass of ligand atoms in the system. This approach can be used to identify egress routes for a ligand from a buried protein binding site following unbiased trajectories. With the former approach the unbinding pathway is biased by the initial selection; with the latter method many alternative solutions can be found and criteria need to be identified to distill out the most realistic ones. In the literature several examples of the use of both these approaches have been described. Unfortunately, no general rule can be devised and the reaction coordinate selection depends on the information available on the target and its ligands. Further to the selection of the reaction coordinate, the use of SMD requires the calibration of both the constant pulling velocity and the force constant as these parameters are strictly dependent on the system under study (i.e. no velocity-force constant combination fits all the complexes). Generally this optimization is empirically performed following a step-wise iteration i.e. when one parameter is changed the other one is kept fixed and vice-versa. This approach is well-described in the seminal work of Park *et al.*[23]. When SMD is used to study unbinding pathways, preliminary SMD experiments are performed on a few representative ligands, usually one slow and one fast dissociating compound. The optimal combination of the pulling velocity and the force constant is selected in order to ensure small deviation of the reaction coordinate from the constraint position for both ligand simulations. Besides, conformational changes of the protein have to be constantly monitored along all the simulation timing by calculating RMSD of protein reference atoms and by visual inspecting the frames sampled. Further to system calibration, SMD experiments need to be replicated on several target-ligand equilibrated complexes in order to get statistically significant results. Finally, accurate analyses of the force profiles as well as SMD trajectories allow for the identification of breakage or formation of relevant protein-ligand atoms interactions along the unbinding pathways, observation of conformational

changes of protein amino acids and understanding of the role played of water molecules in assisting the whole process.

In this PhD thesis SMD simulations were used to unravel at atomistic level the unbinding pathways of three clinically relevant glucocorticoid receptor (GR) modulators from GR ligand binding domain and to derive a qualitative relationship between their PMFs and kinetic (k_{off}) experimental parameters. Besides, this methodology was used to model the unbinding pathways of two clinically relevant inhibitors of vascular endothelium growth factor receptor 2 (VEGFR2) from their receptor.

References

(1) Chothia C, Lesk A.M. The relation between the divergence of sequence and structure in proteins. *EMBO J.* **1986**, 5, 823–36.

(2) Sander C., Schneider R. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **1991**, *9*, 56–68.

(3) Krieger E., Nabuurs S.B., and Vriend G. Structural Bioinformatics, Edited by Philip E. Bourne and Helge Weissig, ISBN 0-471-20199-5 Copyright © 2003 by Wiley-Liss, Inc.

(4) Fischer D., Barret C., Bryson K., Elofsson A., Godzik A., Jones D., Karplus K.J., Kelley L.A., Mac-Callum R.M., Pawowski K., Rost B., Rychlewski L., Sternberg M.J.E. CAFASP1: Critical assessment of fully automated structure prediction methods. *Proteins (Suppl. 3)* **1999**, 209–17.

(5) a) Brown F.K. "Chapter 35. Chemoinformatics: What is it and How does it Impact Drug Discovery". Annual Reports in Med. Chem. Annual Reports in Medicinal Chemistry 1998, 33, 375. doi:10.1016/S0065-7743(08)61100-8. ISBN 978-0-12-040533-6. b) Brown F. "Editorial Opinion: Chemoinformatics – a ten year update". *Current Opinion in Drug Discovery & Development* **2005**, *8*, 296–302.

(6) Willett P. Similarity and Clustering in Chemical Information Systems. Research Studies Press, Wiley, Letchworth, Hertfordshire, England; New York, 1987.

(7) Brown R. D. and Martin Y. C. Use of Structure−Activity Data To Compare Structure-Based Clustering Methods and Descriptors for Use in Compound Selection. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 572–584.

(8) Jaccard P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin de la Socie´te´ Vaudoise des Sciences Naturelles* **1901**, *37*, 547–579.

(9) Tversky A. Features of Similarity. *Psychol. Rev.*, **1977**, *84*, 327–352.

(10) Izrailev S., Farnum M. A. Enzyme classification by ligand binding. *Proteins* **2004**, *57*, 711–724.

(11) a) Keiser M.J., Roth B.L., Armbruster B.N., Ernsberger P., Irwin J.J., Shoichet B.K. Relating protein pharmacology by ligand chemistry. *Nat. Biotech.* **2007**, *25*, 197–206. b) Keiser M.J., Setola V., Irwin J.J., Laggner C., Abbas A., Hufeisen S.J., Jensen N.H., Kuijer M.B., Matos R.C., Tran T.B., Whaley R., Glennon R.A., Hert J., Thomas K.L.H. , Edwards D. D., Shoichet B.K., Roth B.L. Predicting new molecular targets for known drugs. *Nature* **2009**, *462*, 175-181

(12) Leach A. R., Gillet V. J. An Introduction to Chemoinformatics, Kluwer Academic Publishers Dordrecht, Neth., 2003.

(13) Vogt M., Bajorath J. Chemoinformatics: A view of the field and current trends in method development. *Bioorganic and Medicinal Chemistry* **2012**, *20*, 5317-5323.

(14) Gozalbes R., Pineda-Lucena A. Small molecule databases and chemical descriptors useful in chemoinformatics: an overview. *Combinatorial Chemistry & High Throughput Screening* **2011**, *14*, 548-558.

(15) Kauvar L.M., Higgins D.L., Villar H.O., Sportsman J.R., Engqvist-Goldstein A., Bukar R., Bauer K.E., Dilley H., Rocke D.M. Predicting ligand binding to proteins by affinity fingerprinting. *Chem Biol.* **1995,** *2*, 107-18.

(16) Moietessier N., Englebienne P., Lee D., Lawandi J., Corbeil C.R. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Br.J.Pharmacol.* **2008**, *153*, S7-S26.

(17) Huang N., Shoichet B.K. and Irwin J.J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006,** 49, 6789-6801.

(18) Warren G. L., Andrews C. W., Capelli, A.M., Clarke B., LaLonde J.; Lambert M. H., Lindvall M., Nevins N., Semus S.F., Senger S., Tedesco G., Wall I. D., Woolven J. M., Peishoff C. E., Head M. S. A Critical Assessment of Docking Programs and Scoring Functions. *J.Med.Chem.* **2006**, *49*, 5912-5931.

(19) Zwanzig R.W. High Temperature Equation of State by a Perturbation Method. I. Non Polar Gases. *J.Chem.Phys.* **1954**, *22*, 1420-26.

(20) Rosenbaum D.M., Zhang C., Lyons J., Holl R., Aragao D., Arlow D.H., Rasmussen S.G.F, Choi H.-J.,1, DeVree B.T., Sunahara R.K., Chae P.S., Gellman S.H., Dror R.O., Shaw D.E., Weis W.L.,1, Caffrey M., Gmeiner P., Kobilka B.K. Structure and Function of an Irreversible Agonist-β2 Adrenoceptor complex. *Nature* **2011**, *469*, 236–240.

(21) Colizzi F., Perozzo P., Scapozza L., Recanatini M., Cavalli A. Single-Molecule Pulling Simulations Can Discern Active from Inactive Enzyme Inhibitors. *J.Am.Chem.Soc.* **2010**, *132*, 7361-71.

(22) Jarzynski C. Nonequilibrium Equality for Free Energy Differences. *Physical Reviews Letters* **1997**, *78* , 2690-93.

(23) Park S., Khalili-Araghi F., Tajkhrshid E., Schulten K. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality *J. Chem.Physics* **2003**, *119*, 3559-66.

(24) Luedemann S. K., Carugo, O., Wade, R. C. Substrate access to cytochromeP450cam: A comparison of a thermal motion pathway analysis with molecular dynamics simulation data. *J. Mol. Model.* **1997**, *3*, 369.

MODELLING STUDIES OF BITTER TASTE RECEPTORS AND THEIR LIGANDS

Introduction

In mammalians the sense of taste is mediated by receptors located in the oral cavity. Each of the five taste qualities (sour, salty, sweet, umami and bitter) recognizes specific food components in the diet[1]. Extragustatory expression of taste signaling molecules and taste receptors is associated with either solitary chemosensory cells (SCC) or brush cells. Cells that express signaling components related to the taste system have been found in nasal respiratory epithelium, vomeronasal organ, airways, gastrointestinal tract and testis. However, direct evidences for bitter taste receptor gene expression in these organs are scarce.

A family of genes has been identified and demonstrated to encode bitter taste receptors TAS2Rs (or T2Rs)[2-4]. Vertebrate species differ in the number of TAS2Rs genes, ranging from three and four in chicken and Fugu fish, approximately 25 and 35 members in humans and mice, respectively, to 49 in frog[5]. TAS2Rs are not co-expressed with the other taste receptors TAS1Rs. In addition, TAS1R2 (sweet) and TAS1R1 (umami) are co-expressed with the common subunit TAS1R3 but not with each other. Finally, sour taste receptor (PKD2L1) is present in a population of cells distinct from sweet, bitter and umami taste receptors.

The human bitter taste receptor gene family belongs to the Frizzled/Taste2 subfamily of the G-protein coupled receptors (GPCR)[6,7]. They are expressed on the tongue in bitter taste receptor cells co-expressing specific signal transduction components like Gα–gustducin and able to detect stimuli of only one taste quality[8].

Recently these receptors have been identified in isolated human airway smooth muscle cells (ASM) which also express the cognate G protein gustducin. TAS2R agonists such as saccharin, chloroquine and denatonium evoked increased intracellular calcium in ASM eliciting relaxation of isolated cultured human ASM and dilation of airways greater than that showed by β-adrenergic receptor agonists. In addition, inhaled bitter taste compounds decreased airway

obstruction in a mouse model of asthma[9]. Furthermore, relaxation of precontracted guinea pig trachea (GPT) has been reported as a result of bitter taste receptors activation by bitter tastants (denatonium, chloroquine) in a dose-dependent manner[10].

Quite interestingly, a genome-wide expression array analyses performed on blood leukocytes collected from problematic severe asthmatic subjects (GINA 4 level) revealed the upregulation of TAS2Rs, which was subsequently confirmed by quantitative PCR (qPCR). Besides, chloroquine and denatonium significantly inhibited the LPS-induced release of the inflammatory mediators TNF-α, IL-13 and MCP-1 in blood leukocytes. Furthermore, chloroquine inhibited the release of IL-1β, IL-2, IL-4, IL-5, IL-10, IL-17, G-CSF, GM-CSF, IFN-c and PGE2 in a dose-dependent manner[11]. These findings suggest that the activation of bitter taste receptors may have anti-inflammatory actions further to bronchodilation. This study revealed for the first time a connection between TAS2R expression and clinical characteristics.

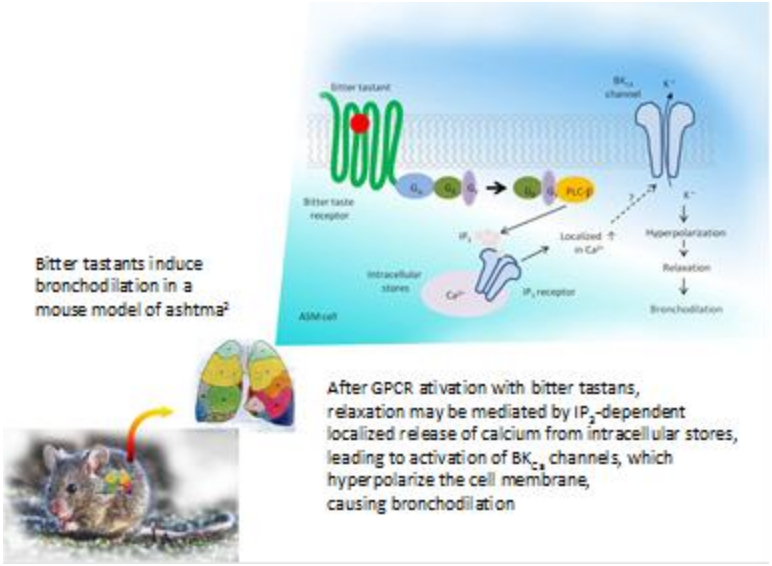


FIGURE 1. Proposed mechanism of action leading to relaxation of isolated ASM and dilation of airways[9].

It has been proposed that the relaxation induced by TAS2R activation in ASM cells could be related to a localized Ca^{2+} response at the cell membrane which

opens large-conductance Ca^{2+}-activated K^+ (BK_{Ca}) channels, leading to ASM hyperpolarization (**Figure 1**)[9].

However, further studies have shown that BK_{Ca} channel blockers either failed to prevent[12] or only partially blocked[13] relaxation caused by the bitter tastant chloroquine, and direct recording of the channel current with patch clamp revealed that bitter tastants did not activate BK_{Ca} channels in mouse airway preparations[12].

In light of these evidences the mechanism proposed by Deshpande *et al.*[9] was questioned by Zhang and coworkers[14], who observed that in mouse primary ASM cells bitter tastants neither evoke localized Ca^{2+} events nor alter spontaneous local Ca^{2+} transients. Instead, they increase global intracellular $[Ca^{2+}]i$, although to a much lower level than bronchoconstrictors. This effect seems to be mediated by the canonical bitter taste signaling cascade (i.e. TAS2R gustducin-phospholipase Cβ [PLCβ]-inositol 1,4,5-triphosphate receptor [IP3R]), which is not sufficient to impact airway contractility. On the contrary, activation of TAS2Rs fully reverses the increase in $[Ca^{2+}]i$ induced by bronchoconstrictors[14], and the reduction of $[Ca^{2+}]i$ is necessary for bitter tastant-induced ASM cell relaxation. Zhang and coworkers[14] proposed that the activation of TAS2R leads to inhibition of L-type voltage-dependent Ca^{2+} channels (VDCCs), the major contributor to Ca^{2+} influx and sustained contraction in response to metacholine (Mch) in mouse airways. This effect seems to be mediated by G-protein βγ pathway rather than by PLCβ-IP3R as it can be prevented by pertussis toxin and G-protein βγ subunit inhibitors.

Even though the mechanism underlying airways relaxation needs further investigation, all these experiments suggest that TAS2Rs agonists can have a protective role in asthma, following a pathway different from that of known β-adrenergic receptor agonists. Hence, TAS2Rs appear to be attractive targets for drug discovery in the field of respiratory obstructive diseases.

Cognate bitter compounds have been identified for 20 out of 25 human TAS2Rs. Besides, the screening of a diverse set of approximately 100 bitter taste substances in a recombinant system expressing 25 hTAS2Rs provided valuable information about structural and pharmacophoric requirements to bind these receptors[15].

Recently Tizzano *et al.*[16] have shown that nasal SCCs which express TAS2Rs along with their downstream signaling components including Gα-gustducin and

TrpM5, detect denatonium benzoate and acyl–homoserine lactones (AHL), utilized as quorum-sensing molecules by Gram-negative bacteria. Isolated SCCs respond to AHL signals by increasing intracellular Ca^{2+}. Furthermore, activation of the trigeminal nerve by these compounds is observed in the form of changes in respiration (respiratory depression and apnea). Genetic deletion of either Gα-gustducin or TrpM5 eliminates the trigeminal nerve evoked reflex to AHLs. Thus functional SCCs are necessary for the AHL-evoked trigeminal response. AHLs are the first natural ligand described for the nasal SCCs and the detection of bacterial quorum-sensing molecules by SCCs is the first demonstration of a clear function for these cells. Since the SCCs are innervated by peptidergic polymodal nociceptor fibers, activation of this system by bacterial AHLs will trigger a local neurogenic inflammatory and immune response to fight the bacterial invasion. Unfortunately, no information has been reported about the selectivity profile of AHLs towards hTAS2Rs.

Bitter taste receptors and Gα-gustducin are also expressed in the gut where they are involved in the chemosensation of nutrients. Janssen and coworkers[17a] have reported that bitter taste receptors and Gα-gustducin can regulate the secretion of ghrelin, a hunger hormone with gastroprokinetic properties, and consequently they could play a role on food intake and gastric emptying. Hence, bitter taste agonists may be developed for the treatment of weight and gastrointestinal motility disorders. Further to hormonal secretion in the gut, Dotson and collaborators[17b] have reported that a loss-of-function variant of a human TAS2R is associated with glucose dysregulation and the presence of Type 2 diabetes mellitus in a human population. Furthermore, bitter-tasting compounds can elicit glucagon-like peptide-1 (GLP-1) secretion from enteroendocrine L cells in a receptor-dependent manner, thus suggesting a mechanism by which TAS2R function could influence glucose homeostasis. The interplay between gustatory and endocrine systems support a role for the canonical “taste” molecules in the maintenance of metabolic homeostasis and suggest that sensory function may be modulated in the context of an animal’s metabolic status.

Finally, bitter taste receptor subtype TAS2R46 appears to be expressed also on the surface of human mesenchimal stromal cells and on vascular smooth muscle cells. Rats exposed to the TAS2R46 agonist denatonium showed a transient and significant drop in blood pressure followed by recovery, providing in vivo evidence that a bitter compound can modulate the vascular tone through engagement of its receptor[17c].

Here we report the development of receptor models of human TAS2R10, TAS2R14 and TAS2R31, the bitter taste receptors most abundantly expressed in ASM, which were used to carry out docking studies with strychnine, flufenamic acid and aristolochic acid, the most potent and selective ligands for these receptor subtypes, respectively (**Figure 2**). Results of the docking studies will be presented in detail and discussed in light of available site-directed mutagenesis experiments[18] and literature SAR data. Besides, here we report different chemoinformatics analyses performed on a huge set of bitter tastants described in the literature.

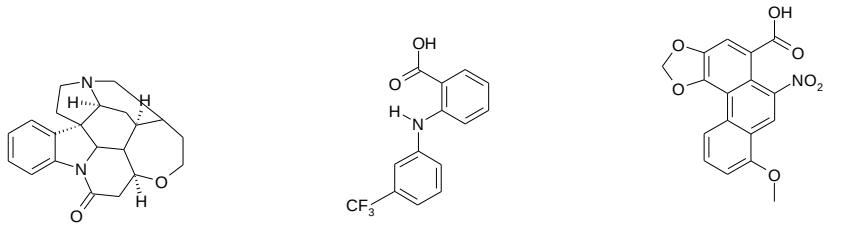


FIGURE 2. Structure of strychnine, flufenamic acid and aristolochic acid. EC_{50} values were taken from Meyerhof *et al.*[15].

Methods

2.1 Sequence analyses

Human bitter taste receptor sequences were retrieved from UniProtKB (http://www.uniprot.org/help/uniprotkb) (**Appendix 1**) and aligned with the use of BLASTP 2.2.25[19]. According to the new numbering system, hTAS2R31 corresponds to hTAS2R44 described in Meyerhof *et al.*[15]. Sequence identity of human TAS2Rs with respect to 7TM receptors whose X-ray structures are available was calculated with GPCRDB (12.01.2011) alignment engine available at http://www.gpcr.org/7tm/ (Accession dates: 28/02/2011 and 07/06/2011). Furthermore sequence conservation of human with respect to mouse and guinea pig sequences was evaluated with the use of BLAST and default parameters.

2.2 Computational analyses of bitter tastants

The biological fingerprint of each hTAS2R was built by using potency data generated for 104 bitter tastants[3, 15, 20-27], which were converted into a bit string of 1 and 0 depending on the ligand response at each receptor i.e. 1 = activation, 0 = no activation. Therefore, each human receptor was described by a 104-bit length string. Then, the biological similarity of each hTAS2R with respect to the others was calculated with the use of Tanimoto index and the biological fingerprints described above (for details, see references #6 and #7, Computational Methods Chapter of this thesis). Tanimoto similarity index is c/(a+b-c) where a is the number of bits on in receptor A, b is the number of bits on in receptor B, c is the number of bits in common between A and B. Two receptors exhibiting the same profile have Tanimoto = 1, while if nothing is in common, Tanimoto = 0.

The structures of 97 compounds out of the 104 described by Meyerhof *et al.*[15] were retrieved from PubChem database (http://pubchem.ncbi.nlm.nih.gov/). Bitter tastants cluster analysis was performed with LibraryMCS cluster algorithm and ChemAxon v5.4 1024 Chemical Fingerprints[28].

Furthermore, the structure of 550 bitter tastants were retrieved from BitterDB database (http://bitterdb.agri.huji.ac.il/bitterdb/dbbitter.php, accessed 15/06/2011)[29] in sdf format. The compounds were desalted with the use of in house Perl script. Cations and labeled derivatives were removed. After standardization with the use of ChemAxon v5.4 Standardizer (neutralize, aromatize, clean2D option) 513 derivatives were obtained. It was noticed that configuration of the stereogenic centers appearing in these derivatives had not been specified for the compounds retrieved. Therefore, it was assumed that the associated biological data had been generated for their racemates. Then, they were described with the use of the following ChemAxon v5.4 descriptors: Mol Weight, MASS, LOGP, AVERAGE_MOL_POLARIZABILITY, MOL_POLARIZABILITY, FORMAL_CHARGE, ATOMCOUNT, ALIPHATIC_ATOMCOUNT, AROMATIC_ATOMCOUNT, BONDCOUNT, ALIPHATIC_BONDCOUNT, AROMATIC_BONDCOUNT, RINGCOUNT, HETERO_RINGCOUNT, HETEROALIPHATIC_RINGCOUNT, HETEROAROMATIC_RINGCOUNT, RINGATOMCOUNT, RINGBONDCOUNT, CHAINATOMCOUNT, CHAINBONDCOUNT, SMALLEST_RINGSIZE, LARGEST_RINGSIZE,ALIPHATIC_RINGCOUNT, AROMATIC_RINGCOUNT, PSA, Volume, REFRACTIVITY,

ACCEPTOR_COUNT, DONOR_COUNT, ROTATABLE_BONDCOUNT, LogD.

Calculation settings are reported in **Appendix 1**. Bitter taste compound ID of the BitterDB was maintained.

Structural and pharmacophoric diversity of the compounds was assessed with the use of ChemAxon v5.4 tools as reported in **Appendix 2**. In particular, structure and pharmacophores of hTAS2R14, hTAS2R10 and hTAS2R31 agonists were analyzed in detail.

Subsequently, Principal Component Analysis (PCA) of the descriptor matrix (513 objects X 31 descriptors) was performed with the use of SIMCA-P^{+} [30]. Compounds were color coded according to their activity profile with respect to hTAS2R10, hTAS2R14 and hTAS2R31 as follows: agonists of hTAS2R10 only (pink); agonists of hTAS2R14 only (orange); agonists of hTAS2R31 only (light blue); agonists of both hTAS2R10 and hTAS2R14 (green); agonists of both hTAS2R10 and hTAS2R31 (blue); agonists of both hTAS2R14 and hTAS2R31 (red); agonists of hTAS2R10, hTAS2R14 and hTAS2R31 (bordeaux); agonists of neither these three subtypes (black).

2.3 Modelling of TAS2Rs

Homology models of hTAS2R10, hTAS2R14 and hTAS2R31 were based on the crystal structure of the nanobody-stabilized active state of the human β_2 adrenergic receptor (PDB code: 3P0G)[31]. Sequence alignments of bitter taste receptors to the structure template were generated within MOE[32] and manually adjusted by taking into account TM helices predictions reported by Adler *et al.*[2]. The 3D models were generated and refined within MOE. Prior to run docking experiments, receptor models were processed with the use of Protein preparation workflow available within Maestro[33]. Binding sites were identified with the use of Sitemap[33] and utilized as starting points for Induced Fit Docking simulations (IFD)[33] to dock strychnine, flufenamic acid and aristolochic acid in hTAS2R10, hTAS2R14 and hTAS2R31, respectively (**Figure 1**). Poses were visually inspected within Maestro. Finally, strychnine and flufenamic acid were docked in hTAS2R10 and hTAS2R14 after removal of their extracellular loops and with the use of IDF (default parameters) for comparison. The center of the box was defined as the centroid of F107 and S274 for hTAS2R10; F108 and M270 for hTAS2R14. Finally, aristolochic acid was docked in the hTAS2R31 receptor using IFD and a wider binding site including extracellular loops

according to the procedure described by Vilar *et al.*[34]. (residues:_:100,_:101,_:102,_:103,_:104,_:105,_:106,_:107,_:108,_:109,_:110,_:111,_:144,_:145,_:146,_:147,_:148,_:149,_:150,_:151,_:152,_:153,_:154,_:155,_:156,_:157,_:158,_:159,_:160,_:161,_:162,_:163,_:164,_:165,_:166,_:167,_:168,_:169,_:170,_:171,_:172,_:173,_:174,_:175,_:176,_:177,_:178,_:179,_:180,_:181,_:182,_:183,_:184,_:185,_:186,_:187,_:188,_:189,_:190,_:191,_:192,_:193,_:194,_:195,_:235,_:236,_:237,_:238,_:239,_:240,_:241,_:242,_:243,_:244,_:245,_:246,_:247,_:248,_:249,_:250,_:251,_:252,_:253,_:254,_:255,_:256,_:257,_:258,_:259,_:260,_:261,_:262,_:263,_:264,_:265,_:266,_:267,_:268,_:269,_:270,_:271,_:272,_:58,_:59,_:60,_:61,_:62,_:63,_:64,_:65,_:66,_:67,_:68,_:69,_:70,_:71,_:72,_:73,_:74,_:75,_:76,_:77,_:78,_:79,_:80,_:81,_:82,_:83,_:84,_:85,_:86,_:87,_:88,_:89,_:90,_:91,_:92,_:93,_:94,_:95,_:96,_:97,_:98,_:99)

Results and discussion

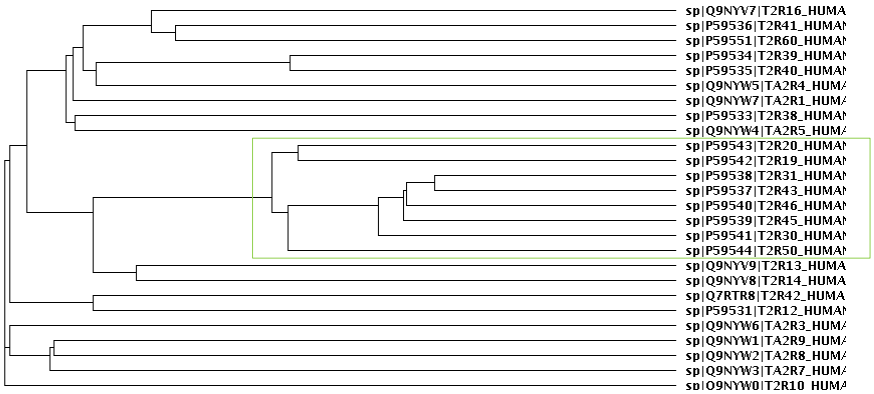
(i) Sequence analyses

Sequences of human bitter taste receptors were retrieved from UNIProtKB/Swiss-prot (**Appendix 1**). According to the sequence alignments obtained, TAS2Rs exhibit overall low identity among each other (**Figure 3**, top) as well as to GPCRs whose X-ray structures have been reported in the literature (**Figure 3**, bottom).

hTAS2Rs are relatively short GPCR proteins (290 -340 amino acids), with short N termini and C termini. The extracellular loop and the top of TM3 are frequently associated with Single Nucleotide Polymorphism (SNP) in mouse cycloheximide non-tasters (mouse T2R5) and human T2R4 (**Figure 4**). Amino acids crucial for 6-nitrosaccharin and N-isopropyl-2-methyl-5-nitrobenzenesulfonamide (IMNB) binding to T2R43 were identified in extracellular loops[24a].

All hTAS2Rs contain a conserved N-glycosylation site in the 2nd extracellular loop (NXS/T)[35]. In the literature it has been reported that the hypoglycosilated mutant N163Q hTAS2R16 transiently expressed in HEK293T cells stably expressing the G-protein chimera Gα16gust44 is non-functional[18]. On the contrary, hTAS2R16 WT shows a clear stimulation by its agonist salicin[20]. Similar observation were made using the glycosylated receptor hTAS2R46 and its corresponding non-glycosylated mutant hTAS2R46 N161Q[25]. According to these experiments, the glycan structures may not be directly involved in agonist

recognition, but rather play a role during the biosynthetic pathway of hTAS2Rs. Potential glycosylation sites have been identified in hTAS2R10 (amino acid sequence number N92 and N158), hTAS2R14 (N153, N162 and N171) and hTAS2R31 (N14 and N161). In particular, N153 in hTAS2R14 has also been described in the literature[36].



% Identity	T2R10_human	T2R14_human	T2R31_human
opsd_bovin	23.69	16.98	31.25
adrb2_human	22.22	16	33.33
adrb1_melga	18.2	16.36	21
cxcr4_human	23.53	22.02	21.38
drd3_human	18.37	37.5	21.05
aa2ar_human	18.54	16.67	29.07

FIGURE 3. Top. Dendrogram of human TAS2Rs sequence relationship. Bottom. TAS2Rs sequence identity with GPCRs for which X-ray structures are available. Sequence identity was calculated with GPCRDB (12.01.2011) alignment engine available at http://www.gpcr.org/7tm/ Accession dates: 28/02/2011 and 07/06/2011.

Limited overall sequence conservation was observed among hTAS2R31, hTAS2R10 and hTASR2R14 and mouse bitter taste receptors (**Figure 5**, **Figure 6** and **Figure 7**).

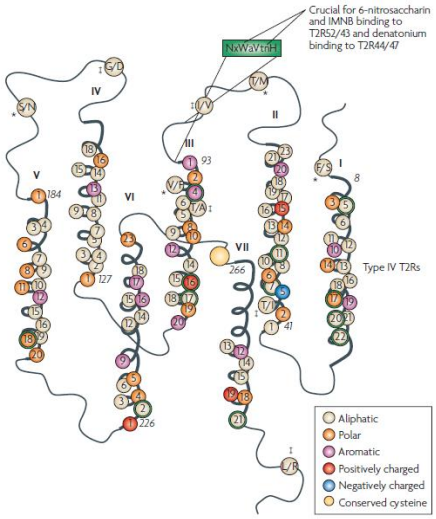
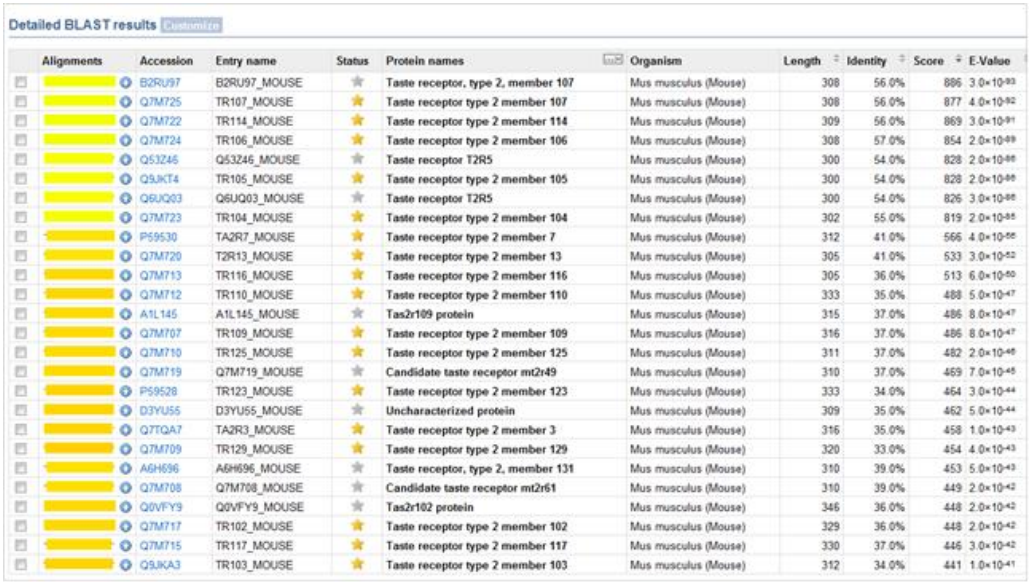


FIGURE 4. Conserved features and structural motifs of hTAS2Rs. Consensus of the alignment generated with ClustalW v1.8 of 8 related TAS2R i.e. 43, 44(31), 45, 46, 47(30), 48 (19), 49(20) and 50.[7]

Detailed BLAST results Customize

Aligements	Accession	Entry name	Status	Protein names	Organism	Length	Identity	Score	E-Value
	Q7M721	TR120_MOUSE		Taste receptor type 2 member 120	Mus musculus (Mouse)	295	52.0%	758	2.0×10^{-78}
	Q7TQA8	TR136_MOUSE		Taste receptor type 2 member 136	Mus musculus (Mouse)	305	47.0%	689	2.0×10^{-70}
	Q7M710	TR125_MOUSE		Taste receptor type 2 member 125	Mus musculus (Mouse)	311	44.0%	604	2.0×10^{-60}
	Q7M713	TR116_MOUSE		Taste receptor type 2 member 116	Mus musculus (Mouse)	305	40.0%	564	7.0×10^{-56}
	Q7M719	Q7M719_MOUSE		Candidate taste receptor mt2r49	Mus musculus (Mouse)	310	41.0%	540	4.0×10^{-53}
	Q0VFY9	Q0VFY9_MOUSE		Tas2r102 protein	Mus musculus (Mouse)	346	41.0%	539	6.0×10^{-53}
	Q7M717	TR102_MOUSE		Taste receptor type 2 member 102	Mus musculus (Mouse)	329	41.0%	539	6.0×10^{-53}
	Q7TQA4	TR140_MOUSE		Taste receptor type 2 member 140	Mus musculus (Mouse)	312	41.0%	537	1.0×10^{-52}
	Q7M712	TR110_MOUSE		Taste receptor type 2 member 110	Mus musculus (Mouse)	333	40.0%	535	2.0×10^{-52}
	Q7M720	T2R13_MOUSE		Taste receptor type 2 member 13	Mus musculus (Mouse)	305	40.0%	531	5.0×10^{-52}
	Q7M711	TR113_MOUSE		Taste receptor type 2 member 113	Mus musculus (Mouse)	309	40.0%	529	8.0×10^{-52}
	Q7M718	TR124_MOUSE		Taste receptor type 2 member 124	Mus musculus (Mouse)	309	39.0%	523	4.0×10^{-51}
	P59528	TR123_MOUSE		Taste receptor type 2 member 123	Mus musculus (Mouse)	333	40.0%	512	8.0×10^{-50}
	Q7M715	TR117_MOUSE		Taste receptor type 2 member 117	Mus musculus (Mouse)	330	38.0%	500	2.0×10^{-48}
	A1L145	A1L145_MOUSE		Tas2r109 protein	Mus musculus (Mouse)	315	38.0%	488	5.0×10^{-47}
	Q7M707	TR109_MOUSE		Taste receptor type 2 member 109	Mus musculus (Mouse)	316	38.0%	488	5.0×10^{-47}
	P59530	TA2R7_MOUSE		Taste receptor type 2 member 7	Mus musculus (Mouse)	312	39.0%	462	5.0×10^{-44}
	Q9JKA3	TR103_MOUSE		Taste receptor type 2 member 103	Mus musculus (Mouse)	312	39.0%	460	8.0×10^{-44}

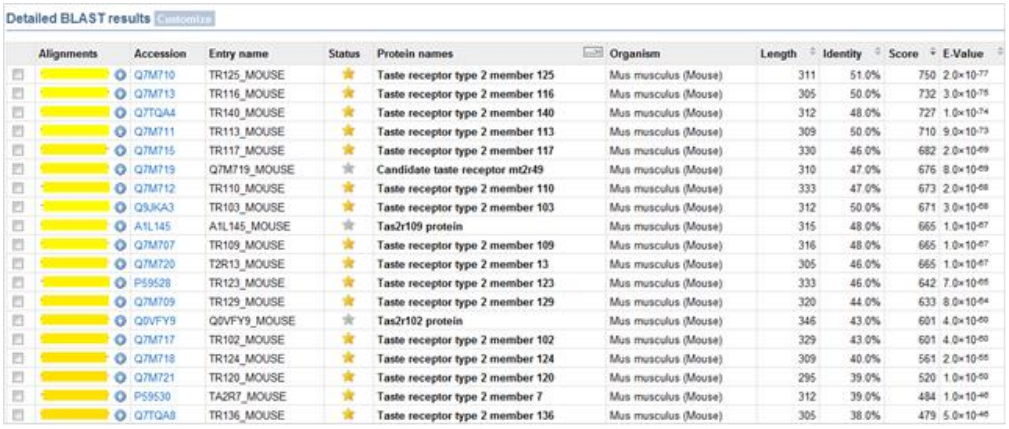
FIGURE 5. Human TAS2R31 (sp|P59538|T2R31_HUMAN) BLAST search for mouse orthologous exhibiting at least 30% of sequence identity.



Detailed BLAST results Customize

Alignments	Accession	Entry name	Status	Protein names	Organism	Length	Identity	Score	E-Value
	B2RU97	B2RU97_MOUSE		Taste receptor, type 2, member 107	Mus musculus (Mouse)	308	56.0%	886	3.0×10^{-93}
	Q7M725	TR107_MOUSE		Taste receptor type 2 member 107	Mus musculus (Mouse)	308	56.0%	877	4.0×10^{-92}
	Q7M722	TR114_MOUSE		Taste receptor type 2 member 114	Mus musculus (Mouse)	309	56.0%	869	3.0×10^{-91}
	Q7M724	TR106_MOUSE		Taste receptor type 2 member 106	Mus musculus (Mouse)	308	57.0%	854	2.0×10^{-89}
	Q53Z46	Q53Z46_MOUSE		Taste receptor T2R5	Mus musculus (Mouse)	300	54.0%	828	2.0×10^{-86}
	Q9JKT4	TR105_MOUSE		Taste receptor type 2 member 105	Mus musculus (Mouse)	300	54.0%	828	2.0×10^{-86}
	Q6UQ03	Q6UQ03_MOUSE		Taste receptor T2R5	Mus musculus (Mouse)	300	54.0%	826	3.0×10^{-86}
	Q7M723	TR104_MOUSE		Taste receptor type 2 member 104	Mus musculus (Mouse)	302	55.0%	819	2.0×10^{-85}
	P59530	TA2R7_MOUSE		Taste receptor type 2 member 7	Mus musculus (Mouse)	312	41.0%	566	4.0×10^{-56}
	Q7M720	T2R13_MOUSE		Taste receptor type 2 member 13	Mus musculus (Mouse)	305	41.0%	533	3.0×10^{-52}
	Q7M713	TR116_MOUSE		Taste receptor type 2 member 116	Mus musculus (Mouse)	305	36.0%	513	6.0×10^{-50}
	Q7M712	TR110_MOUSE		Taste receptor type 2 member 110	Mus musculus (Mouse)	333	35.0%	488	5.0×10^{-47}
	A1L145	A1L145_MOUSE		Tas2r109 protein	Mus musculus (Mouse)	315	37.0%	486	8.0×10^{-47}
	Q7M707	TR109_MOUSE		Taste receptor type 2 member 109	Mus musculus (Mouse)	316	37.0%	486	8.0×10^{-47}
	Q7M710	TR125_MOUSE		Taste receptor type 2 member 125	Mus musculus (Mouse)	311	37.0%	482	2.0×10^{-46}
	Q7M719	Q7M719_MOUSE		Candidate taste receptor mt2r49	Mus musculus (Mouse)	310	37.0%	469	7.0×10^{-45}
	P59528	TR123_MOUSE		Taste receptor type 2 member 123	Mus musculus (Mouse)	333	34.0%	464	3.0×10^{-44}
	D3YU55	D3YU55_MOUSE		Uncharacterized protein	Mus musculus (Mouse)	309	35.0%	462	5.0×10^{-44}
	Q7TQA7	TA2R3_MOUSE		Taste receptor type 2 member 3	Mus musculus (Mouse)	316	35.0%	458	1.0×10^{-43}
	Q7M709	TR129_MOUSE		Taste receptor type 2 member 129	Mus musculus (Mouse)	320	33.0%	454	4.0×10^{-43}
	A6H696	A6H696_MOUSE		Taste receptor, type 2, member 131	Mus musculus (Mouse)	310	39.0%	453	5.0×10^{-43}
	Q7M708	Q7M708_MOUSE		Candidate taste receptor mt2r61	Mus musculus (Mouse)	310	39.0%	449	2.0×10^{-42}
	Q0VFY9	Q0VFY9_MOUSE		Tas2r102 protein	Mus musculus (Mouse)	346	36.0%	448	2.0×10^{-42}
	Q7M717	TR102_MOUSE		Taste receptor type 2 member 102	Mus musculus (Mouse)	329	36.0%	448	2.0×10^{-42}
	Q7M715	TR117_MOUSE		Taste receptor type 2 member 117	Mus musculus (Mouse)	330	37.0%	446	3.0×10^{-42}
	Q9JKA3	TR103_MOUSE		Taste receptor type 2 member 103	Mus musculus (Mouse)	312	34.0%	441	1.0×10^{-41}

FIGURE 6. Human TAS2R10 (sp|Q9NYW0|T2R10_HUMAN) BLAST search for mouse orthologues exhibiting at least 30% of sequence identity



Detailed BLAST results Customize

Alignments	Accession	Entry name	Status	Protein names	Organism	Length	Identity	Score	E-Value
	Q7M710	TR125_MOUSE		Taste receptor type 2 member 125	Mus musculus (Mouse)	311	51.0%	750	2.0×10^{-77}
	Q7M713	TR116_MOUSE		Taste receptor type 2 member 116	Mus musculus (Mouse)	305	50.0%	732	3.0×10^{-75}
	Q7TQA4	TR140_MOUSE		Taste receptor type 2 member 140	Mus musculus (Mouse)	312	48.0%	727	1.0×10^{-74}
	Q7M711	TR113_MOUSE		Taste receptor type 2 member 113	Mus musculus (Mouse)	309	50.0%	710	9.0×10^{-73}
	Q7M715	TR117_MOUSE		Taste receptor type 2 member 117	Mus musculus (Mouse)	330	46.0%	682	2.0×10^{-69}
	Q7M719	Q7M719_MOUSE		Candidate taste receptor mt2r49	Mus musculus (Mouse)	310	47.0%	676	8.0×10^{-69}
	Q7M712	TR110_MOUSE		Taste receptor type 2 member 110	Mus musculus (Mouse)	333	47.0%	673	2.0×10^{-68}
	Q9JKA3	TR103_MOUSE		Taste receptor type 2 member 103	Mus musculus (Mouse)	312	50.0%	671	3.0×10^{-68}
	A1L145	A1L145_MOUSE		Tas2r109 protein	Mus musculus (Mouse)	315	48.0%	665	1.0×10^{-67}
	Q7M707	TR109_MOUSE		Taste receptor type 2 member 109	Mus musculus (Mouse)	316	48.0%	665	1.0×10^{-67}
	Q7M720	T2R13_MOUSE		Taste receptor type 2 member 13	Mus musculus (Mouse)	305	46.0%	665	1.0×10^{-67}
	P59528	TR123_MOUSE		Taste receptor type 2 member 123	Mus musculus (Mouse)	333	46.0%	642	7.0×10^{-65}
	Q7M709	TR129_MOUSE		Taste receptor type 2 member 129	Mus musculus (Mouse)	320	44.0%	633	8.0×10^{-64}
	Q0VFY9	Q0VFY9_MOUSE		Tas2r102 protein	Mus musculus (Mouse)	346	43.0%	601	4.0×10^{-60}
	Q7M717	TR102_MOUSE		Taste receptor type 2 member 102	Mus musculus (Mouse)	329	43.0%	601	4.0×10^{-60}
	Q7M718	TR124_MOUSE		Taste receptor type 2 member 124	Mus musculus (Mouse)	309	40.0%	561	2.0×10^{-55}
	Q7M721	TR120_MOUSE		Taste receptor type 2 member 120	Mus musculus (Mouse)	295	39.0%	520	1.0×10^{-50}
	P59530	TA2R7_MOUSE		Taste receptor type 2 member 7	Mus musculus (Mouse)	312	39.0%	484	1.0×10^{-46}
	Q7TQA8	TR136_MOUSE		Taste receptor type 2 member 136	Mus musculus (Mouse)	305	38.0%	479	5.0×10^{-46}

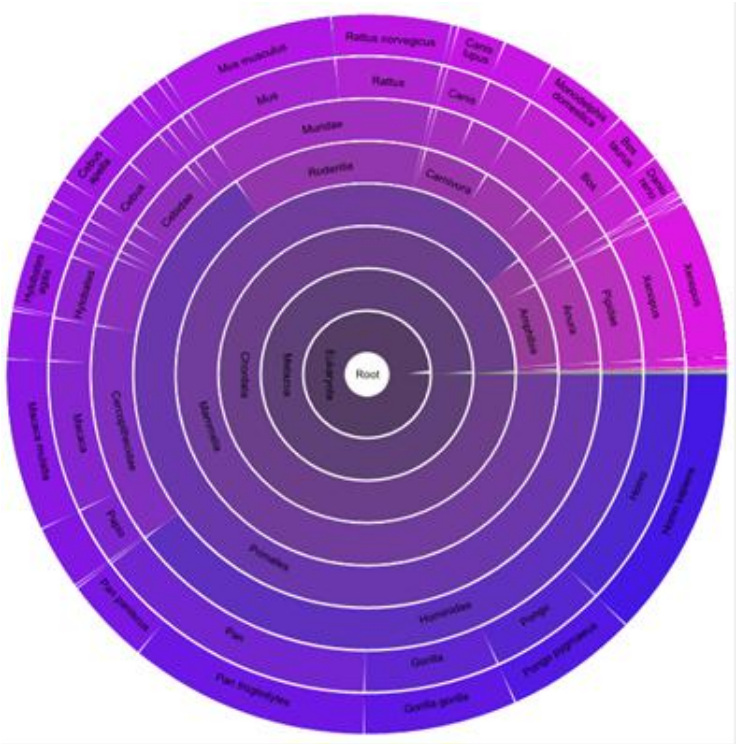
FIGURE 7. Human TAS2R14 (sp|Q9NYV8|T2R14_HUMAN) BLAST search for mouse orthologues exhibiting at least 30% of sequence identity.

Two mouse bitter taste receptors have been extensively described in the literature: mT2R5 (alias TR105_MOUSE, Q9JKT4) which is activated by cycloheximide and exhibits 54% sequence identity to hTAS2R10, and mT2R8

(alias TR140_MOUSE, Q7TQA4) activated by denatonium and 6-n-propyl-2-thiouracil, exhibiting 48% and 41% sequence identity to hTAS2R14 and hTAS2R31, respectively[3]. In addition, TA2R7_MOUSE (P59530) shows 41%, 39% and 39% of sequence identity to hTAS2R10, hTAS2R14 and hTAS2R31 sequences, respectively. Furthermore, TA2R16_MOUSE (P59529) exhibits 52% of sequence identity to hTAS2R16, 33% to hTAS2R41 and 32% to hTAS2R60. None of these subtypes has been found in ASM[9].

No annotated bitter taste receptor sequences were identified in UniProtKB with the use of the following query (database accessed 03/11/2011): taste AND taxonomy:"Guinea pig[10141]" even though recently the discovery of guinea pig bitter taste receptors has been described in the literature[10].

Besides, in the Pfam database the distribution of the bitter taste receptor family across species does not cover guinea pigs as shown in the picture below: (http://pfam.sanger.ac.uk/family?acc=PF05296#tabview=tab7, Accession date: Nov, 23rd 2011)



(ii) Biological fingerprints generation and analysis

Even though hTAS2R10 and hTAS2R14 exhibit low sequence identity, they show a certain degree of similarity calculated with the use of biological fingerprints and Tanimoto index, which reflects the fact that they share several common activators (15 out of the 97 agonists considered). On the contrary, hTAS2R31 (hTAS2R44) exhibits low biological fingerprints similarity with respect to hTAS2R10 and hTAS2R14 as well as to the other subtypes (**Figure 8**).

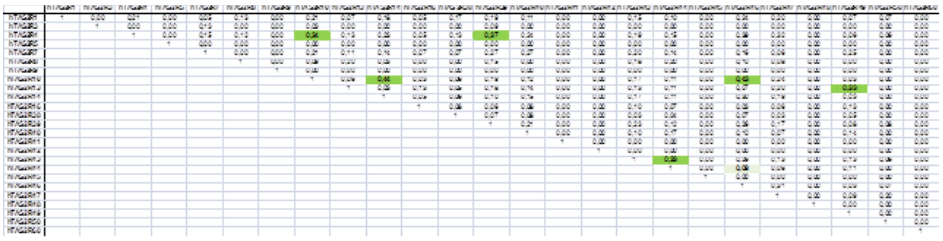


FIGURE 8. Similarity of bitter taste receptors calculated with their biological fingerprints and Tanimoto index. hTAS2R44 corresponds to hTAS2R31.

(iii) Cluster analysis

According to the cluster analysis performed (**Figure 9**), bitter tastants look structurally different among each other suggesting that they can exhibit different binding modes. The 15 hTAS2R10 and hTAS2R14 common activators out of the 97 agonists considered are spread over 7 clusters and 6 singletons. At the same time, these receptors can discriminate even among very similar bitter compounds. For example, cucurbitacin E and B belong to the same cluster (**Figure 9**, #158). While both these ligands are agonists of hTAS2R10, hTAS2R14 is activated by cucurbitacin B only.

These results are consistent with those reported by Meyerhof *et al.*[15]. Even though TAS2Rs bind structurally diverse bitter compounds in either μM or mM affinity range, they are able to discriminate even among structurally similar bitter compounds with very high accuracy. Furthermore, hTAS2R10 and hTAS2R14 are promiscuous receptors showing wide molecular receptive ranges. hTAS2R31 exhibits quite a broad agonist spectrum, lacking common structural motifs responsible for specific recognition. In addition, bitter compounds activate various hTAS2Rs in different concentration ranges,

generally between 10- and 100- fold. Furthermore, it seems that agonist specificity is not related to receptor sequence homology. Therefore, agonist spectra appear to be unique for every single hTAS2R. In light of that, it can reasonably be concluded that while hTAS2R31 is characterized by a broad agonist spectra and recognizes compounds lacking common motifs, hTAS2R10 and hTAS2R14 seem to be tuned to detect even a wider variety of structurally unrelated activators.

As shown, none of these derivatives can be considered a drug-like chemical starting point to initiate a medicinal chemistry project.

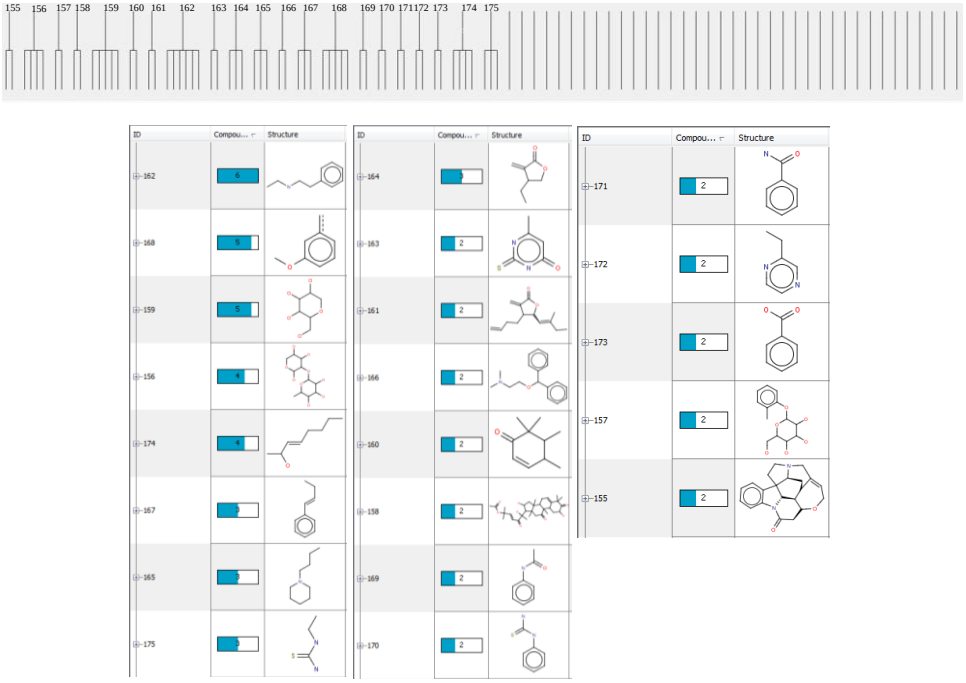


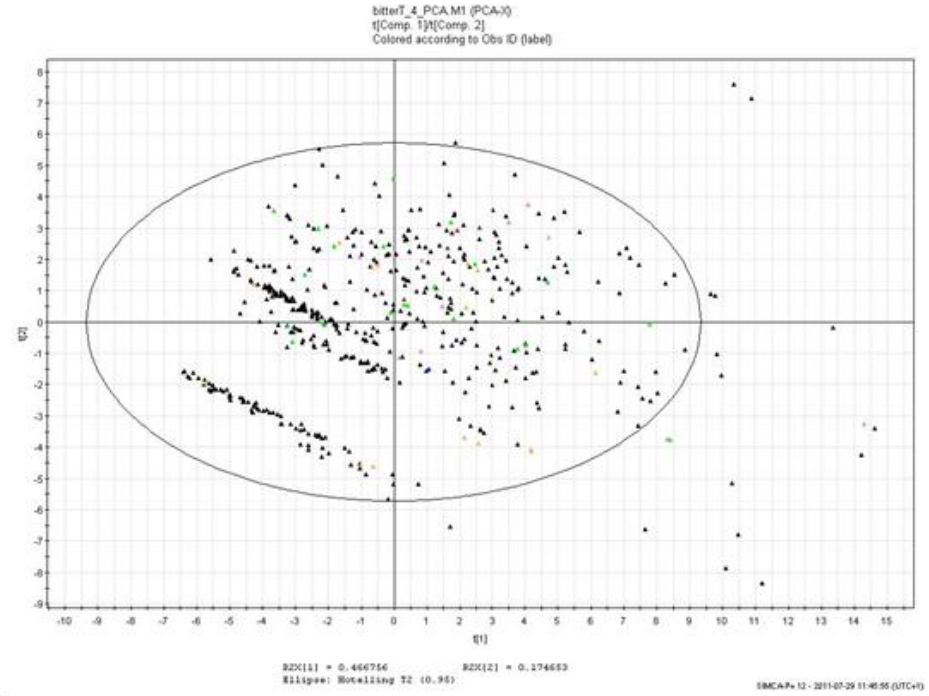
FIGURE 9. Top: Dendrogram of the cluster analysis of 97 bitter tastants described by Meyerhof *et al.*[15] (58 chemotypes, 21 clusters, 37 singletons). **Bottom:** The Maximum Common Substructure of each cluster compound is shown. 9 out of the 15 common activators of both hTAS2R10 and hTAS2R14 belong to clusters #158, #160, #161, #164, #165, #168, #174; the other 6 compounds are singletons.

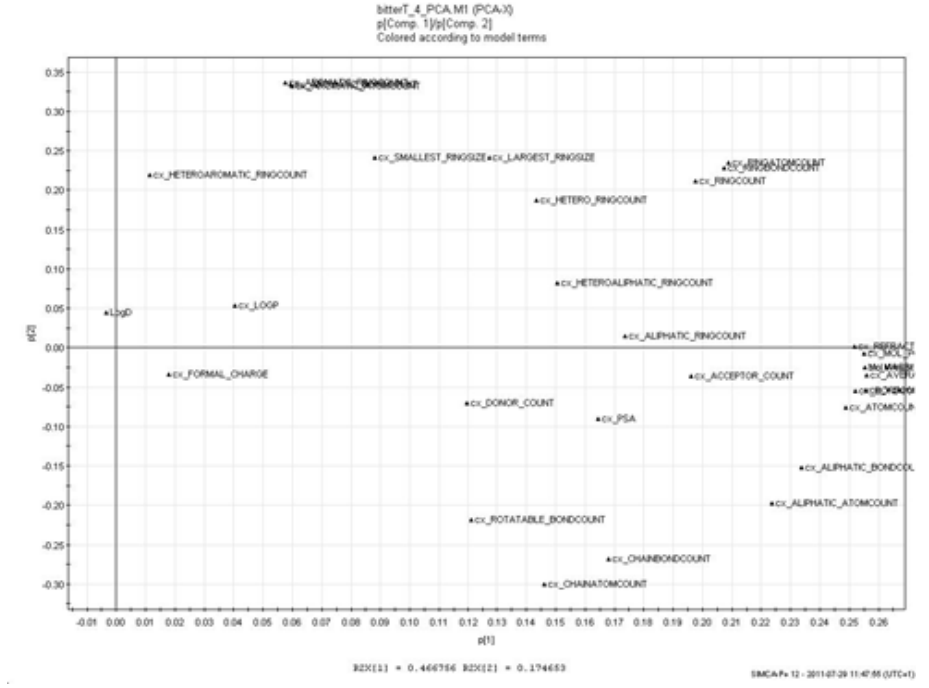
(iv) Principal component analysis

513 bitter tastants retrieved from BitterDB database were described with 31 descriptors and then PCA was carried out on the 513X31 descriptor matrix. The first 4 components account for 84% of the total variance ($R^2 = 0.84$, $Q^2 = 0.72$). In **Figure 10** (top) is reported the score plot of PC1 vs. PC2 while in **Figure 10** (middle) the loading plot of PC1 vs. PC2 is shown. In the score plot several bitter tastants appear to be strong outliers and no clusters were found except for a huge group of compounds appearing at the bottom left of the plot. Loadings of PC1 are positively correlated to bulk descriptors (volume, MW). The PC2 loadings are positively correlated to aromatic atom and ring counts and negatively correlated to aliphatic atom count and chain. Both LogP and LogD have almost negligible coefficients in both PC1 and PC2. Upon removal of the strongest outliers (i.e. the derivatives lying outside the Hotelling-T curve of the score plot) no clusters were identified (**Figure 10**, bottom). 4 components account for 82% of the total variance. Furthermore, ligands of hTAS2R31, hTAS2R10 and hTAS2R14 do not occupy a preferred physico-chemical space supporting the fact that these bitter receptors are activated by ligands quite different in terms of structure as well as physico-chemical properties.

The average structural diversity defined as (1 – Tanimoto_index) of the 513 bitter compounds retrieved from BitterDB was assessed with the use of ChemAxon structural fingerprints and Tanimoto index. As a result, the set is characterized by structural ave_dissimilarity = 0.80 ± 0.10, suggesting that they are very structurally heterogeneous. Similarly, the pharmacophoric ave_dissimilarity measured with the use of ChemAxon pharmacophoric fingerprints accounts for 0.82 ± 0.14.

Subsequently, the agonists of hTAS2R10, hTAS2R14 and hTAS2R31 were analyzed in detail.


bitterT_4_PCA.M1 (PCA-X)
t[Comp. 1]/t[Comp. 2]
Colored according to Obs ID (label)
t[2]
t[1]
R2X[1] = 0.466756
R2X[2] = 0.174653
Ellipse: Hotelling T2 (0.95)
SIMCA-P+ 12 - 2011-07-29 11:46:55 (UTC+1)


bitterT_4_PCA.M1 (PCA-X)
p[Comp. 1]/p[Comp. 2]
Colored according to model terms
p[2]
p[1]
CX_SMALLEST_RINGSIZE
CX_LARGEST_RINGSIZE
CX_HETEROAROMATIC_RINGCOUNT
CX_RINGCOUNT
CX_HETERO_RINGCOUNT
CX_HETEROALIPHATIC_RINGCOUNT
LogD
CX_LOGP
CX_ALIPHATIC_RINGCOUNT
CX_FORMAL_CHARGE
CX_ACCEPTOR_COUNT
CX_DONOR_COUNT
CX_PSA
CX_ALIPHATIC_BONDCOU
CX_ALIPHATIC_ATOMCOUNT
CX_ROTATABLE_BONDCOUNT
CX_CHAINBONDCOUNT
CX_CHAINATOMCOUNT
R2X[1] = 0.466756 R2X[2] = 0.174653
SIMCA-P+ 12 - 2011-07-29 11:47:55 (UTC+1)

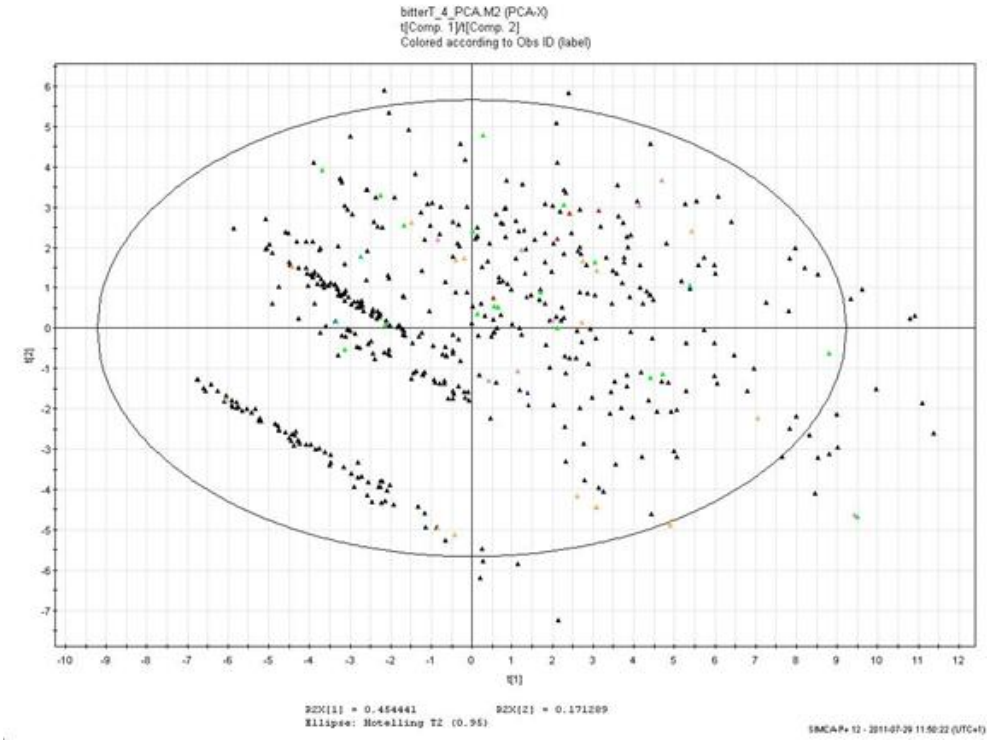


FIGURE 10. Top: Score plot of PC1 vs. PC2. Color coding: agonists of hTAS2R10 only (pink); agonists of hTAS2R14 only (orange); agonists of hTAS2R31 only (light blue); agonists of hTAS2R10 and hTAS2R14 (green); agonists of hTAS2R10 and hTAS2R31 (blue); agonists of hTAS2R14 and hTAS2R31 (red); agonists of hTAS2R10, hTAS2R14 and hTAS2R31 (bordeaux); agonists of neither the three subtypes (black). **Middle:** Loading plot of PC1 vs. PC2. **Bottom:** Score plot of PC1 vs. PC2 upon removal of the strongest outliers appearing in the score plot at the top.

8 hTAS2R31 agonists were extracted from BitterDB database (i.e. quinidine, aristolochic_acid, parthenolide, diphenidol, famotidine, acesulfame_K, aloin, saccharin). On average they show structural and pharmacophore dissimilarity among each other of 0.65±0.25 and 0.73±0.29, respectively.

40 hTAS2R14 agonists were fished out from BitterDB (i.e. Quinidine, Lupulone, Benzoin, Absinthin, Arborescin, Arglabin, Aristolochic_acid,

Artemorin, Caffeine, Campher, Cascarillin, Coumarin, Cucurbitacin_B, Falcarindiol, Noscapine, Papaverine, Parthenolide, Quassin, Azathioprine, Benzamide, Carisoprodol, Chlorhexidine, Chlorpheniramine, Diphenhydramine, diphenidol, thiocarbanilide, Divinyl_sulfoxide, Flufenamic_Acid, Haloperidol, Sodium_Benzoate, Isoxanthohumol, Xanthohumol, trans-isohumulone, trans-isocohumulone, Adlupulone, cis-isoloadhumulone, chloroquine, 8-prenylnaringenin, Picrotoxinin, Thujone). These derivatives exhibit on average structural and pharmacophoric diversity among each other of 0.74±0.17 and 0.78±0.21.

31 hTAS2R10 agonists were extracted from BitterDB database (Quinidine, Benzoin, Absinthin, Arborescin, Arglabin, Artemorin, Caffeine, Campher, Cascarillin, Coumarin, Cucurbitacin_B, Papaverine, Parthenolide, Quassin, Azathioprine, Chlorpheniramine, diphenidol, Haloperidol, chloramphenicol, yohimbine, dextromethorphan, chloroquine, Picrotoxinin, Thujone, dapsone, denatonium_benzoate, Cucurbitacins, Cycloheximide, erythromycin, strychnine, famotidine). Their average structural and pharmacophoric dissimilarity among each other accounts for 0.71 ± 0.18 and 0.76 ± 0.21, respectively.

(v) Modelling studies

The crystal structure of the nanobody-stabilized active state of the β_2 adrenoceptor[31] was selected as template to build the model of TAS2Rs. This X-ray complex (PDB code: 3P0G) is very similar to that of β_2 adrenoceptor in the T4L-β_2AR-Gs complex[37] (PDB code: 3SN6). This high structural similarity is consistent with their functional similarity as they both show the same high affinity for the agonist isoproterenol[37]. Hence, the nanobody behaves as a surrogate of the Gs protein.

Sequence alignment of hTAS2R31 to the template was followed by manual adjustment to take into account TM helices prediction by Adler *et al.*[2] based on multiple sequence alignments of bitter taste receptors. As a result, the most conserved residue in each helix was identified: $N24^{1.50}$, $R55^{2.50}$, $F105^{3.50}$, $G135^{4.50}$, $P187^{5.50}$, $Y241^{6.50}$ and $P276^{7.50}$ (sequence and Ballesteros-Weinstein numberings).

In **Figure 11** a docking solution of aristolochic acid bound to the homology model of hTAS2R31 is shown. The carboxylic acidic portion of this ligand makes charged-reinforced hydrogen bond interaction with R268 while the NO_2 moiety makes a hydrogen bond interaction with K265. This result is in agreement with alanine-scanning site directed mutagenesis studies reported in the literature[18].

Strychnine is not able to activate native hTAS2R31 up to 0.03 mM concentration[15]. However, when K265 and R268 were mutated to Glutamic acid and Alanine, respectively, strychnine was able to activate this mutant at 30 mM concentration[18]. In **Figure 12** strychnine bound to K265E R268A hTAS2R31 is shown. According to the model obtained, the removal of both these positively-charged residues and the introduction of a negatively-charged amino acid allows for the protonated nitrogen of the ligand to be nicely accommodated in this cavity. Besides, the hydrophobic core structure sits close to the hydrophobic residues F242 (TM6) and F188 (TM5).

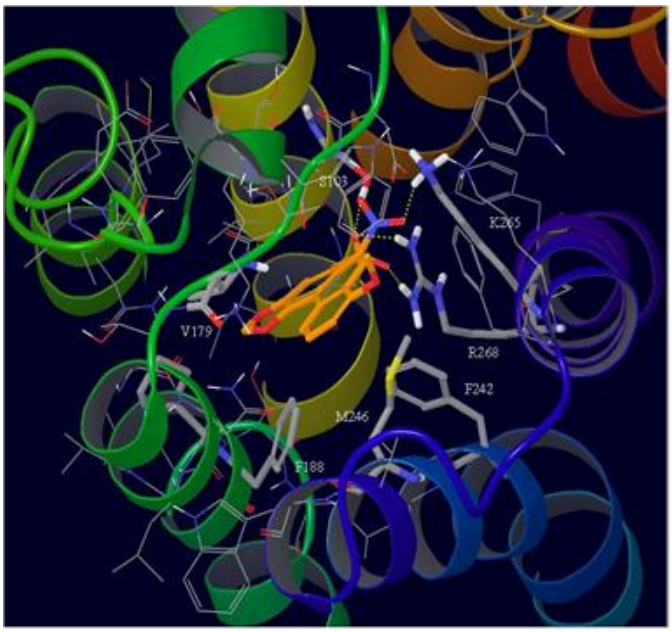


FIGURE 11. Aristolochic acid bound to hTAS2R31. The carboxylic acidic portion of this ligand makes a charged-reinforced hydrogen bond interaction with R268 while the NO_2 group makes a hydrogen bond interaction with K265.

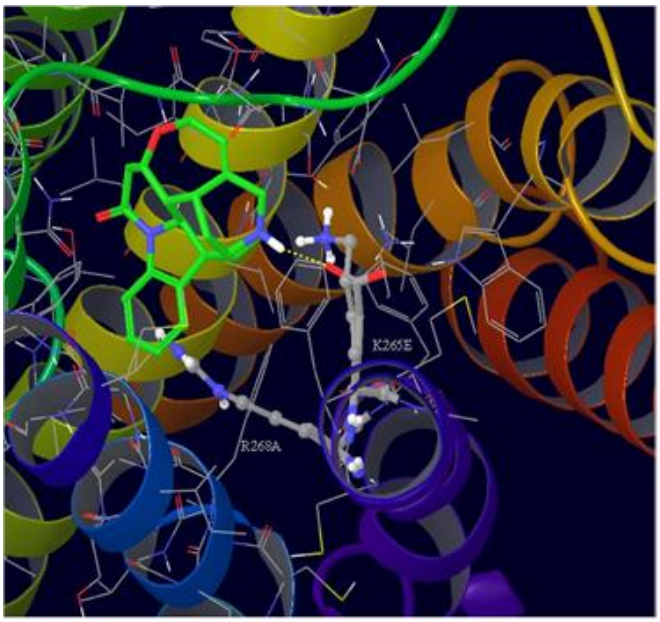


FIGURE 12. Strychnine bound to K265E and R268A hTAS2R31. The side chains of R268 and K265 of the native receptor are shown as ball&stick.

Denatonium benzoate is not able to activate native hTAS2R31 up to 3 mM concentration (Meyerhof *et al.*[15]). However, when R268 was mutated to Alanine, this agonist was able to activate this mutated receptor at 1 mM concentration[18]. When both R268 and K265 were mutated to Alanine and Glutamic acid, respectively, denatonium benzoate was still able to activate this mutant. In **Figure 13** a docking solution of denatonium bound to K265E hTAS2R31 is shown while in **Figure 14** is reported a pose of this ligand in the double mutated receptor.

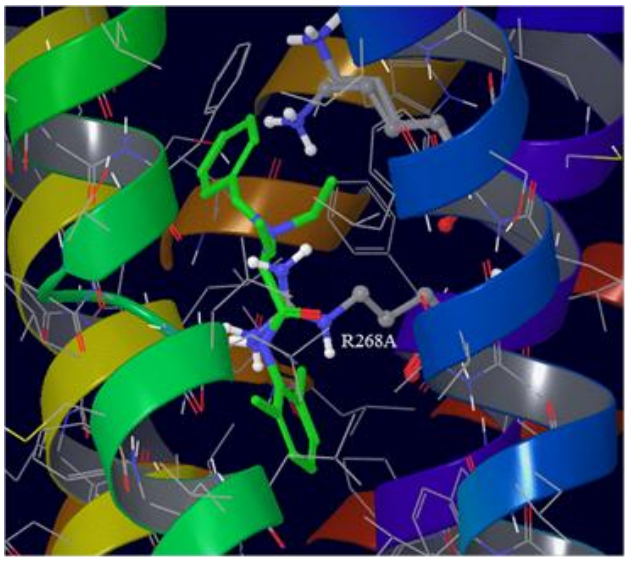


FIGURE 13. Denatonium benzoate bound to R268A hTAS2R31. The corresponding side chain of the wild type receptor is displayed as ball&stick.

According to the model obtained, the removal of the flexible linear chain of Lysine 265 creates space in the cavity allowing the bulky quaternary moiety of denatonium to be nicely accommodated. In the double mutant the positively charged nitrogen lies in the same spatial area as the guanidine moiety of the arginine.

Finally, aristolochic acid was docked in the hTAS2R31 receptor model by using a wider binding site (residues highlighted in cyan in **Figure 15**, left) in order to take into account the flexibility of the loops together with wide portions of all the helices except for TM1 which appears to be far away from the putative binding site.

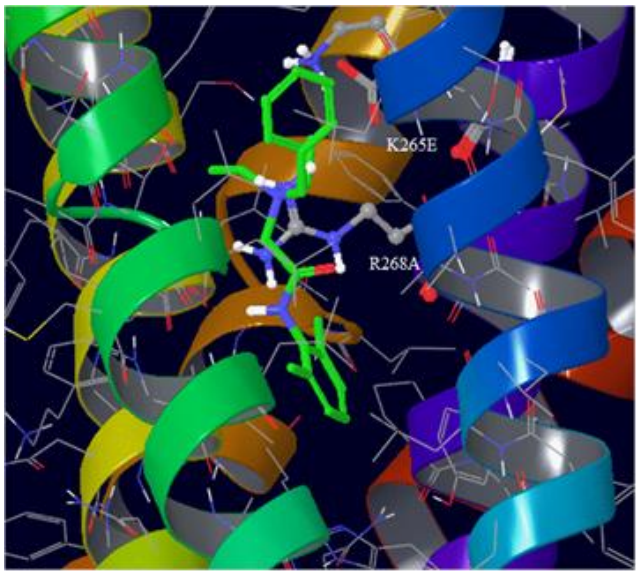


FIGURE 14. Denatonium benzoate bound to K265E R268A hTAS2R31. The side chain of K265 and R268 of the wild type receptor are displayed as ball&stick.

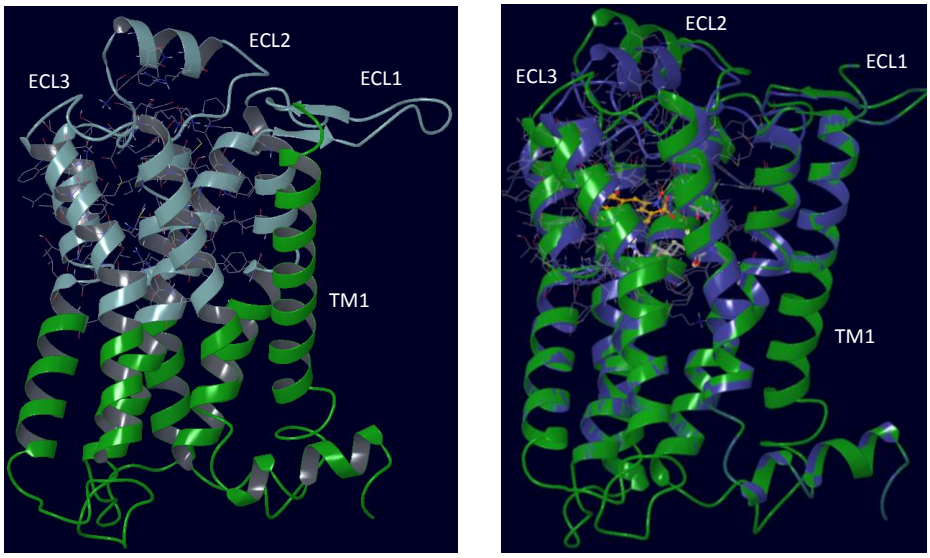


FIGURE 15. Left: Binding site used to dock aristolochic acid (residues highlighted in cyan). **Right:** Overlay of the most favorable IFD solution of aristolochic acid bound to hTAS2R31 (blue ribbon) on the starting structure of the receptor (green ribbon).

As shown, the use of a wide binding site comprising large portions of Helices 2-7 as well as all the extracellular loops allow for the ligand to flexibly explore a huge binding site. In all the 16 IFD solutions except 1 the ligand is close to K265 or R268 and makes hydrogen bond interactions with either one or both these residues, similarly to the poses sampled with a smaller binding site (**Figure 11**). Following this procedure, the extracellular loops, the ECL2 in particular, can adopt a more packed conformation.

Sequence alignment of hTAS2R10 and hTAS2R14 to the template was performed following the same procedure described for hTAS2R31. As a result, the most conserved residue in each helix of these receptors was identified:

hTAS2R10: $N24^{1.50}$, $R54^{2.50}$, $F105^{3.50}$, $F135^{4.50}$, $F185^{5.50}$, $Y239^{6.50}$ and $S274^{7.50}$
hTAS2R14: $N24^{1.50}$, $R55^{2.50}$, $F106^{3.50}$, $V136^{4.50}$, $P190^{5.50}$, $F243^{6.50}$ and $S277^{7.50}$

Then, the 3D models of hTAS2R10 and hTAS2R14 were built in the same way as that of hTAS2R31.

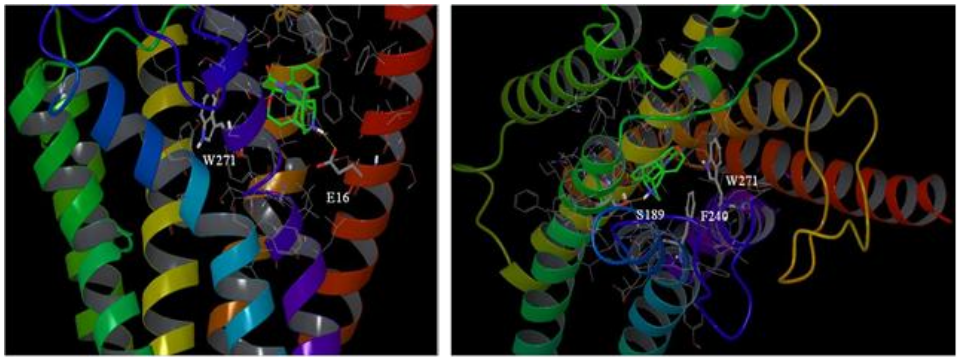


FIGURE 16. Potential binding modes of strychnine in hTAS2R10.

Docking studies of strychnine, the most selective agonist of hTAS2R10, suggest that this activator can either lie close to E16 (TM1) or make a hydrogen bond with S189 (TM5) (**Figure 16**). The latter binding mode projects this ligand close to W271 (TM7), which corresponds to K265 in hTAS2R31, and was also found by docking the ligand in the receptor devoid of its extracellular loops. Therefore, it is tentative to speculate that these both strychnine and aristolochic acid might share a common binding site at hTAS2R10 and hTAS2R31. In order to validate this model and the predicted ligand binding sites, site-directed mutagenesis experiments (e.g. alanine-scanning) should be performed for the amino acids in contact with this ligand. Recently, a homology model of

hTAS2R10, based on T4L-β_2AR-Gs complex (PDB code: 3SN6) as structural template, has been reported in the literature[38]. According to the putative strychnine binding modes proposed and the alanine-scanning SDM experiments performed[38], it seems that S85 in particular could be in contact with strychnine. In the homology model described by Born *et al.*[38] this residue lies in TM3 while in our model it is located in ECL1. It is worth highlighting that the strychnine EC_{50}s values reported in that paper for the wild type receptor, S85A and other mutants are not very robust due to high standard deviations (mutant S85A EC_{50} = 44.5 ± 38.8 μM *vs.* WT EC_{50} = 33.4 ± 23.9 μM). Hence, they appear to be inconclusive and more robust biological data will be needed in order to validate that model.

Docking studies of flufenamic acid in the model of hTAS2R14 show that the carboxylic group of this ligand can either bind R55 (TM2) or make a hydrogen bond interaction with T105 (TM3). The latter binding mode was also identified by performing docking experiments without the extracellular loops (**Figure 17**). Site-directed mutagenesis experiments should be undertaken to support these results.

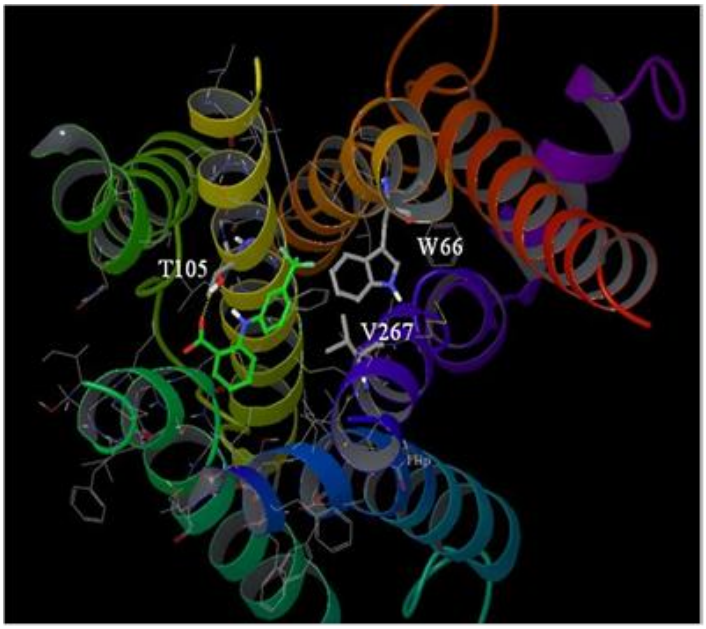


FIGURE 17. A docking solution of flufenamic acid bound to hTAS2R14 after the removal of the ECLs

Conclusions and future perspective

In recent years several publications have shown that bitter taste receptors are expressed in extragustatory regions such as airways smooth muscle cells and the gut where they could play a role in relaxation and secretion of hormones, respectively. Hence, TAS2Rs have emerged as attractive targets to initiate drug discovery projects aimed at identifying potent and selective activators to potentially treat respiratory and metabolic diseases. According to the studies performed in airways muscle cells and in an animal model of asthma, TAS2Rs activators clearly appear to play a key role as direct bronchodilators following a novel pathway deserving deeper investigations but clearly distinct from that of classical β-adrenergic receptor agonists. Despite the availability of β-adrenergic receptor agonists on the market to treat asthma, the β-agonist therapy has been associated with increased sensitivity to bronchospasm (bronchial hyper-reactivity), decreased acute responsiveness (tachyphylaxis), worsening asthma and death. Hence, there still the need to find bronchodilators for improving care on obstructive lung diseases following a novel mechanism of action.

Overall human bitter taste receptors exhibit low sequence identify among each other as well as with GPCRs whose X-ray structures have been solved. hTAS2R10 and hTAS2R14 show a certain degree of biological similarity as they share several common activators. However, cluster analysis has highlighted that their common tastants look structurally different, hence it is likely that they exhibit different binding modes at these receptors. Results of PCA highlight the fact that bitter tastants are characterized even by different physico-chemical properties with respect to one another. Furthermore, agonists of the same receptors look quite heterogeneous in terms of these properties.

Docking studies of representative bitter tastants in hTAS2T10, hTAS2R14 and hTAS2R31 subtypes have provided insights about their potential binding modes. While poses of aristolochic acid in complex with hTAS2R31 are consistent with site-directed mutagenesis studies reported in the literature, results of strychnine and flufenamic acid simulations in hTAS2R10 and hTAS2R14 models, respectively, need validation either by performing robust SDM experiments or indirectly by carrying out SAR expansion of these tool compounds. Data coming out from these studies would complement the information being generated on the function, genetics and cellular expression of bitter taste receptors.

As none of the bitter tastants considered in this study are potent agonists of hTAS2R10, hTAS2R14 and hTAS2R31 receptor subtypes, computational and synthetic efforts should be aimed at the identification of suitable pharmacological tool compounds for target validation experiments. This could be achieved by performing virtual screening exercises in parallel with focused chemical exploration around known activators.

Recently Liggett[39] has highlighted some issues to be addressed in order to develop TAS2Rs activators and treat asthma. First of all, it is not clear whether one or multiple bitter taste receptors have to be targeted by the activators to effectively treat this disease. Depending on receptor localization, it might be beneficial to target more than one subtype even though the appropriate selectivity profile has still to be defined. Besides, considering that bitter taste receptors are expressed in cells different from the airways smooth muscle cells of asthmatic patients (e.g. blood leukocytes) and in organs such as the tongue and the gut, the route of administration should be carefully evaluated. At this purpose, the use of inhaled delivery devices could reduce systemic exposure and hence give rise to selected receptor activations. Another potential issue is represented by bitter taste receptor desensitization upon repeated treatment with activators[40]. This effect has been observed in other GPCRs upon agonist administration and might be due to receptor phosphorylation by various kinases, receptor internalization and a loss of receptor expression. From a clinical perspective, desensitization has been observed as tolerance or tachyphylaxis, to an administered compound[41]. This observation raises a question as to whether it would be more appropriate to develop a partial rather than a full agonist. It would also be quite interesting to investigate the existence of endogenous ligands for TAS2Rs. If this could be proved, positive allosteric modulators approach might be utilized to selectively improve either the potency or the efficacy or both of the native substrate. This approach has already been successfully adopted for other GPCRs[42a]. In particular, Cinacalcet is a positive allosteric modulator of CaSR (family C GPCR) on the market, as calcimimetic to treat secondary hyperparathyroidism and hypercalcemia[42b]. Finally, positive allosteric modulators of the sweet receptor has been disclosed[43], demonstrating the validity of this approach also for taste receptors.

Abbreviations

TAS2R, bitter taste receptors; ASM, human airway smooth muscle cells; COPD, chronic obstructive pulmonary disease; PCA, Principal Component Analysis; PC1, Principal Component #1; PC2, Principal Component #2

References

(1) Lindermann B. Taste reception. *Physiol Rev.* **1996**, *76*, 718-66.

(2) Adler E., Hoon M.A, Mueller K.L., Chandrashekar J., Ryba N.J., Zuker C. S. A Novel Family of Mammalian Taste Receptors. *Cell* **2000**, *100*, 693-702

(3) Chandrasheker J., Mueller K.L., Hoon M.A., Adler E., Feng L., Guo W., Zuker C.S., Ryba N.J. T2Rs function as bitter taste receptors. *Cell* **2000**, *100*, 703-711

(4) Matsunami H., Montmayeur J.P., Buck L.B. A family of candidate taste receptors in human and mouse. *Nature* **2000**, *404*, 601-604.

(5) Shi P., Zhang J. Contrasting Modes of Evolution Between Vertebrate Sweet/Umami Receptor Genes and Bitter Receptor Genes. *Mol.Biol.Evol.* **2006**, *23*, 292-300.

(6) Kyle Palmer R. The pharmacology and signaling of bitter, sweet, and umami taste sensing. *Molecular Interventions* **2007**, *7*, 87-98.

(7) Lagerstroem M.C., Schiöth H. B. Structural diversity of G proteincoupled receptors and significance for drug discovery. *Nature Reviews Drug Discovery* **2008**, *7*, 339-358.

(8) Chandrasheker J., Hoon M.A., Ryba N.J.P., Zuker C.S. The receptors and cells for mammalian taste. *Nature* **2006**, *444*, 288-294.

(9) Deshpande D., Wang W. C. H., McIlmoyle E. L., Robinett K. S., Schillinger R. M., An S. S., Sham J. S. K., Liggett S. B. Bitter taste receptors on airway smooth muscle bronchodilate by a localized calcium flux and reverse obstruction. *Nat. Med.* **2010**, *16*, 1299-1304

(10) Pulkkinen V., Manson M.L., Säfholm J., Adner M., Dahlén S.-E. The bitter taste receptor (TAS2R) agonists denatonium and chloroquine display

distinct patterns of relaxation of the guinea pig trachea. *Am. J. Physiol. Lung Cell Mol. Physiol.* **2012**, *303*, L956–L966.

(11) Orsmark-Pietras C., James A., Konradsen J.R., Nordlund B., So¨derha¨ll C., Pulkkinen V., Pedroletti C., Daham K., Kupczyk M., Dahle´n B., Kere J., Dahle´n S.-E., Hedlin G., Mele´n E. Transcriptome analysis reveals upregulation of bitter taste receptors in severe asthmatics. *Eur. Respir. J.* **2013**, *42*, 65–78

(12) Zhang C.H., Chen C., Lifshitz L.M., Fogarty K.E., Zhu M.S., ZhuGe R. Activation of BK channels may not be required for bitter tastant-induced bronchodilation. *Nat. Med.* **2012**, *18*, 648–650.

(13) An S.S., Robinett K.S., Deshpande D.A., Wang W.C., Liggett S.B. Reply to: Activation of BK channels may not be required for bitter tastant-induced bronchodilation. *Nat. Med.* **2012**, *18*, 650–651.

(14) Zhang C.-H., Lifshitz L.M., Uy K. F., Ikebe M., Fogarty K.E., ZhuGe R. The Cellular and Molecular Basis of Bitter Tastant-Induced Bronchodilation. *PLOS One* **2013**, *11*, 1.

(15) Meyerhof W., Batram C., Kuhn C., Brockhoff A., Chudoba E., Bufe B., Appendino G., Behrens M. The Molecular Receptive Ranges of Human TAS2R Bitter Taste Receptors. *Chem.Senses* **2010**, *35*, 157-170

(16) Tizzano M., Gulbransen B.D., Vandenbeuch A., Clapp T.R., Herman J.P., Sibhatu H.M., Churchill M. E. A., Silver W.L., Kinnamon S.C., Finger T.E. Nasal chemosensory cells use bitter taste signaling to detect irritants and bacterial signals. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 3210-15.

(17) a) Janssen, S., Laermans, J., Verhulst, P. J., Thijs, T., Tack, J., Depoortere, I. Bitter taste receptors and α-gustducin regulate the secretion of ghrelin with functional effects on food intake and gastric emptying. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 2094−2099. b) Dotson, C. D., Zhang, L., Xu, H., Shin, Y. K., Vigues, S., Ott, S. H., Elson, A. E., Choi, H. J., Shaw, H., Egan, J. M., Mitchell, B. D., Li, X., Steinle, N. I., Munger, S. D. Bitter taste receptors influence glucose homeostasis. *PLoS One* **2008**, *3(12)*, e3974. c) Lund T.C., Kobs A.J., Kramer A., Nyquist M., Kuroki M.T., Osborn J., Lidke D.S., Low-Nam S.T., Blazar B.R., Tolar J. *PLOS One* **2013**, *8(3)*, e58945, 1-12.

(18) Brockhoff A., Behrens M., Niv M.Y., Meyerhof W. Structural requirements of bitter taste receptor activation. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 11110-11115.

(19) Altschul S.F., Madden T. L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389-3402.

(20) Bufe B., Hofman T., Krautwurst D., Raguse J.D., Meyerhof W. *Nat. Genet.* **2002**, *32*, 397-401.

(21) Kim U.K., Jorgenson E., Coon H., Leppert M., Risch N., Drayna D. Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science* **2003**, *299(5610)*, 1221-5.

(22) a) Behrens M., Brockhoff A., Kuhn C., Bufe B., Winnig M. and Meyerhof, W. The human taste receptor hTAS2R14 responds to a variety of different bitter compounds. *Biochem. Biophys. Res. Commun.* **2004**, *319*, 479–485. b) Behrens M., Reichling C., Batram C., Brockhoff A., Meyerhof W. Bitter Taste Receptors and Their Cells. *Annals of the New York Academy of Sciences* **2009**, *1170*, International Symposium on Olfaction and Taste pages 111–115.

(23) Kuhn C., Bufe B., Winnig M., Hofmann T., Frank O., Behrens M., Lewtschenko T., Slack J.P., Ward C.D., Meyerhof W. Bitter taste receptors for saccharin and acesulfame K. *Journal of Neuroscience.* **2004**, *24*, 10260-5.

(24) a) Pronin A.N., Tang H., Connor J., Keung W. Identification of ligands for two human bitter T2R receptors. *Chem Senses.* **2004**, *29*, 583-93. b) Pronin A.N., Xu H., Tang H., Zhang L., Li Q., Li X. Specific alleles of bitter receptor genes influence human sensitivity to the bitterness of aloin and saccharin. *Curr. Biol.* **2007**, *17*, 1403-8.

(25) Brockhoff A., Behrens M., Massarotti A., Appendino G., Meyerhof W. Broad Tuning of the Human Bitter Taste Receptor hTAS2R46 to Various Sesquiterpene Lactones, Clerodane and Labdane Diterpenoids, Strychnine, and Denatonium *J.Agric.Food.Chem.* **2007**, *55*, 6236-6243.

(26) Sainz E., Cavenagh M.M., Gutierrez J., Battey J.F., Northup J.K., Sullivan S.L. Functional characterization of human bitter taste receptors. *Biochem. J.* **2007**, *403*, 537–543.

(27) Maehashi K., Matano M., Wang H., Vo L. A., Yamamoto Y., Huang L. Bitter peptides activate hTAS2Rs, the human bitter receptors. *Biochem Biophys Res Commun.* **2008**, *365*, 851–855.

(28) ChemAxon v5.4. www.chemaxon.com

(29) Wiener A., Shudler M., Levit A., Niv M.Y. BitterDB: a database of bitter compounds. *Nucl. Acids Res.* 2012, 40(Database issue), D413-9. doi: 10.1093/nar/gkr755. Epub 2011 Sep 22.

(30) SIMCA-P^{+} v11. UMETRICS AB, Box 7960, S-907 Umeå, Sweden. www.umetrics.com

(31) Rasmussen S.G.F., Choi H.-J., Fung J.S., Pardon E., Casarosa P., Chae P.S., DeVree B.T., Rosenbaum D.M., Thian F.S., Kobilka T.S., Schnapp A., Konetzki I., Sunahara R.K., Gellman S.H., Pautsch A., Steyaert J., Weis W.I., Kobilka B.K. Structure of a nano body-stabilized active state of the β2 adrenoceptor. *Nature* **2011**, *469*, 175-181.

(32) MOE 2010.12 Chemical Computing Group 1010 Sherbrooke St. W, Suite 910 Montreal, Quebec, Canada H3A 2R7

(33) Schrodinger LLC. 120 West 45th Street 17th Floor, Tower 45 New York, NY 10036-4041 Suite 2010.

(34) Vilar S., Karpiak J., Berk B., Costanzi S. In silico analysis of the binding of agonists and blockers to the β2-adrenergic receptor. *J.Mol.Graph.Model.* **2011**, *29*, 809-817.

(35) Reichling C., Meyerhof W., Behrens M. Functions of bitter taste receptors depend on N-glycosilation. *J.Neurochem.* **2008**, *106*, 1138-48.

(36) Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E., Moore R.J., Smith R.D. J. *Proteome Res.* **2005**, *4*, 2070-2080.

(37) Rasmussen S.G.F., DeVree B.T., Zou Y., Kruse A.C., Chung K.Y., Kobilka T.S., Thian F.S., Chae P.S., Pardon E., Calinski D., Mathiesen J.M., Shah S.T.A., Lyons J.A., Caffrey M., Gellman S.H., Steyaert J.,Skiniotis G., Weis W.I., Sunahara R.K., Kobilka B.K. Structure of a nano body-stabilized active state of the β2 adrenoceptor. *Nature* **2011**, *477*, 549-557.

(38) Born S., Levit A., Niv M.Y., Meyerhof W., Behrens M. The Human Bitter Taste Receptor TAS2R10 Is Tailored to Accommodate Numerous Diverse Ligands. *J. Neuroscience* **2013**, *33(1)*, 201-213.

(39) Liggett S.B *Expert Opin. Ther. Targets* **2013**, *17(6)*, 721-731.

(40) Robinett K.S., Deshpande D.A., Malone M.M., Liggett S.B. Agonist-promoted homologous desensitization of human airway smooth muscle bitter taste receptors. *Am. J. Respir. Cell Mol. Biol.* **2011**, *45*, 1069-74.

(41) Liggett S.B., Lefkowitz R.J. Adrenergic receptor-coupled adenylyl cyclase systems: regulation of receptor function by phosphorylation, sequestration and downregulation. In: Sibley D, Houslay M, editors. Regulation of cellular signal transduction pathways by desensitization and amplification. John Wiley & Sons, London; 1993. p. 71-97.

(42) a) Conn P.J., Christopoulos A., Lindsley C.W. Allosteric modulators of GPCRs: A novel approach for the treatment of CNS disorders. *Nat. Rev.* **2009**, *8*,41–54. b) Brauner-Osborne H., Wellendorph P., Anders J.A. Structure, Pharmacology and Therapeutics Prospects of Family C G-Protein Coupled Receptors. *Curr. Drug Targets* **2007**, *8*, 169

(43) a) Servant G., Tachdjian C., Tang X.-Q., Werner S., Zhang F., Li X., Kamdar P., Petrovic G., Ditschun T., Java A., Brust P., Brune N., DuBois G.E., Zoller M., Karanewsky D.S.. Positive allosteric modulators of the human sweet taste receptor enhance sweet taste *Proc. Natl. Acad. Sci. U.S.A* **2010**, *107*, 4746. b) Shigemura R., Podagurski C., Kitisin B., Suparno M., Ward J., Lebien T., Wirtz K., Zeller J.R. Sweetener Compositions and Methods of Making Them. WO/2009/100333.

Appendix 1

Sequence codes of human bitter taste receptors retrieved from UNIProtKB/Swiss-prot

Accession	Entry name	Status	Protein names	Gene names
Q9NYW0	T2R10_HUMAN	reviewed	Taste receptor type 2 member 10 (T2R10) (Taste receptor family 6 member 2) (T6R2)	TAS2R10
P59535	T2R12_HUMAN	reviewed	Putative taste receptor type 2 member 12 (T2R12) (Putative taste receptor type 2 member 28) (T2R28)	TAS2R12 TAS2R28
Q9NYW9	T2R13_HUMAN	reviewed	Taste receptor type 2 member 13 (T2R13) (Taste receptor family 6 member 5) (T6R5)	TAS2R13
Q9NYW8	T2R14_HUMAN	reviewed	Taste receptor type 2 member 14 (T2R14) (Taste receptor family 6 member 1) (T6R1)	TAS2R14
Q9NYV7	T2R16_HUMAN	reviewed	Taste receptor type 2 member 16 (T2R16)	TAS2R16
P59542	T2R19_HUMAN	reviewed	Taste receptor type 2 member 19 (Taste receptor type 2 member 35) (Taste receptor type 2 member 46) (T2R46)	TAS2R19 TAS2R35 TAS2R46
P59543	T2R20_HUMAN	reviewed	Taste receptor type 2 member 20 (Taste receptor type 2 member 49) (T2R49) (Taste receptor type 2 member 56) (T2R56)	TAS2R20 TAS2R49
P59541	T2R30_HUMAN	reviewed	Taste receptor type 2 member 30 (Taste receptor type 2 member 47) (T2R47)	TAS2R30 TAS2R47
P59538	T2R31_HUMAN	reviewed	Taste receptor type 2 member 31 (T2R31) (Taste receptor type 2 member 44) (T2R44) (Taste receptor type 2 member 55) (T2R55)	TAS2R31 TAS2R44
P59539	T2R38_HUMAN	reviewed	Taste receptor type 2 member 38 (T2R38) (PTC bitter taste receptor) (Taste receptor type 2 member 61) (T2R61)	TAS2R38 PTC
P59534	T2R39_HUMAN	reviewed	Taste receptor type 2 member 39 (T2R39) (Taste receptor type 2 member 57) (T2R57)	TAS2R39
P59533	T2R40_HUMAN	reviewed	Taste receptor type 2 member 40 (T2R40) (G-protein coupled receptor 60) (Taste receptor type 2 member 58) (T2R58)	TAS2R40 GPR60
P59536	T2R41_HUMAN	reviewed	Taste receptor type 2 member 41 (T2R41) (Taste receptor type 2 member 59) (T2R59)	TAS2R41
Q7RTR8	T2R42_HUMAN	reviewed	Taste receptor type 2 member 42 (T2R42) (Taste receptor type 2 member 53) (T2R53)	TAS2R42 TAS2R53
P59537	T2R43_HUMAN	reviewed	Taste receptor type 2 member 43 (T2R43) (Taste receptor type 2 member 52) (T2R52)	TAS2R43
P59532	T2R45_HUMAN	reviewed	Taste receptor type 2 member 45 (T2R45) (G-protein coupled receptor 59)	TAS2R45 GPR59
P59540	T2R46_HUMAN	reviewed	Taste receptor type 2 member 46 (T2R46) (Taste receptor type 2 member 54) (T2R54)	TAS2R46
P59544	T2R50_HUMAN	reviewed	Taste receptor type 2 member 50 (T2R50) (Taste receptor type 2 member 51) (T2R51)	TAS2R50
P59545	T2R60_HUMAN	reviewed	Taste receptor type 2 member 60 (T2R60) (Taste receptor type 2 member 56) (T2R56)	TAS2R60
Q9NYW7	TA2R1_HUMAN	reviewed	Taste receptor type 2 member 1 (T2R1) (Taste receptor family 6 member 7) (T6R7)	TAS2R1
Q9NYW6	TA2R3_HUMAN	reviewed	Taste receptor type 2 member 3 (T2R3)	TAS2R3
Q9NYW5	TA2R4_HUMAN	reviewed	Taste receptor type 2 member 4 (T2R4)	TAS2R4
Q9NYW4	TA2R5_HUMAN	reviewed	Taste receptor type 2 member 5 (T2R5)	TAS2R5
Q9NYW3	TA2R7_HUMAN	reviewed	Taste receptor type 2 member 7 (T2R7) (Taste receptor family 6 member 4) (T6R4)	TAS2R7
Q9NYW2	TA2R8_HUMAN	reviewed	Taste receptor type 2 member 8 (T2R8) (Taste receptor family 6 member 3) (T6R3)	TAS2R8
Q9NYW1	TA2R9_HUMAN	reviewed	Taste receptor type 2 member 9 (T2R9) (Taste receptor family 6 member 6) (T6R6)	TAS2R9

Appendix 2

2.1 Descriptors calculated with ChemAxon v5.4

/home/software/chemaxon5.4/jchem/bin/cxcalc –S –t cx BitterDB_all_15062011_3_des_clean_std_unique.sdf mass logp avgpol molpol formalcharge atomcount aliphaticatomcount aromaticatomcount bondcount aliphaticbondcount aromaticbondcount ringcount heteroringcount heteroaliphaticringcount heteroaromaticringcount ringatomcount ringbondcount chainatomcount chainbondcount smallestringsize largestringsize aliphaticringcount aromaticringcount polarsurfacearea volume refractivity acceptorCount donorCount rotatableBondCount > BitterDB_all_15062011_3_des_clean_std_unique_cxprofile_2.sdf

LogD@pH7.4 was calculated within InstantJChem (IJC) graphical interface.

2.2 Evaluation of structural and pharmacophoric diversity of BitterDB compounds

2.2.1 Generation of structural fingerptints

generatemd c BitterDB_all_15062011_3_des_clean_std_unique.sdf -g -k CF -c ./cfp_not.xml -D -o BitterDB_all_15062011_3_des_clean_std_unique.CF

2.2.2 Comparison of the whole BitterDB set

/home/software/chemaxon5.4/jchem/bin/compr -f 1024 -g -z -i BitterDB_all_15062011_3_des_clean_std_unique.CF BitterDB_all_15062011_3_des_clean_std_unique.CF > compr_BitterDB_all_15062011_3_des_clean_std_unique.txt

2.2.3. Evaluation of structural (CF) and pharmacophoric (PF) diversity of each BitterDB compound with respect to each other

CF

/home/software/chemaxon5.4/jchem/bin/screenmd ./BitterDB_all_15062011_3_des_clean_std_unique.sdf ./BitterDB_all_15062011_3_des_clean_std_unique.sdf -g –k CF -M Tanimoto -c ./cfp_not.xml > screenmd_CF_BitterDB_all_15062011_3_des_clean_std_unique.txt

FP

/home/software/chemaxon5.4/jchem/bin/screenmd ./BitterDB_all_15062011_3_des_clean_std_unique.sdf ./BitterDB_all_15062011_3_des_clean_std_unique.sdf -g –k PF -M Tanimoto -c ./pharma-frag.xml > screenmd_PF_BitterDB_all_15062011_3_des_clean_std_unique.txt

2.2.4 Evaluation of structural (CF) and pharmacophoric (PF) diversity of hTAS2R10, hTAS2R14 and hTAS2R31 agonists

/home/software/chemaxon5.4/jchem/bin/screenmd ./BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R14.sdf ././BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R14.sdf -g -k CF -M Tanimoto -c ./cfp_not.xml > screenmd_CF_BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R14.txt

/home/software/chemaxon5.4/jchem/bin/screenmd ./BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R14.sdf ././BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R14.sdf -g -k PF -M Tanimoto -c ./pharma-frag.xml > screenmd_FP_BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R14.txt

/home/software/chemaxon5.4/jchem/bin/screenmd ./BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R10.sdf ././BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R10.sdf -g -k CF -M Tanimoto -c ./cfp_not.xml > screenmd_CF_BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R10.txt

/home/software/chemaxon5.4/jchem/bin/screenmd ./BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R10.sdf ././BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R10.sdf -g -k PF -M Tanimoto -c ./pharma-frag.xml > screenmd_PF_BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R10.txt

/home/software/chemaxon5.4/jchem/bin/screenmd ./BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R31.sdf ./BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R31.sdf -g -k CF -M Tanimoto -c ./cfp_not.xml > screenmd_CF_BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R31.txt

/home/software/chemaxon5.4/jchem/bin/screenmd ./BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R31.sdf ./BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R31.sdf -g -k PF -M Tanimoto -c ./pharma-frag.xml > screenmd_PF_BitterDB_all_15062011_3_des_clean_std_unique_hTAS2R31.txt

ELUCIDATION OF GLUCOCORTICOSTEROIDS UNBINDING PATHWAYS FROM THE GLUCOCORTICOID RECEPTOR

Introduction

Glucocorticoids receptors (GR) belong to the steroid-thyroid-retinoid family of the nuclear receptors (NRs), globular proteins acting as ligand-dependent transcription factors and controlling specific gene expression in most mammalian cells.[1,2] NRs share a common domain organization, and the most conserved regions are represented by the DNA-binding domains and the ligand binding domain (LBD).[1,2] Glucocorticoids (GCs) are endogenous steroid hormones that regulate a variety of cell, tissue-, and organ-specific biological functions such as metabolism, growth, development behavior, and apoptosis.[1,4–6] GCs represent the most effective anti-inflammatory agents to treat several chronic and acute inflammatory conditions including asthma, psoriasis, rheumatoid arthritis and irritable bowel disease.[5,7,8] Despite the great benefits of GCs severe side effects such as osteoporosis, diabetes, hypertension, and obesity can occur, thus, hampering the GCs clinical use.[8,9] Therefore, the development of novel modulators that can dissociate the therapeutic from the undesired adverse effects is impelling.[8,9] Hence, drug design strategies of new GR modulators require a deep understanding of GCs pharmacological profile and how these ligands act at molecular level.

The biological action of glucocorticoids is mediated by the glucocorticoid nuclear receptor (GR). Upon ligand binding, GR translocates into the nucleus and binds to DNA at glucocorticoid response elements (GRE) in the promoter region of steroid-responsive genes (transactivation).[5] GR may also negatively modulate (transrepression) inflammatory responses through inhibitory interactions with pro-inflammatory transcription factors such as Activator Protein (AP)-1 and Nuclear Factor (NF)-kB, without binding to DNA; in this case the GR controls gene expression by modulating the transcriptional activities of those factors.[8]

Dexamethasone (Dexa, **Figure 1**) is a prototypical glucocorticosteroid characterized by nanomolar affinity for GR and high oral bioavailability (30%).[6,10] A considerable amount of research has been done by medicinal chemists to improve the therapeutic index of glucocorticosteroids by increasing their potency and reducing their systemic side effects.[1,4–6,9,10] These efforts

culminated in the discovery of Fluticasone Furoate (FF, **Figure 1**) and Fluticasone Propionate (FP, **Figure 1**) for topic administration.[11] In particular, FF shows sub-nanomolar potency and longer duration of action with respect to Dexa and FP.[12] Furthermore, FF is characterized by slow dissociation rate (k_{off}) from GR, comparable to that of FP, but slower than that showed by Dexa.[10] k_{off} can be translated into dissociative half-life (dissociative $t_{1/2}$) for the receptor-ligand complex and this is a direct measure of the residence time in vitro and a possible indicator for (*i*) in vivo duration of action and (*ii*) non-target related toxicity due to sub-optimal selectivity profile.[13,14] Considering that either the lack of efficacy or off-target effects represent the main reasons for drug attrition,[13–16] the optimization of $t_{1/2}$ dissociative parameter might be taken into account in drug design strategies and during the lead optimization process.[13,14]

The disclosure of the X-ray crystal structures of Dexa and FF bound to human GR LBD in complex with the co-activator TIF2 (transcriptional intermediary factor 2)[6,11] provided, for the first time, important atomistic details about the binding modes of such drugs. Indeed, the crystal structures revealed that the furoate chain is nicely accommodated in the lipophilic 17α pocket of the receptor.[6,11] On the contrary the hydroxyl group at C17 of Dexa does not completely fill this hydrophobic cavity and this moiety is involved in a hydrogen bond interaction with Q642.[6,11] Although of foremost importance, these X-ray crystal structures represent static pictures of ligand-receptor complexes and do not provide apparent explanation for the different binding kinetic properties of Dexa, FF and FP and the mechanism underlying the unbinding process. Hence, the mechanism which regulates the k_{off} kinetic parameters of glucocorticoids for GR is still elusive. In this regard, computational tools have proven to be useful in deciphering the structural details underlying the different k_{off} values for a series of ligand analogues.[17,18] In particular Steered Molecular Dynamics (SMD) has demonstrated to be a flexible and powerful method to gather information about the energy landscape driving the ligand-receptor unbinding processes, and information about time-resolved complex dissociation,[17–22] also in the case of another NR. [23,24]

Herein, a detailed description of the unbinding process of Dexa, FF and FP from GR LBD is provided by using SMD simulations. Details of their unbinding pathways obtained are instrumental to propose an atomistic explanation for the high reduction in efficacy of Dexa in the naturally occurring I747M-GR, a mutated receptor implicated in rare familial glucocorticoid resistance, clinically characterized by generalized, partial, target tissue insensitivity to

glucocorticoids[25,26]. This syndrome is currently treated by pharmacologically reducing the endogenous secretion of ACTH and by adequately suppressing the HPA axis. Current treatments consist in the administration of high doses of Dexamethasone (1-3mg/d), which has to be carefully titrated according to the clinical manifestation and biochemical profiles of the patients.[27,28] Therefore, the possibility of relying on more potent and selective glucocorticoids to treat such a rare condition would be highly desirable in order to reduce the doses and minimize side-effects of the administered drug. The SMD protocol here described can be used to prioritize the synthesis of novel glucocorticosteroids, structurally-related to Dexa, on the basis of their calculated Potential of Mean Forces (PMFs) and unbinding energies.

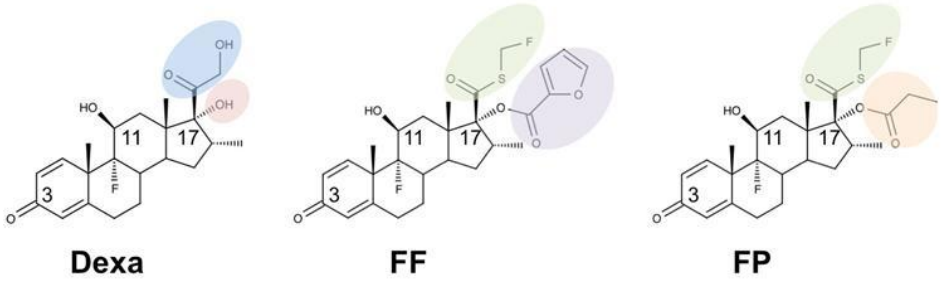


Figure 1. 2D representation of the GCs used in the study. For each compound the substituents on the C17 carbon have been highlighted with colored ovals.

Methods

(i.) Protein preparation

According to the accepted mechanism of ligand unbinding from NRs, the ligand is released before co-activator dissociation.[2] Therefore, ligand unbinding would take place when the co-activator is still bound. Thus, 3D coordinates of Dexa bound to the LBD of human GR and co-activator TIF2 were retrieved from the Protein Data Bank (PDB code: 1M2Z). Chain A (GR LBD, F602Y mutant) and chain B (TIF2) were used in the subsequent simulations (monomeric receptor complex). All the water molecules present in the X-ray structures were removed. Dexa was modeled starting from its crystallographic coordinates. Similarly, the GR LBD with bound FF was retrieved from the Protein Data Bank (PDB code: 3CLD). The chain A of GR-FF/TIF2 X-ray structure was

superimposed with that of GR-Dexa/TIF2 complex using Maestro v9.1.[34] Then the crystallized Dexa was replaced with FF. The GR-FF/TIF2 X-ray crystal structure was not used because, further to F602Y replacement, C638 near the putative escaping pathway had been mutated to glycine.[11] The impact of C638G mutation on the subsequent MD/SMD studies cannot be predicted *a priori*. Therefore, Dexa inside the binding site of 1M2Z Chain A was replaced with FF. This was possible since the superposition of GR-Dexa/TIF2 and GR-FF/TIF2 X-ray structures does not show significant differences in the organization of the tertiary structure of the protein, but only minor changes of the amino acids side chains conformations, which can be reduced through energy minimization of the generated complex. Finally, FP was manually overlaid on the new GR-FF/TIF-2 structures, and FF was replaced by FP. The mutated receptor I747M was built starting from the GR-Dexa/TIF2 complex and I747 was replaced by a methionine residue using the protein mutation site tool available in MOE.[35]

(ii.) Ligand preparation

The geometry of Dexa, FP and FF was optimized using the *Jaguar* tool available in Maestro v9.1.[34] Partial charges were computed using the HF/6-31G** level of theory and fixed using the RESP methodology. Atom types and ligand parameters were assigned using the general Amber force field (gaff).[36]

(iii.) Ligand-receptor complexes preparation

Each ligand-receptor construct was finally parameterized using the *amber99SB* and *gaff* force fields.[36] The complexes thus obtained were solvated in a cubic box of 12 Å using Leap[36] and TIP3P as water models. Na^+ and Cl^- ions were added to neutralize the net charge of the system and to mimic their physiological concentration (0.15M).

(iv.) MD simulations parameters and protocols

All MD simulations were performed with Amber v11.[36] All systems were gradually minimized and then equilibrated for 5 ns. During each equilibration phase the NPT ensemble was used with a pressure target equal to 1 atm and temperature of 310 K, pressure relaxation time of 1ps and collision frequency set to $5ps^{-1}$. The non-bonded interactions cutoff was set to 10 Å, and the PME algorithm was used with a grid spacing of 1 Å. We performed a two steps equilibration procedure. Each step lasted for 2.5 ns, with an integration step of 1 fs. The first 2.5 ns of the equilibration were conducted by imposing harmonic constraints equal to 2 kcal/mol*$Å^2$ to the protein backbone and the ligand

atoms. For the last 2.5 ns the same MD conditions were used and the constraints on protein backbone and ligand atoms were reduced to 1 kcal/mol*Å^2. Finally a production run was conducted for a total of 20 ns for each system. The first 10 ns were done using the NPT ensemble, while for the last 10 ns the NVT ensemble was used, so as to prepare the systems for the following SMD runs. During the production phase, the integration step was set equal to 2 fs and only bonds involving hydrogen atoms were constrained. The MD conditions were kept equal to those used for the equilibration phase, without imposing any harmonic constraints on the system. Along the last 10ns of NVT production phase (500 frames) frame #1, #125, #250, #375 and #500 were taken and used for constant-velocity SMD simulations.

(v.) SMD simulations and Potential of Mean Force Calculation (PMF, Φ)

All the SMD runs were carried out by using the NVT ensemble. The unbinding reaction coordinate was defined as the distance between the Cα atom of R611 and the ligand center of mass, which was gradually increased with respect to the initial distance, measured in the selected frames, during the SMD simulations. Hence, the agonists were not forced to follow a specific pathway (i.e. no predefined direction) but they have found ways to escape by constantly increasing the scalar value of the distance observed after 20 ns of the production runs. Five independent 20 ns long SMD trajectories (2000 frames each one) were generated for each system. All the SMD runs were conducted imposing a pulling velocity (*v*) of 0.001 Å/ps, the smallest ever used to model ligand unbinding from NRs[23,24] The value of the exerted force (*f*) was printed out every 5 ps (*dt*) and the work done on the system during SMD was calculated by numerical integration (eq.1). The stiff spring approximation (eq. 2)[19,21,22] was satisfied with a spring constant of 15 kcal/mol*Å^2. This value was identified by using the following step-wise optimization: (*i*) preliminary SMD simulations were conducted on the GR-Dexa/TIF2 (fast dissociating ligand) and GR-FF-TIF2 (slow dissociating ligand)[10] complexes by setting the spring constant value at 5, 10, 15, 20 and 25 kcal/mol*Å^2, respectively; (*ii*) selection of the spring constant values allowing for the conservation of the stiff-spring approximation for both the systems and definition of the spring values range (10-20 kcal/mol*Å^2) to be evenly explored (1 kcal/mol*Å^2 step) with further SMD runs with both systems; (*iii*) selection of the smallest spring constant value which allows for the conservation of the stiff-spring approximation in both systems (15 kcal/mol*Å^2). The GR-FP/TIF2 system was initially excluded from this analysis since it has a k_{off} profile very similar to that observed for the FF

ligand.[10] PMF along the unbinding reaction coordinate was reconstructed from the SMD trajectories described above and using Jarzynski's equality (eq. 3).[19,21,22] The cumulant expansion approach up to second order (eq. 4)[19,21,22] was used for reconstructing the PMF. Finally, in order to assess the convergence of our SMD simulations we performed a whole replica of the SMD runs for the GR-FF/TIF2 system, and the new PMF profile was compared with those obtained with the first 5 SMD runs.

$$W_{(x(t))} = \int_0^{x_{(t)}} F(t)dx(t) \quad (1)$$

$$F(\lambda) \cong \phi(\lambda) \quad (2)$$

$$e^{-\beta\Delta F} = \langle e^{-\beta W} \rangle \quad (3)$$

$$\Delta F = \langle W \rangle - \sigma^2 \frac{W}{2K_B T} \quad \text{where} \quad \sigma^2 W = \langle W^2 \rangle - \langle W \rangle^2 \quad \text{and} \quad K_B T = 0.6186 \quad (4)$$

(vi.) Description of the Unbinding Pathway

The dissociation path characterized by the lowest work profile according to eq. 1 was selected for each ligand-receptor among the 5 SMD runs. For each selected path, conformational clustering was carried out by using the *ptraj* tool of AmberTool11,[36] leading to the collection of a reduced number of conformational states. The snapshots of the selected SMD trajectories (one SMD run per GR-Dexa/TIF2, GR-FF/TIF2, GR-FP/TIF2, and GF(I747M)-Dexa/TIF2 systems, respectively) were clustered according to the RMSD of the ligand atoms and the Cα of the amino acids belonging to H11 - H12 loop with respect to the coordinates of frame #1. The H11-H12 loop was also included since it has been hypothesized to be important for GR activation process[3]. As a matter of fact, mutation of I747M lying in this loop decreases Dexa affinity by 2-fold with respect to the wild type receptor.[25,26] The average-linkage cluster analysis algorithm was used and the critical distance was set to 1 Å. The cluster tree was cut at 20 clusters, so as to cover the whole dissociation path for the selected SMD simulations and the representative structure, defined as the SMD frame closest to the each cluster centroid, was utilized to gather structural information. It is worth mentioning that the selected path and its associated representative structures do not give any estimation of the probability to find them, but provide useful information at atomic resolution of the key interactions

underlying the dissociation path characterized by the lowest work profile under our simulation conditions.

Results

(i) PMF along the Unbinding Reaction Coordinates

In **Figure 2a** is reported the PMF calculated by applying the Jarzynski's equality (eq. 3) and the cumulative expansion up to the second order (eq. 4). The end-to-end distance *vs.* the constrained center shown in **Figure 2b** proves that the stiff-spring approximation is satisfied under our simulation conditions by using a very small pulling velocity v = 0.001 Å/ps and a force constant of 15 kcal/mol*Å^2. Analysis of **Figure 2a** reveals that the PMF of FF (88.36 ± 6.33 kcal/mol) is higher than that of Dexa (63.87 ± 10.23 kcal/mol), but it is in the twilight zone of statistical significance with respect to that of FP (78.01 ± 3.65 kcal/mol) according to their standard deviations values (**Figure 2a**, orange and grey shadow regions, respectively). This result is qualitatively in agreement with the experimental values of their dissociation unbinding constants reported in the Inset table.

Furthermore, the PMF of Dexa shows higher variability with respect to those of the other derivatives (green shadow area) and its energy basin is wider that those of FF and FP, suggesting that Dexa can follow different unbinding trajectories, as reported for other NRs [29,30]. At the same time the dissipated work W_{diss} values of all these ligands representing the irreversible work due to non-equilibrium conditions, are comparable. Finally, Δw_{diss} of FF (5.10 kcal/mol), Dexa (2.87 kcal/mol) and FP (2.87 kcal/mol) are lower than their corresponding ΔΦs, (FF-Dexa = 24.48 kcal/mol; FF-FP =10.34 kcal/mol; FP-Dexa = 14.14 kcal/mol) supporting the robustness of these simulations Finally, these results are also in agreement with previously reported study for related Estrogen Receptors[23,24].

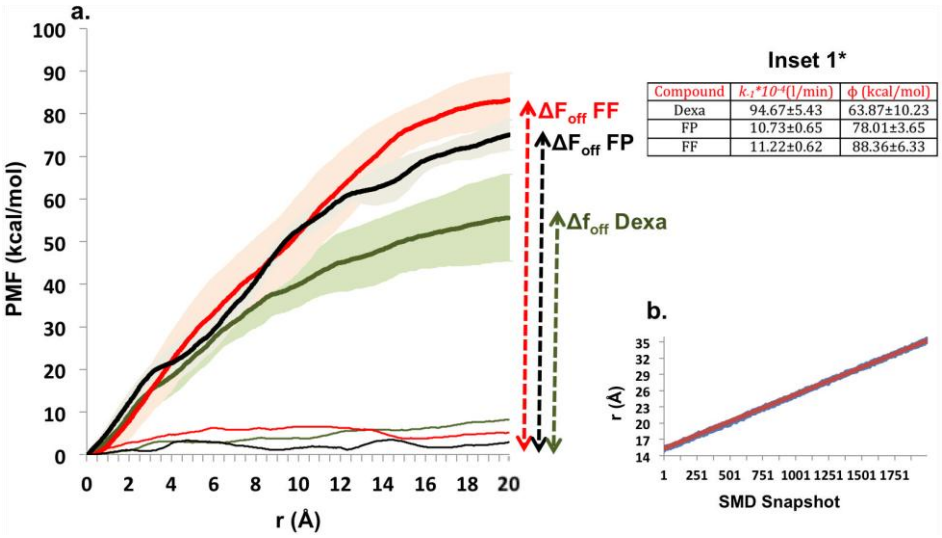


Figure 2. a) PMF profile for Dexa (green), FF (red), and FP (black), with the corresponding standard deviations (shadow regions), and the dissipated works (thin line); **b)** Typical end-to-end distance vs. constrained center plot obtained during the setting of our SMD protocol; **Inset 1** reports the k_{off} values with their standard deviations obtained from ref. [10], and the PMF profile obtained for each ligand (Φ kcal/mol).

(ii.) Analysis of the dissociating pathways

1. Comparison of the dissociating pathway of Dexa, FP and FF

The dissociation path exhibiting the lowest work profile was selected for each complex in order to rationalize the different biological behaviors of Dexa, FF and FP at atomistic level. The conformations sampled along the selected paths were clustered and representative structures used to reconstruct the representative dissociation path of each GR-ligand/TIF2 complex. Then, the structures showing crucial dissociation steps were projected on the force profiles of the selected SMD runs. As shown in **Figure 3a**, Dexa escapes following two putative pathways consistent with its own larger PMF energy basin. During Dexa escaping routes, there is a significant change in the conformation of the loop connecting H11 to H12 (**Figure 3b**). Differently from Dexa, FP and FF leave the binding site by following a unique putative path, which explains their more narrow PMF basins. The observed different behavior

of these drugs can be related to the presence of a bulky substituent on the C17 carbon (furoate ester for FF, ethyl ester for FP, **Figure 3a-c** and **4**, respectively), which can behave as an anchor reducing the ligand freedom during the unbinding process. On the contrary, Dexa lacks this bulky fragment and can freely explore different unbinding pathways.

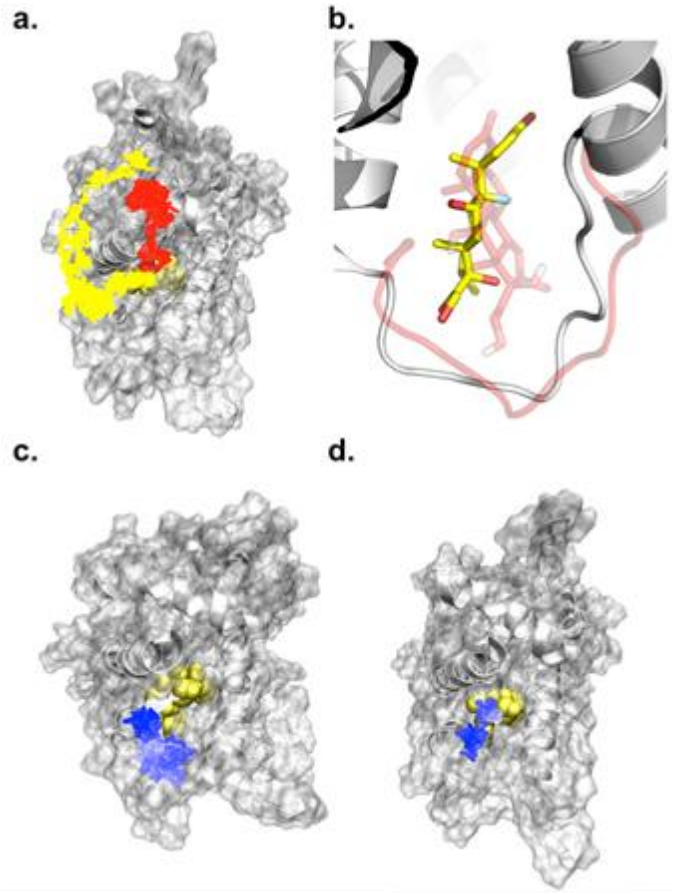


Figure 3. a) Dissociating paths for Dexa, the yellow and red line highlights the different paths observed in our simulations; **b)** Comparison of the orientation of Dexa in the different paths reported in **a.** Dexa in yellow and stick corresponds to the yellow path, while Dexa in red and transparent stick corresponds to the red path. The frame correspond to the same SMD simulation time, and it can be appreciated the remarkable conformational change, which occurs at the H11-H12 connecting loop; **c)** Dissociating paths for FF; **d)** Dissociating paths for FP. It is worth noticing that **c)** and **d)** are quite similar.

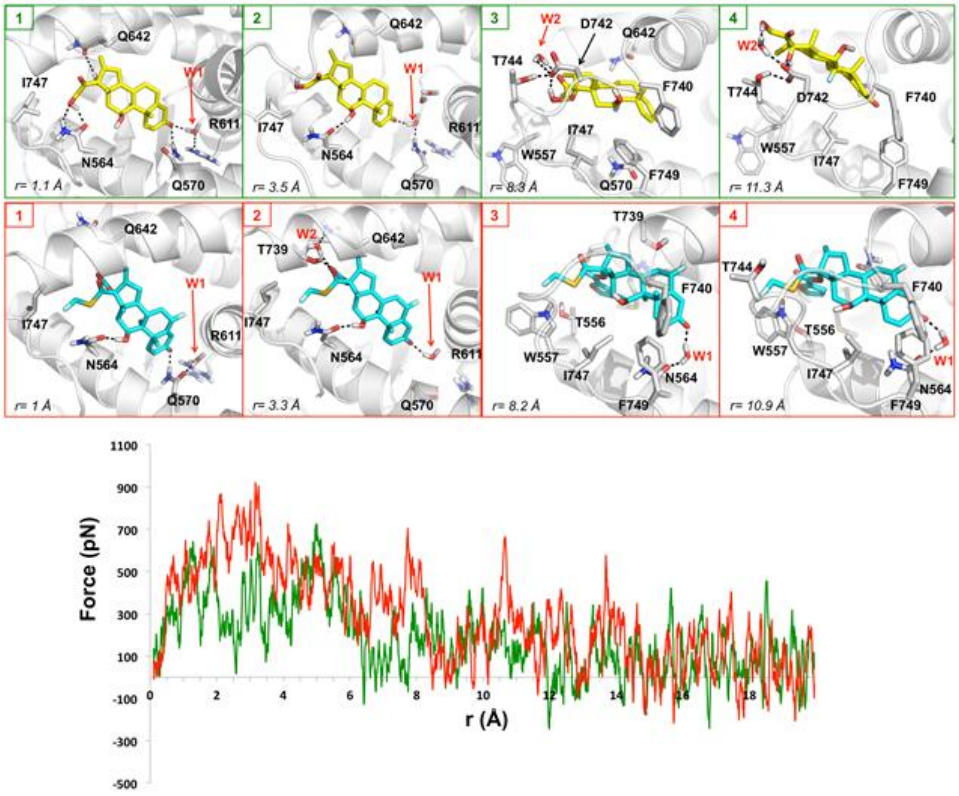


Figure 4. Representative structures for Dexa (green boxes 1,2,3,4) and FF (red boxes 1,2,3,4). The reported structures represent crucial step (with similar r) selected along the representative path obtained as described in Materials and Methods section. The force profiles plot for Dexa (green line) and FF (red line) are reported below the representative structures.

In **Figure 4** are reported representative structures describing these crucial steps in the dissociation pathways for Dexa and FF, respectively, together with their force profiles. Overall these snapshots highlight the crucial role played by water molecules (W1, W2) in assisting the breakage of hydrogen bond interactions of the ligands with the amino acids binding sites. Besides, the steric interactions established by the furoate chain with H11-H12 loop residues trap FF in the GR binding site delaying its way-out. Structures #1 show that both Dexa and FF interact with N564 and Q570. Besides, the OH moiety at C17 of Dexa makes a

hydrogen bond interaction with Q642. FF instead cannot make such an interaction and its furoate fragment changes the conformation of Q642 side chain, similar to that observed in the X-ray structure of GR-FF/TIF2 (PDB code 3CLD). However, these differences in ligand receptor interactions do not affect their unbinding pathways as demonstrated by the similarity of their force profiles between 0 and 2 Å. Besides a water molecule (W1) was found in the binding site of both Structures #1. While in the Dexa complex W1 bridges the carbonyl group at C3 with R611, in the FF one W1 is not always involved in hydrogen bond interactions and fluctuates nearby R611. This is in line with the fact that in the crystal structure of GR-FF/TIF2 two water molecules are close to R661 and Q570. Between 2 and 4 Å, the interactions with R611 are lost. In the representative Structure #2 of Dexa, the H11-H12 connecting loop moves closer to the ligand while in the corresponding snapshot of FF the furoate fragment is involved in a network of hydrogen bonds with T739, Q642 and another water molecule (W2) and is close to H11-H12 loop. As a consequence, Dexa is free to move towards the solvent while FF is trapped inside the GR binding site and its unbinding pathway is delayed. In the representative Structure #3 of Dexa, there is a change in the conformation of H11-H12 loop. In particular, D742 interacts with the OH group of the acetyl alcohol chain of C20 and a water molecule bridges the carboxyl moiety of D742 with the carbonyl group of the acetyl alcohol chain. Besides, the hydrophobic interactions established by F740 with F749 are lost and Dexa rotates so as to expose the acetyl alcohol chain to the H11-H12 loop residues. On the contrary, FF does not interact either with D742 or T744 and the furoate chain moves along a hydrophobic channel made by the methyl group of T556, the indole ring of W557 and the side chains of C643 and M639. Furthermore, the carbonyl moiety at C3 of the ligand is involved in a water-mediated hydrogen bond interaction with N564 side chain. In addition, F740 makes hydrophobic interactions with F749, both adopting conformations similar to those shown in the X-ray structure of GR-Dexa/TIF2 and GR-FF/TIF2. Furthermore, in the Dexa simulations, W557 points outward the hydrophobic pocket of the LBD, while in the FF ones it moves towards the ligand so as to build up the hydrophobic channel passed through by the furoate chain. These observations are also consistent with the force profiles of their unbinding trajectories between 7 and 9 Å and support the hypothesis that along the unbinding path the side chains of several hydrophobic amino acids act as gatekeepers, affecting the dissociation of different ligands from the GR LBD. Once Dexa is almost out of the hydrophobic pocket of the LBD (**Figure 4**, representative Structure #4 green box) the side chains conformations of F740 and F749 are comparable with those observed in the native X-ray complexes

(Phe side chains RMSD of GR-Dexa/TIF2 *vs.* Dexa X-ray structure 1M2Z=2.561 Å; RMSD of GR-FF/TIF2 *vs.* Dexa X-ray structure 1M2Z=1.704 Å) and in the GR-FF/TIF2 simulations (Phe side chains RMSD GR-Dexa/TIF2 *vs.* GR-FF/TIF2=0.857 Å), suggesting that the SMD parameters used to perform these simulations, in particular the very slow pulling velocity, allow for physiological re-adaptation of the protein during the dissociation of the ligands. At the same time Dexa rotates of ~90° and interacts with I747. In the corresponding FF Structure #4, the ligand is partially trapped inside the hydrophobic pocket of the LBD and the water molecule W1 bridges the ligand to N564. The furoate chain is involved in hydrophobic contacts with the side chains of T556, W557, C643 and M639 while W557 rotates so as to completely define the hydrophobic channel around the furoate chain. The force profiles between 10 and 11 Å highlights these remarkable differences observed for the unbinding pathways of FF and Dexa.

2. Comparison of the dissociating pathway of FF and FP

PMFs for FF and FP (**Figure 2a**) are very similar and qualitatively in agreement with their experimental k_{off} values reported in the literature[10]. Nevertheless, interesting differences emerged from the analysis of their force profiles and associated representative structures (**Figure 5**). Between 0 and 2 Å, the force profile of FP is similar to that of FP, which reflects the similarity of their interactions, mainly related to the presence of hydrophobic fragments at C17 (**Figure 5**, reference Structures #1). In the next steps of the dissociation path, between 2 and 6 Å, FP shows a behavior similar to that observed for Dexa (**Figure 5**, Structure #2, black line, and **Figure 4**, Structure #2, green line), which is likely due to the greater flexibility of the propionate chain of FP with respect to the furoate one of FF, facilitating the interactions with the residues of the LBD.

In the second part of the force profile plot (8 – 11 Å of reaction coordinate displacement) the propionate chain being smaller than the furoate one establishes less interactions with the binding site residues. In representative Structures #3 (**Figure 5**, black box) FP dissociates following a path closer to H3 rather than H7, resulting in a change in the conformation of H11-H12 loop and of residues in helix H3 such as T556 and W557. The latter residue adopts a conformation quite different from that of Dexa and FF (**Figures 4** and **5**,

representative Structures #4), which might explain the higher force profile of FP with respect to that of FF (**Figure 5**, 8 – 11 Å reaction coordinate).

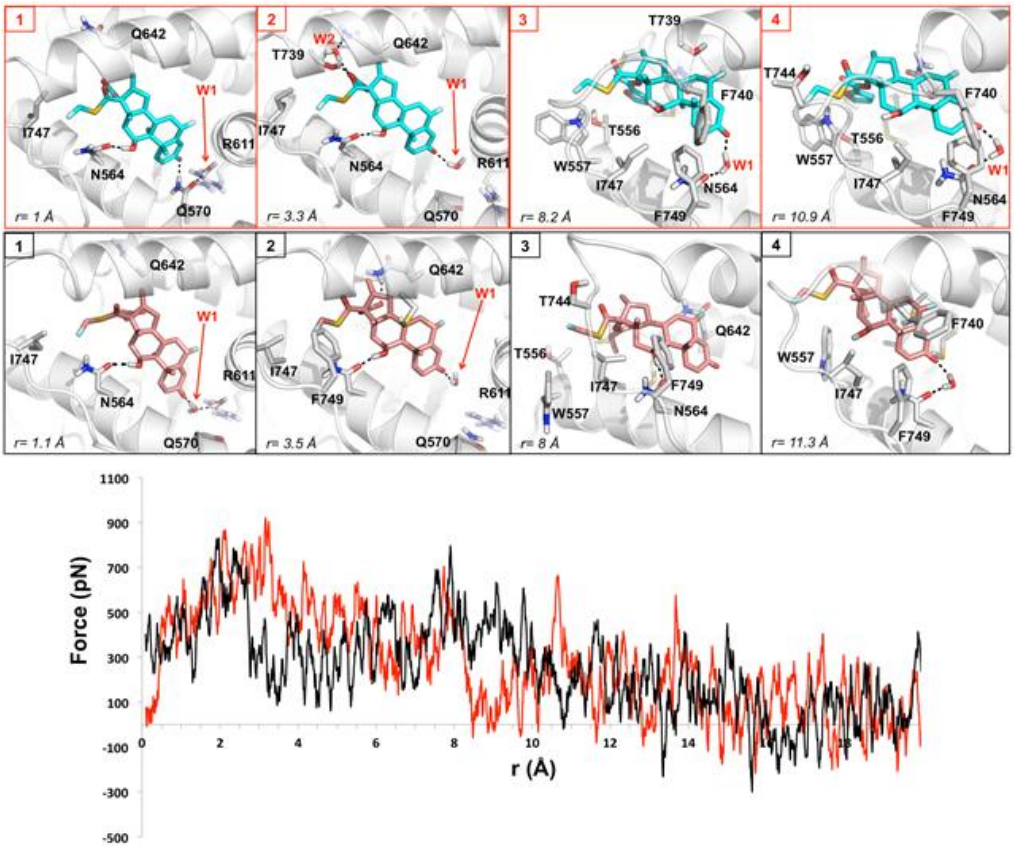


Figure 5. Representative structures for FF (red boxes 1,2,3,4) and FP (black boxes 1,2,3,4). The reported structures represent crucial step (with similar r) selected along the representative path obtained as described in Materials and Methods section. The force profiles plot for FF (red line) and FP (black line) are reported below the representative structures.

When the reaction coordinate lies between 10 and 11 Å, the force profiles of FP and FF are dramatically different. In the representative Structure #4 of FP simulations, W557 changes its conformation and the hydrophobic channel along which the C17 chain is moving becomes narrower. In the FP dissociation

process the energetic cost due to the conformational change of W557 is not completely compensated by the hydrophobic interactions of the propionate fragment with W557 and T556 side chains. On the contrary, in the FF simulations the bulkier furoate chain establishes a greater number of hydrophobic interactions with residues listed above, which compensates for the energy cost derived from W557 conformational change. In summary, despite FP and FF k_{off} values are comparable, these ligands show significant differences in their unbinding pathways, likely due to the fact that FP (*i*) does not completely fill the 17α hydrophobic pocket, and (*ii*) is not able to establish the same hydrophobic interactions as FF with T556 and W557. As a result, the unbinding process of FP requires greater exerted forces than that of FF to dissociate from the same region of the LBD (**Figure 5**, 8 – 11 Å reaction coordinate and corresponding snapshots in red and black lines) and it can be speculated that W557 is involved in the unbinding process of these ligands. In order to evaluate that, the relative frequency distribution of X1 and X2 dihedral angles of W557 was analyzed in all SMD runs for GR-Dexa/TIF2, GR-FP/TIF2 and GR-FF/TIF2. As shown in **Figure 6**, in all GR-Dexa/TIF2 SMD simulations W557 adopts a unique conformation (**Figure 6a**, Insert A). On the other hand, in GR-FP/TIF2 and FR-FF/TIF2 SMD simulations W557 shows great conformational variability. The most populated conformational state (**Figure 6b**, Insert A) is identical to that found in GR-Dexa/TIF2 SMD simulations. Besides, the second most populated state (**Figure 6b**, Insert C) differs from the previous one for X1 values. In addition, two less populated states were discovered. The former shown in **Figure 6b**, (Insert B) is similar to that observed in GR-FF/TIF2 crystal structure. The latter reported in **Figure 6b** (Insert D) differs from the previous ones both for X1 and X2.

In **Figure 7** are reported representative Structures #3 of Dexa (**Figure 7a**), representative Structures #3 and #4 of FF (**Figure 7b,c**) and representative Structures #4 of FP (**Figure 7d**). These pictures highlight the different behavior of W557 as a result of the change in the size of the substituent on C17. Furthermore, these pictures emphasize the plasticity of GR LBD in changing the shape of the hydrophobic floor according to the nature of the ligand substituent on C17, reinforcing the observation that W557 plays a crucial role in the dissociation of these ligands from the LBD.

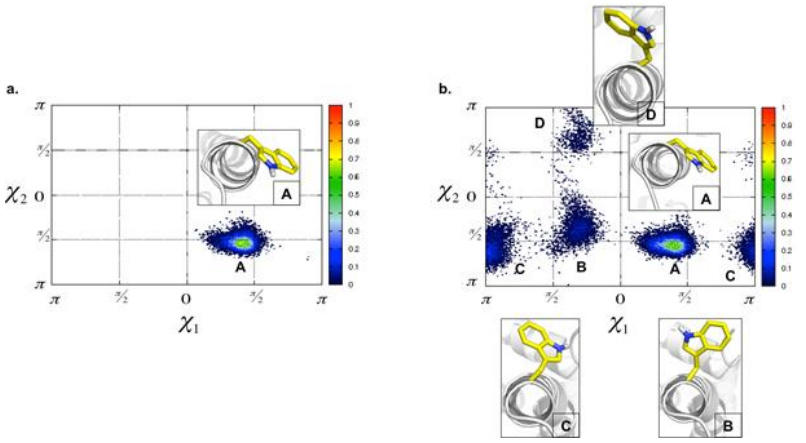


Figure 6. (a) Relative occurrence of W557 conformational states defined by X1 and X2 dihedral angles for GR-Dexa/TIF2, GR-FP/TIF2 and GR-FF/TIF2 SMD simulations. (b) Relative occurrence of W557 conformations defined by X1 and X2 dihedral angles for GR-Dexa/TIF2, GR-FP/TIF2 and GR-FF/TIF2 SMD simulations (color coding according to the normalized occurrence of W557 conformational states).

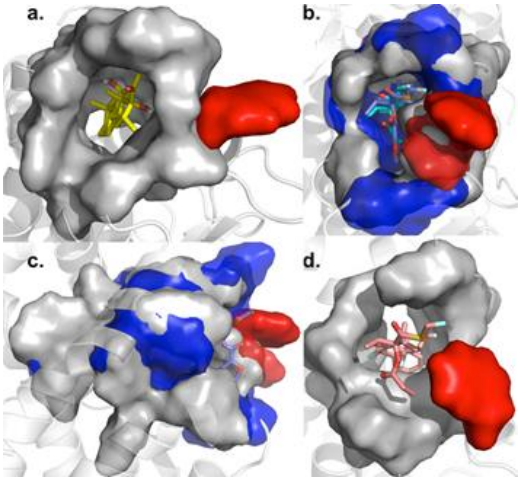


Figure 7. Representative structure #3 of Dexa (a), of FF structure #3 (gray and red surfaces) and FF structure #4 (blue and red surfaces), side view of panel b for representative structure #3 (gray and red) and #4 (blue and red surfaces) of FF (c) and Representative structure #4 of FP (4). These pictures highlight the conformational changes of W557 during SMD simulations.

3. *PMF along the Unbinding Reaction Coordinates for the GR(I747M)-Dexa/TIF2 system*

In the literature it has been reported that inactivating mutations within the LBD of GR lead to GCs resistance condition, which results in a generalized resistance to normal glucocorticoid levels.[25,26] I747M causes familial, autosomal dominant glucocorticoid resistance by decreasing ligand binding affinity and transcriptional activity.[25,26] In light of that, the effects of the I747M GR mutant variant[25,26] on the dissociation mechanism of Dexa were investigated. In **Figure 8a** is reported the PMFs for GR-Dexa/TIF2 and GR(I747M)-Dexa/TIF2 systems. The analysis of **Figure 8a** reveals that the calculated PMF for GR(I747M)-Dexa/TIF2 (49.800 ± 5.33 kcal/mol) is lower than that obtained for the GR-Dexa/TIF2 (63.880 ± 10.23 kcal/mol). The observed differences in the PMF profiles (**Figure 8a**) are statistically significant for most of the unbinding process. Nevertheless, in the last portion of the plot the computed standard deviations of both profiles are overlapping. As previously discussed the PMF for the GR-Dexa/TIF2 complex is affected by the highest standard deviation (SD=10.23) (**Figure 8a** green shadow region and Inset 2), since Dexa can dissociate following slightly different paths. It is worth highlighting that the standard deviation of GR(I747M)-Dexa/TIF2 PMF profile (SD=5.33) (**Figure 8a** purple shadow region and Inset 2) is drastically reduced with respect to that observed for the GR-Dexa/TIF2 system (**Figure 8a** green shadow region and Inset 2). This can be due to the fact that I747M mutation leads to the selection of a specific dissociation path, which differs from those observed for the GR-Dexa/TIF2 system. In particular, the different volume mapped by M747 side chain induces the rearrangement of the H11-H12 loop and the subsequent relocation of the adjacent Dexa (**Figure 8b**), which is closer to the hydrophobic floor of the LBD with respect to the WT complex (**Figure 8b**). The dissipated work values (ΔW_{diss}) of GR-Dexa/TIF2 (8.25 kcal/mol) and GR(I747M)-Dexa/TIF22 (3.290 kcal/mol) simulations are lower than the corresponding differences for the calculated PMFs (14.08 kcal/mol), supporting the reliability of the differences observed for their calculated PMFs in our SMD calculations.

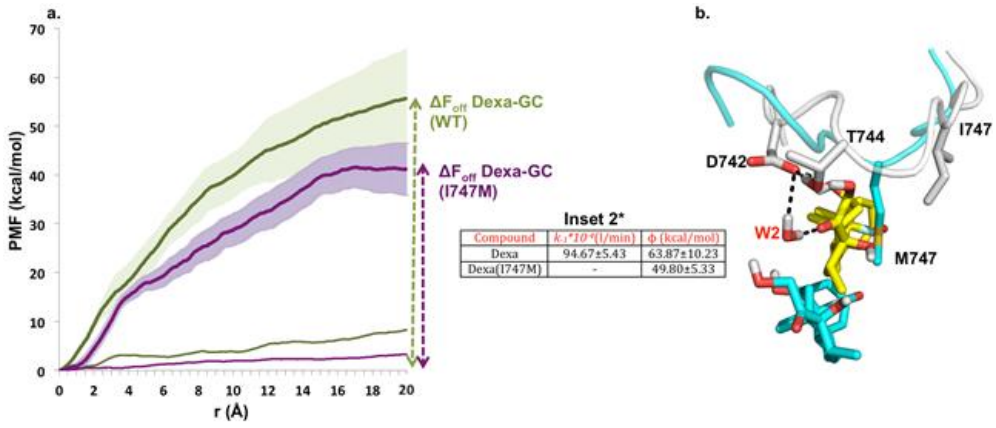


Figure 8. a. PMF profile of Dexa (green) in complex with GR-WT, and Dexa in complex with GR-I747M mutant (purple) with the corresponding standard deviations (shadow regions), and the dissipated works (thin line); **b.** Comparison of the representative structures 3 of both systems, which highlights the remarkable effect on the conformation of the H11-H12 connecting loop and on the location of Dexa. **Inset 1** reports the k_{off} values with their standard deviations obtained from ref.[10], and the PMF profile obtained for each ligand (Φ kcal/mol).

4. Analysis of the dissociating pathways for the GR(I747M)-Dexa/TIF2 system

In **Figure 9** the structures representing crucial steps along the unbinding reaction coordinate are reported, together with the force profiles of the SMD run characterized by the lowest work profile.

In the representative Structures #2 and #3 of GR(I747M)-Dexa/TIF2 there is a dramatic change in the conformation of H11-H12 connecting loop with respect to the corresponding frames of GR-Dexa/TIF2. This is most likely due to the replacement of isoleucine with methionine (**Figure 8b** and **Figure 9**, Structure #3 purple box) and is reflected in the force profile plot between 0-4 Å reaction coordinate. In the GR(I747M) mutant, Dexa follows a path similar to that observed for FF and FP, but lacking a bulky substituent on C17 it rotates similarly to Dexa in the WT LBD, without establishing any interactions with the loop residues. Between 4 and 6 Å of the reaction coordinate there is a marked decrease of the force values since Dexa is rapidly driven towards the subsequent dissociation step following a different dissociation path.

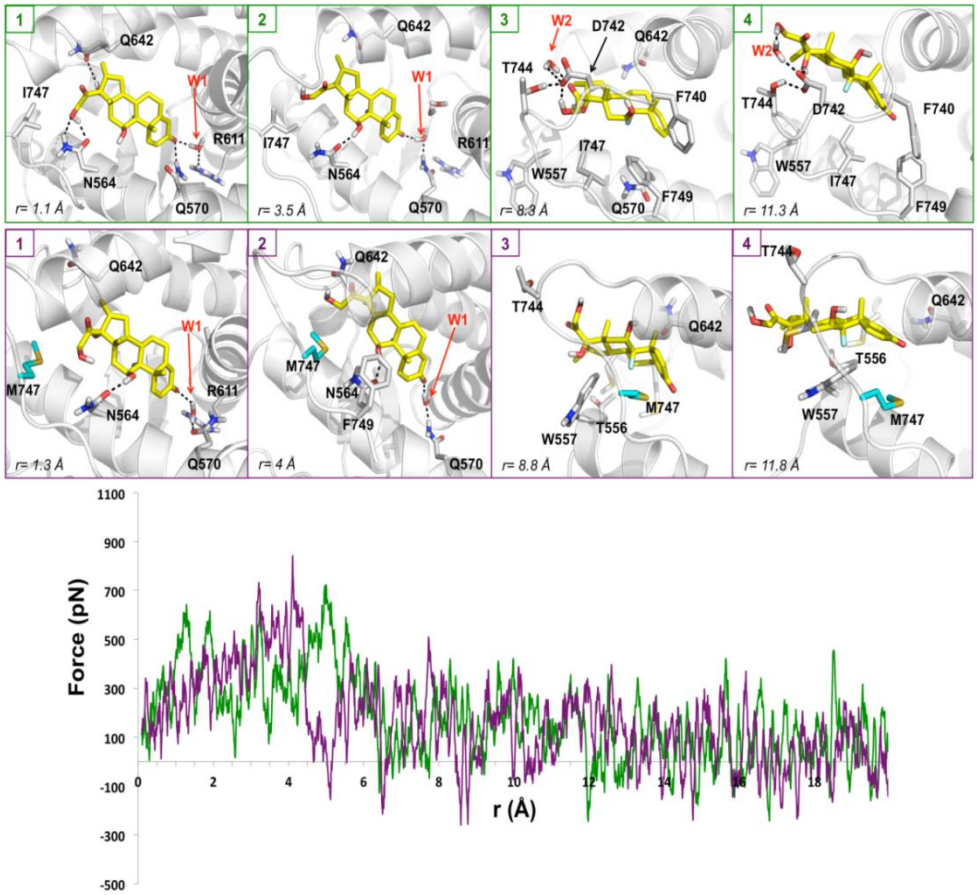


Figure 9. Representative structures for Dexa in complex with GR-WT (green boxes 1,2,3,4) and Dexa in complex with GR-I747M mutant (purple boxes 1,2,3,4). The reported structures represent crucial step (with similar r) selected along the representative path obtained as described in Materials and Methods section. The force profiles plot for Dexa in complex with GR-WT (green line) and Dexa in complex with GR-I747M mutant (purple line) are reported below the representative structures.

On the contrary, in the GR-Dexa/TIF2 simulations, the exerted forces in the same distance range are still used to establish interactions of the loop with the polar tail of Dexa. In the subsequent steps in the GR(I747M)-Dexa/TIF2 simulations (**Figure 9** structures #3 and #4 purple boxes) Dexa is close to the hydrophobic floor of the LBD, and is not able to interact with D742 and T744 residue in the loop, while in the GR-Dexa/TIF2 simulations (**Figure 9** structures

#3 and #4 green boxes) Dexa is anchored to the loop facilitating ligand unbinding. In the GR(I747M)-Dexa/TIF2 simulations (**Figure 9,** Structures #3 and #4 purple boxes), W557 residue adopts conformations similar to those observed in the FF and FP simulations, so that Dexa follows an unbinding path much more similar to those observed for FF and FP simulations with respect to that observed for Dexa in the WT LBD, but lacking a bulky substituent on C17 Dexa is rapidly driven out of the mutated LBD of GR.

Discussion

The SMD simulations of the glucocorticoid agonists were performed without imposing a predefined unbinding direction and the reaction coordinate defined as the distance between the Cα atom of R611 and the centroid of the ligand allowed Dexa, FP and FF to exit the GR binding site without causing major changes to the conformation of the receptor. However, it cannot be ruled out that these ligands could potentially follow other unbinding pathways (e.g. moving close to the the β-strands or displacing Helix 12).

The unbinding pathway identified, which involves the amino acids in the loop connecting H7 with H8, close to H11, is consistent with the most energetically favorable pathway described in the literature for Estrogen Receptor alpha ligands[23,24] as well as for thyroid hormone receptor.[29,30] Taken together these results seem to suggest a common unbinding pathway for nuclear receptor ligands.

Furthermore, the SMD simulations performed have shown that Dexa can leave the cavity by following two unbinding trajectories. On the contrary, the chain linked to C17 of FF and FP behaves like an anchor point which forces these ligands to follow a unique unbinding trajectory. This behavior is reflected by narrower PMF energy basins with respect to that of Dexa. Furthermore, the energies of FP and FF PMFs are higher with respect to that of Dexa, suggesting that a higher amount of energy is required to drive these ligands out to the GR binding site. The force profiles of Dexa and FF show that the breakage of hydrogen bond interactions of both these ligands with GR is assisted by water molecules. More importantly, the hydrophobic interactions established between the furoate chain of FF and F740 are responsible for delaying ligand unbinding. Lacking a hydrophobic chain at C17, Dexa can more freely translate and rotate

when leaving the cavity. In addition, FP carrying an ethyl ester group rather than the furoate moiety showed a behavior similar to FF.

Furthermore, the calculated PMF for FF, FP and Dexa are qualitatively in good agreement with their experimental rates of dissociation retrieved from the literature.[10] Besides, the calculated SD values for all the ligands are reflected in their experimental values. According to the results obtained, the steered-MD protocol developed appears to be statistically robust and validated to model this chemical class of compounds. At the same time it allows to gather detailed structural information responsible for long $t_{1/2}$ in vitro, which represents valuable information for drug design purposes.

Finally, the simulations performed for I747M GR showed that the steered-MD protocol here described is also able to detect subtle structural and potency differences. The energy of Dexa PMF for this mutated receptor is lower than that for the WT and the energy basin of the former is wider than the latter. Hence, it can be hypothesized that Dexa enters the binding site following the same pathway both in the WT and in this mutated receptors. Assuming that the k_{on} of Dexa is the same for I747M as well as WT GR, the predicted lower k_{off} value for Dexa bound to this mutant gives rise to lower K_d (i.e. reduced Dexa binding affinity), in agreement with experimental data.[25,26] In order to validate this hypothesis, kinetic studies of Dexa in I747M GR should be undertaken in parallel with that of the WT. Furthermore, potential binding pathways could be identified by carrying out site directed mutagenesis experiments of GR (e.g. alanine scanning) followed by kinetic studies with Dexa and the other ligands. This approach has already been described in the literature for estradiol and Estrogen Receptor alpha.[32]

Conclusions

Glucocorticoid receptors have been deeply investigated by medicinal chemists to find GR ligands able to treat several chronic and acute inflammatory conditions like asthma, psoriasis, rheumatoid arthritis and irritable bowel disease.[5,7,8] As a result of these efforts several glucocorticoid agonists have been identified and are currently in clinical use for the treatment of anti-inflammatory diseases.[5,7,8] Nevertheless, these drugs are characterized by severe side effects[8,9] and the development of novel modulators that can dissociate therapeutic from undesired adverse effects is impelling.[8,9] One of the approaches pursued to meet

this objective consisted in changing the structure of the steroid template hoping to improve the safety profile while retaining high potency and efficacy at GR. The most advanced compound identified by undertaking this approach is represented by ZK245186, which is currently in early clinical development for atopic dermatitis[38]. However, none of these novel agonists have been extensively characterized from a pharmacological view point. Hence, the approach relying on extensive in vitro pharmacological profiling of novel compounds against genes transactivated and transrepressed by prototypical corticosteroids such as dexamethasone can represent an alternative way to discover safer GR ligands. This method has been recently described in the literature[39] and has successfully led to the identification of GW870086, a glucocorticosteroid showing a distinct in vitro pharmacological profile with respect to dexamethasone, budesonide and fluticasone propionate. Besides, this steroid derivative has already demonstrated in clinical studies a better safety profile (reduced suppression of HPA axis) with respect to classical glucorticosteroids[40].

Drug design strategies aimed at the identification of novel GR modulators have to be based on a deep understanding of how these ligands interact with their receptor. In this regard, molecular modeling of GRs-GCs complexes can provide valuable support even though they are quite challenging due to the plasticity of GR which can adapt its binding site to different ligands.[3] Besides, GR ligands can bind and dissociate from the GR LBD in a different way as reported for other NRs.[29,30] Furthermore, the exact role of water-mediated hydrogen bond ligand-protein interactions and hydrophobic contacts within the binding site are still elusive.[3,30] Finally, the functional role of the mutated LBDs is still unknown. Herein, by using SMD simulations we propose a detailed atomistic description of the unbinding process of Dexa, FF and FP from GR LBD. The dissociating paths of these prototypical glucocorticosteroids observed during our simulations are qualitatively in agreement with those reported for other NRs.[23,24,29,30] Besides, Dexa can dissociate from the LBD of GR following slightly different dissociating paths, a behavior similar to that observed for other NRs.[29,30] Furthermore the PMF for Dexa dissociation exhibits the highest calculated standard deviation with respect to FF and FP, which is consistent with the experimental data reported in the literature.[10] Detailed analyses of the dissociation pathways of these derivatives clearly show that water molecules facilitate ligand dissociation by assisting the disruption of ligand-protein hydrogen bonds in the binding site. The dissociating path of Dexa is characterized by the presence of fluctuating hydrogen bonding patterns

involving D742 and T744 side chains, while FF preferably interacts with T739 and Q642. The dissociating path of Dexa is also characterized by polar contacts with H11-H12 connecting loop, while during the unbinding process of FF and FP a series of hydrophobic contacts are established among the residues located on H7 and H3 and the bulky ligand chain on C17, which behaves as an anchor and selects a specific dissociating path. Furthermore, our studies allowed for the identification of several residues, in particular W557, F740 and F749, which behave like gatekeepers and control the dissociation of the ligands from the LBD of GR. The behavior of W557 is quite remarkable as it is involved in the formation of a hydrophobic channel crossed by the bulky ligand chain on C17.

The SMD simulations performed provided also valuable information to propose an atomistic explanation for the reduction in efficacy of Dexa in the naturally occurring I747M-GR, a mutated receptor implicated in rare familial glucocorticoid resistance.[25,26] On the basis of our studies one can conclude that I747M affects the dissociation rate of Dexa inducing a faster dissociation of the ligand from the LBD of GR. This seems to be related to the different volume occupied by M747 side chain, which induces a significant rearrangement of the loop and subsequently affects the relocation of the adjacent Dexa. In the GR(I747M) mutant, Dexa follows an unbinding path similar to that observed for FF and FP in the WT. However, lacking a bulky substituent on C17 it rotates in the binding site as observed in the WT LBD, without establishing any interactions with the loop residues. It cannot be ruled out that I747M might have a role in the association rate of Dexa. Unfortunately no evidences have been described in the literature. Nevertheless, it has been reported for other NRs that mutations occurring in H12 affect the ligand association rate, but not the dissociation rate,[32,33] suggesting that ligand binding and unbinding might take different routes. The clear effect of the I747M on the unbinding process of Dexa and the results obtained with our SMD protocol add an important piece of information in the puzzling scenario of GR function, which can be extremely useful in deciphering the structural features underlying the dissociation rate for GCs.

Finally, these results confirm the robustness of the SMD methodology to model ligand-receptor unbinding events. The SMD protocol here described and validated with prototypical glucocorticosteroids can be used to prioritize the synthesis of their structural analogues on the basis of their PMFs and calculated unbinding energies.

Abbreviations

GR, glucocorticoid receptor; NRs, nuclear receptors; LBD, ligand binding domain; GCs, glucocorticoids; PMF, potential mean force; FF, fluticasone furoate; Dexa, dexamethasone; FP, fluticasone propionate.

References

(1) Huang, P.; Chandra, V.; Rastinejad, F. Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics. *Annu. Rev. of Physiol.* **2010**, *72*, 247–272.

(2) Tsai, M.; O'Malley, B. W. Molecular Mechanisms of Action of Steroid/Thyroid Receptor Superfamily Members. *Annu. Rev. Biochem.* **1994**, *63*, 451–486.

(3) Veleiro, A. S.; Alvarez, L. D.; Eduardo, S. L.; Burton, G. Structure of the Glucocorticoid Receptor, a Flexible Protein That Can Adapt to Different Ligands. *Chem. Med. Chem.* **2010**, *5*, 649–659.

(4) Kohn, J. A.; Deshpande, K.; Ortlund, E. A. Deciphering Modern Glucocorticoid Cross-pharmacology Using Ancestral Corticosteroid Receptors. *J. Biol. Chem.* **2012**, *287*, 16267–16275.

(5) Barnes, P. J. Glucocorticosteroids: current and future directions. *Br. J. Pharmacol.* **2011**, *163*, 29–43.

(6) Bledsoe, R. K.; Montana, V. G.; Stanley, T. B.; Delves, C. J.; Apolito, C. J.; McKee, D. D.; Consler, T. G.; Parks, D. J.; Stewart, E. L.; Willson, T. M. Crystal Structure of the Glucocorticoid Receptor Ligand Binding Domain Reveals a Novel Mode of Receptor Dimerization and Coactivator Recognition. *Cell* **2002**, *110*, 93–105.

(7) Barnes, P. J. How corticosteroids control inflammation: Quintiles Prize Lecture 2005. *Br. J. Pharmacol.* **2009**, *148*, 245–254.

(8) Barnes, P. J. Mechanisms and resistance in glucocorticoid control of inflammation. *J. Steroid Biochem. Mol. Biol.* **2010**, *120*, 76–85.

(9) Baudy, A. R.; Reeves, E. K. M.; Damsker, J. M.; Heier, C.; Garvin, L. M.; Dillingham, B. C.; McCall, J.; Rayavarapu, S.; Wang, Z.; Vandermeulen, J.

H.; Sali, A.; Jahnke, V.; Duguez, S.; DuBois, D.; Rose, M. C.; Nagaraju, K.; Hoffman, E. P. Δ-9,11 Modification of Glucocorticoids Dissociates Nuclear Factor- B Inhibitory Efficacy from Glucocorticoid Response Element-Associated Side Effects. *J. Pharmacol. Exp. Ther.* **2012**, *343*, 225–232.

(10) Valotis, A.; Högger, P. Human receptor kinetics and lung tissue retention of the enhanced-affinity glucocorticoid fluticasone furoate. *Respir. Res.* **2007**, *8*, 54.

(11) Biggadike, K.; Bledsoe, R. K.; Hassell, A. M.; Kirk, B. E.; McLay, I. M.; Shewchuk, L. M.; Stewart, E. L. X-ray Crystal Structure of the Novel Enhanced-Affinity Glucocorticoid Agonist Fluticasone Furoate in the Glucocorticoid Receptor−Ligand Binding Domain. *J. Med. Chem.* **2008**, *51*, 3349–3352.

(12) Rossios, C.; To, Y.; To, M.; Ito, M.; Barnes, P. J.; Adcock, I. M.; Johnson, M.; Ito, K. Long-acting fluticasone furoate has a superior pharmacological profile to fluticasone propionate in human respiratory cells. *Eur. J. Pharmacol.* **2011**, *670*, 244–251.

(13) Lu, H.; Tonge, P. J. Drug–target residence time: critical information for lead optimization. *Current Opinion in Chemical Biology* **2010**, *14*, 467–474.

(14) Tummino, P. J.; Copeland, R. A. Residence Time of Receptor−Ligand Complexes and Its Effect on Biological Function. *Biochemistry* **2008**, *47*, 5481–5492.

(15) Kola, I.; Landis, J. Opinion: Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Disc.* **2004**, *3*, 711–716.

(16) Pammolli, F.; Magazzini, L.; Riccaboni, M. The productivity crisis in pharmaceutical R&D. *Nat. Rev. Drug Disc.* **2011**, *10*, 428–438.

(17) Schmidtke, P.; Luque, F. J.; Murray, J. B.; Barril, X. Shielded Hydrogen Bonds as Structural Determinants of Binding Kinetics: Application in Drug Design. *J. Am. Chem. Soc.* **2011**, *133*, 18903–18910.

(18) Colizzi, F.; Perozzo, R.; Scapozza, L.; Recanatini, M.; Cavalli, A. Single-Molecule Pulling Simulations Can Discern Active from Inactive Enzyme Inhibitors. *J. Am. Chem. Soc.* **2010**, *132*, 7361–7371.

(19) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.

(20) Liphardt, J. Equilibrium Information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality. *Science* **2002**, *296*, 1832–1835.

(21) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.* **2003**, *119*, 3559.

(22) Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946.

(23) Burendahl, S.; Danciulescu, C.; Nilsson, L. Ligand unbinding from the estrogen receptor: A computational study of pathways and ligand specificity. *Proteins* **2009**, *77*, 842–856.

(24) Shen, J.; Li, W.; Liu, G.; Tang, Y.; Jiang, H. Computational Insights into the Mechanism of Ligand Unbinding and Selectivity of Estrogen Receptors. *J. Phys. Chem. B* **2009**, *113*, 10436–10444.

(25) Bray, P. J.; Cotton, R. G. H. Variations of the human glucocorticoid receptor gene (NR3C1): Pathological and in vitro mutations and polymorphisms. *Human Mutation* **2003**, *21*, 557–568.

(26) Vottero, A. A Novel, C-Terminal Dominant Negative Mutation of the GR Causes Familial Glucocorticoid Resistance through Abnormal Interactions with p160 Steroid Receptor Coactivators. *J. Clin. Endocrinol. Metab.* **2002**, *87*, 2658–2667.

(27) Charmandari, E. Natural Glucocorticoid Receptor Mutants Causing Generalized Glucocorticoid Resistance: Molecular Genotype, Genetic Transmission, and Clinical Phenotype. *J. Clin. Endocrinol. Metab.* **2004**, *89*, 1939–1949.

(28) Charmandari, E.; Kino, T.; Ichijo, T.; Chrousos, G. P. Generalized Glucocorticoid Resistance: Clinical Aspects, Molecular Mechanisms, and Implications of a Rare Genetic Disorder. *J. Clin. Endocrinol. Metab.* **2008**, *93*, 1563–1572.

(29) Martínez, L.; Sonoda, M. T.; Webb, P.; Baxter, J. D.; Skaf, M. S.; Polikarpov, I. Molecular Dynamics Simulations Reveal Multiple Pathways of Ligand Dissociation from Thyroid Hormone Receptors. *Biophys. J.* **2005**, *89*, 2011–2023.

(30) Martínez, L.; Webb, P.; Polikarpov, I.; Skaf, M. S. Molecular Dynamics Simulations of Ligand Dissociation from Thyroid Hormone Receptors: Evidence of the Likeliest Escape Pathway and Its Implications for the Design of Novel Ligands. *J. Med. Chem.* **2006**, *49*, 23–26.

(31) Álvarez, L. D.; Martí, M. A.; Veleiro, A. S.; Presman, D. M.; Estrin, D. A.; Pecci, A.; Burton, G. Exploring the Molecular Basis of Action of the Passive Antiglucocorticoid 21-Hydroxy-6,19-epoxyprogesterone. *J. Med. Chem.* **2008**, *51*, 1352–1360.

(32) Zhong, L.; Skafar, D. F. Mutations of Tyrosine 537 in the Human Estrogen Receptor-α Selectively Alter the Receptor's Affinity for Estradiol and the Kinetics of the Interaction. *Biochemistry* **2002**, *41*, 4209–4217.

(33) Yudt, M. R.; Vorojeikina, D.; Zhong, L.; Skafar, D. F.; Sasson, S.; Gasiewicz, T. A.; Notides, A. C. Function of Estrogen Receptor Tyrosine 537 in Hormone Binding, DNA Binding, and Transactivation. *Biochemistry* **1999**, *38*, 14146–14156.

(34) Schrödinger, Mestro version 9.1 Schrödinger, LLC, New York 2009 http://www.schrodinger.com/

(35) Chemical Computing Group. Scalable Software Scalable Science http://www.chemcomp.com/

(36) The Amber Molecular Dynamics Package ambermd.org http://ambermd.org/

(37) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961–1972.

(38) Schacke, H.; Zollner, T.M.; Docke W.; Rehwinke H.; Jaroch S.; Skuballa W. et al. ZK 245186, a novel, selective glucocorticoid receptor agonist (SEGRA) for the topical therapy of inflammatory skin diseases. *Exp. Dermatol.* **2009**, *18*, 305.

(39) Uings I.J.; Needham D.; Matthews J.; Haase M.; Austin R.; Angell D.; Leavens K.; Holt J.; Biggadike K.; Farrow S. N. Discovery of GW870086: a potent anti-inflammatory steroid with a unique pharmacological profile. *Br. J. Pharmacology* **2013**, *169*, 1389-1403.

(40) Farrow S.N. Discovery of GW870086: a potent anti-inflammatory steroid with a unique pharmacological profile. European Respiratory Socienty meeting held in Barcellona (Spain), September 2013. Oral presentation.

MODELLING STUDIES OF VEGFR2 AND ITS INHIBITORS

Introduction

The vascular endothelium growth factor (VEGF) is an important signaling protein implicated in the process of vasculogenesis and angiogenesis. It binds preferably the human Vascular Endothelial Growth Factor Receptor 2 (VEGFR2), a type III receptor tyrosine kinase (RTK) in the PDGFR family with PDGFRα/β, c-KIT, FLT3 and CSF-1 (cFMS) as other members. Upon VEGF binding to the extracellular domain of VEGFRs, these receptors dimerize and are activated by phosphorylation. The extracellular portion of VEGFRs, containing a 7-immunoglobulin like domain, is connected to the intracellular kinase catalytic domain (CD) through the transmembrane (TM) domain, adopting an alpha helical conformation embedded in the phospholipid bilayer, and a stretch of amino acids referred to as the juxtamembrane domain (JM).

VEGFR2 is a clinically validated target to treat Renal Cell Carcinoma (RCC)[1]. In addition to efficacy seen in RCC, VEGF signaling inhibition has been associated to side effects, in particular hypertension, consistently observed for Tyrosine Kinase Inhibitors (TKI) and the monoclonal antibody bevacozumab. Both efficacy and safety profiles of approved VEGFR TKI depend on their selectivity profile, which has been clinically confirmed in a published comparative study[2]. Hence, VEGFR TKI represent a unique opportunity to study the impact of molecular interactions and physico-chemical properties on potency and selectivity and how these in vitro measures translate to differentiated clinical profiles.

Recently the crystal structures of the clinically relevant TKIs sunitinib and sorafenib, as well as that of a very close analogue of the potent inhibitor TAK-593 (**Figure 1**), have been solved bound to the unphosphorylated kinase domain of VEGFR2 together with the enzyme JM domain[3]. These crystal structures show that all these ligands map the ATP binding site (ATP competitive inhibitors) and the kinase DFG motif adopts the "out" conformation projecting F1047 towards the ATP cavity. Besides, in both crystal structures the JM domains have not been fully solved and the kinase insert domain is missing. It is really interesting to highlight that the conformation of the JM domain is remarkably different in the two crystal structures. As shown in **Figure 2**, in the

sunitinib complex (VEGFR2/SUN, PDB code 4AGD) the N-terminus portion of the JM domain is close to the DFG motif and I804 in the JM domain is projected beneath D1046 (JM_{in} conformation) of the DFG motif.

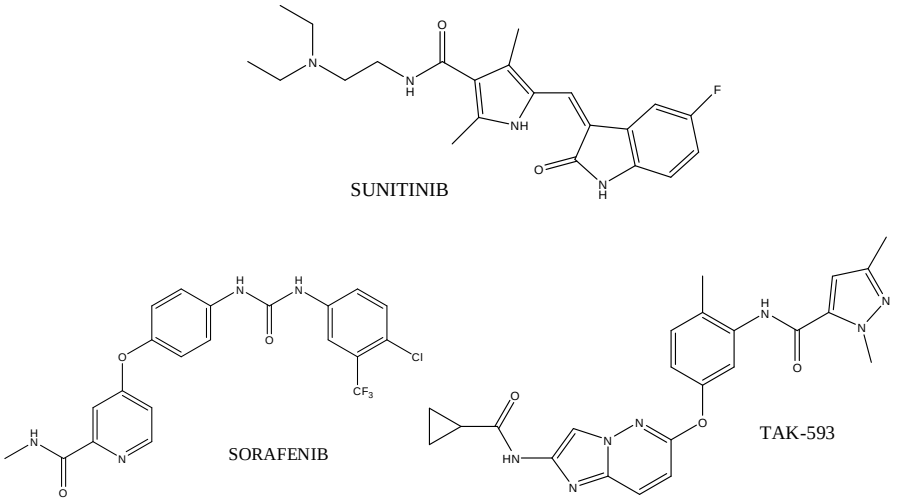


Figure 1. Chemical structures of sunitinib (Sutent, top left), sorafenib (Nexavar, top right) and TAK-593 (bottom).

On the contrary, sorafenib (VEGFR2/SOR, PDB code 4ASD) binds deeper in the VEGFR2 ATP binding site (**Figure 3**). It is likely that the pendant aromatic ring of the ligand decorated with trifluoromethyl and chlorine substituents clashes with I804 forcing the JM domain to move away from the DFG-motif and to adopt the "out" conformation (JM_{out}).

As shown in **Figure 4**, the overlay of these crystal complexes emphasizes how the former inhibitor binds deeper than the latter in the VEGFR2 ATP binding site.

A close structural analogue of TAK-593 (PDB code: 3VO3) binds similarly to sorafenib and the JM domain adopts the "out" conformation (structure not shown).

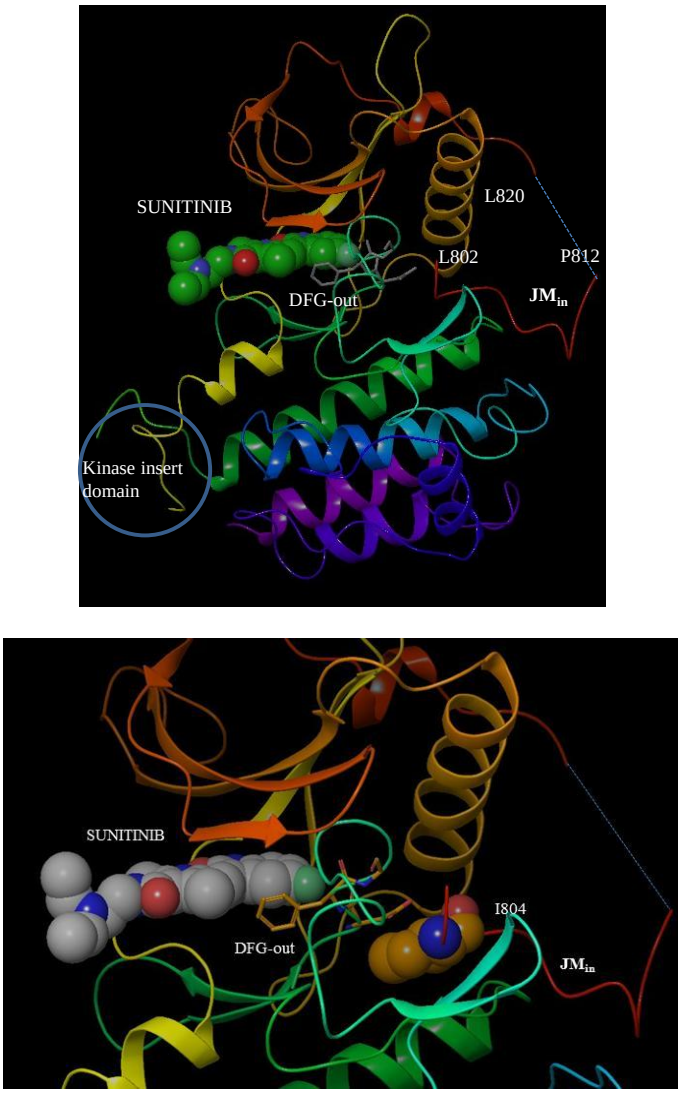


Figure 2. Top: Overview of the crystal complex of sunitinib with the catalytic and the JM domains of human VEGFR2 (PDB code 4AGD). The JM domain has not been solved between P812 and L820, as highlighted by the cyan dotted line. Furthermore, this structure lacks the kinase insert domain. **Bottom:** I804 in the JM domain adopting the "in" conformation is close to the DGF-out motif.

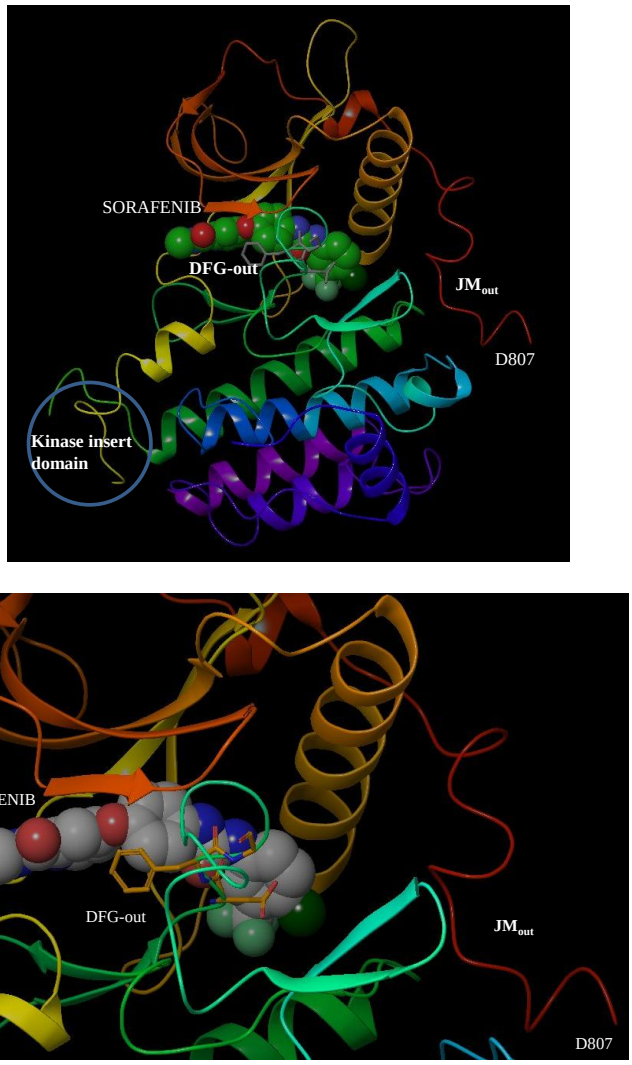


Figure 3. Top: Overview of the crystal complex of sorafenib with the catalytic and the JM domains of the human VEGFR2 (PDB code 4ASD). The JM domain has not been solved between L802 and M806. Furthermore, the structure lacks the kinase insert domain highlighted by the white ring. **Bottom:** I804 in the JM domain adopting the "in" conformation is close to the DGF-out motif.

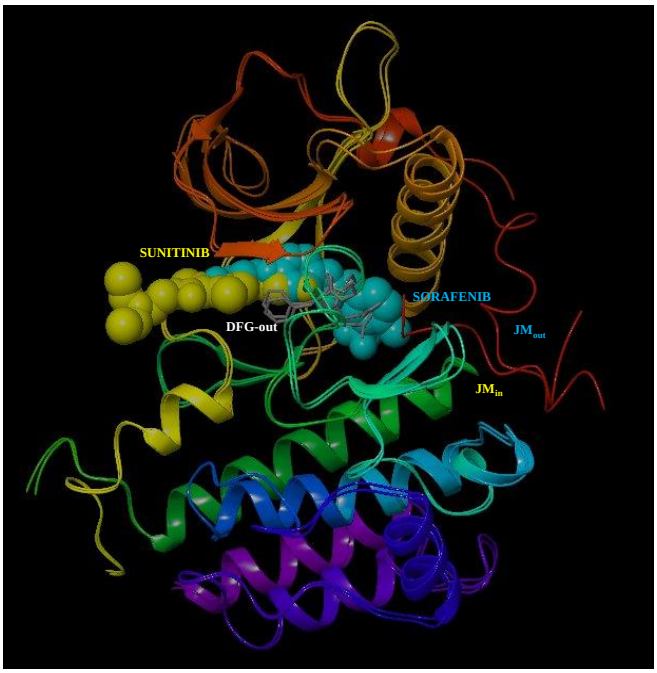


Figure 4. Overlay of sorafenib (cyan) and sunitinib (yellow) crystal complexes. As shown, sorafenib binds deeper in the binding site displacing the JM domain.

Regulation of RTKs is controlled by multiple phosphorylation events. Autophosphorylation of residues on the activation loop of the catalytic domain is a well-characterized mechanism regulating the kinase enzymatic activity. Autophosphorylation of the JM domain may represent an alternative mechanism. In particular, the JM domain of type III RTKs FLT3, c-Kit and CSF-1 (cFMS) can bind to the interface between the N- and C-terminus of the catalytic domain, preventing the activation loop of the C-terminus from adopting a catalytically productive conformation (DFG-in motif) and disrupting the N-terminal αC-helix. In the X-ray structure of sunitinib bound to VEGFR2, the JM domain adopts a binding mode similar to that observed in the other type III RTKs suggesting that this ligand binds to the autoinhibited state of VEGFR2 (**Figure 5**). Similarly, sorafenib binds the inactivated state of this kinase (DFG-out motif). However, its binding requires the displacement of the JM domain.

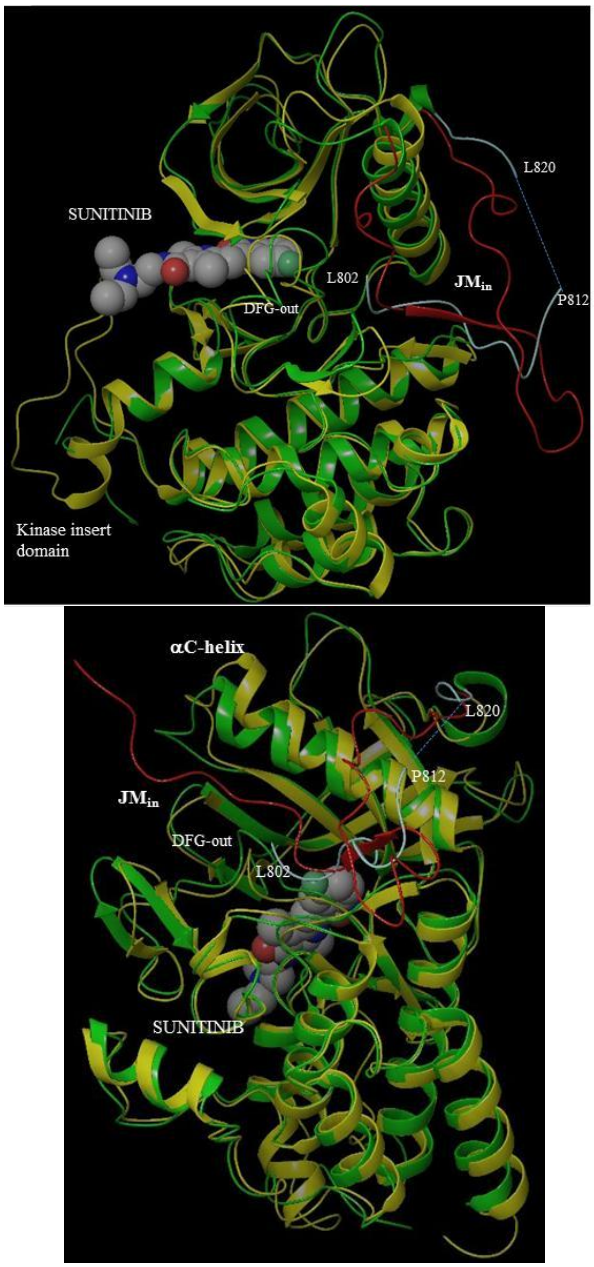


Figure 5. Overlay of the crystal structure of sunitinib bound to VEGFR2 (CD domain: green; JM domain: cyan, PDB code: 4AGD) to Apo enzyme c-Kit (CD domain: yellow, JM domain: red; PDB code: 1T45). In the c-Kit structure the JM domain is close to the DFG-motif adopting the "out" conformation and is projected beneath the αC-helix. The JM domain in 4AGD has not been completely solved.

The construct utilized to generate the VEGFR2 crystal structures described above was used to determine the potencies, time dependencies and selectivity of a panel of TKI, including sunitinib and sorafenib[4]. Potency experiments highlighted that in vitro kinase activity of sunitinib is affected by the presence of the JM domain in the enzyme construct used to carry out the enzymatic assay. For this derivative the enzymatic kinase K_i is equal to 1.1nM when JM is not present in the enzyme construct. However, a 2-fold increase in the enzymatic activity of this derivative (K_i = 0.02nM) was observed for the enzyme construct carrying the JM domain. Besides, the enzymatic inhibition detected in the presence of the JM domain better correlates with the cellular potency of this derivative (VEGF-induced VEGFR2 autophosphorylation assay pVEGFR2 in human umbilical vein endothelium cells HUVEC IC_{50} = 0.013nM). According to these X-ray structures, the JM domain does not seem to directly interact with sunitinib. Instead, it is likely that it locks the DFG triad in the out conformation, hence stabilizing the autoinhibited kinase conformation (JM_{in} conformation). On the contrary, the enzymatic activity of sorafenib does not depend on the presence of the JM domain in the enzyme construct, demonstrating that it does not directly interact with this ligand, as shown by the X-ray structure (JM_{out} conformation). In addition, sorafenib showed a time-dependent inhibition only with the enzyme construct carrying the JM domain. This suggests that the JM domain in the autoinhibitory conformation has to rearrange so that sorafenib can be accommodated in the binding site. The increase in the size of type II inhibitors is associated with weaker efficiency and selectivity, supported by VEGF-dependent cell survival assay in HUVEC IC_{50} values. Unfortunately, no type-I inhibitors are known to prove their potential preference for a DGF-in (activated) kinase conformation.

Besides, kinetic parameters (k_{off}) of these ligands were measured with the construct bearing the JM domain. While sunitinib is a prototypical type IV inhibitor like axitinib, showing short dissociative half-life from VEGFR2 ($t_{1/2}$ about a few minutes)[4b], sorafenib is an exemplar of type II inhibitor like TAK-593, characterized by long dissociative half-life from VEGFR2 receptor ($t_{1/2}$ of about a few hours)[4b]. Long residence time at the enzyme might lead to a long duration of action in vivo.

The objective of this study consists in elucidating at atomistic level the unbinding pathways of the prototypical VEGFR2 inhibitors sunitinib and sorafenib with the use of steered-MD and to possibly identify a qualitative

relationship (either a trend or a rank order) between their calculated unbinding energies and the experimental rate of dissociation (k_{off}) from VEGFR2.

Methods

(i.) Protein preparation

3D coordinates of sunitinib bound to the inactivated catalytic domain of the human VEGFR2 crystal structure, solved together with the juxtamembrane domain, were retrieved from the Protein Data Bank (PDB code: 4AGD, JM_{in}). Five alanine residues were introduced to join Y938 to Y996, corresponding to the kinase insert domain (KID) which plays a structural role by connecting two α helices in the C-terminus CD. This loop had been deleted for crystallization purposes without affecting the intrinsic kinase activity. Furthermore, the amino acids 813-819 in the JM domain, not solved in the X-ray structure, were manually added within Maestro v9.1[5]. Subsequently, these fragments were gently relaxed with OPLS FF (Maestro v9.1) while keeping fixed the rest of the protein. All the water molecules present in the X-ray structure were removed. Sunitinib was modeled starting from its crystallographic coordinates (VEGFR2/SUN). Similarly, the crystal structure of human VEGFR2 (Juxtamembrane and catalytic domains) in complex with sorafenib was retrieved from the Protein Data Bank (PDB code: 4ASD, JM_{out}). Similarly to sunitinib complex, five alanine residues were added to bridge Y938 to Y996. Besides, amino acids 802-806 in the N-terminus domain, not solved in the X-ray structure, were manually added within Maestro v9.1 in an extended conformation and away from the catalytic domain. Then, the added amino acids were gently minimized with OPLS FF (Maestro) while keeping fixed the rest of the protein. All the water molecules present in the X-ray were removed. Sorafenib was modeled starting from its crystallographic coordinates (VEGFR2/SOR). Following this preparation both the complexes have the same number of atoms.

(ii.) Ligand preparation

The geometry of sorafenib and sunitinib was optimized using the *Jaguar* tool available in Maestro v9.1, partial charges were computed using the HF/6-31G** level of theory and fixed using the RESP methodology. Atom types and ligand parameters were assigned using the general Amber force field (gaff)[6].

(iii.) Ligand-receptor complexes preparation

Each ligand-receptor construct was finally parameterized using the *amber99SB* and *gaff* force fields[6]. The complexes thus obtained were solvated in a cubic box of 12 Å using Leap[6] and TIP3P as water models. Na^+ ions were added to neutralize the net charge of the system.

(iv.) MD simulations parameters and protocols

All MD simulations were performed with Amber11[6]. All systems were gradually minimized and then equilibrated for 0.5 ns. During each equilibration phase the NPT ensemble was used with a pressure target equal to 1 atm and temperature of 310 K, pressure relaxation time of 1ps and collision frequency set to $5ps^{-1}$. The non-bonded interactions cutoff was set to 10 Å, and the PME algorithm was used with a grid spacing of 1 Å. A two steps equilibration procedure was used. Each step lasted for 0.25 ns, with an integration step of 1 fs. The first 0.25 ns of the equilibration were conducted by imposing harmonic constraints equal to 2 kcal/mol*$Å^2$ to the protein backbone and the ligand atoms. For the last 0.25 ns the same MD conditions were used and the constraints on ligand atoms were set to 2 kcal/mol*$Å^2$. Finally a production run was conducted for a total of 20 ns for each system. The first 10 ns were done using the NPT ensemble, while for the last 10 ns the NVT ensemble was used, so as to prepare the systems for the following SMD runs. During the production phase, the integration step was set equal to 2 fs and only bonds involving hydrogen atoms were constrained. The MD conditions were kept equal to those used for the equilibration phase, without imposing any harmonic constraints on the system. Along the last 5ns of NVT production phase (5000 frames) frame #3000, #3500, #4000, #4500 and #5000 were taken and used for constant-velocity SMD (NVT ensemble, Amber12)[7-9].

(v.) Calibration of the SMD simulations parameters

Both the constant velocity and the force constant were identified using the following step-wise optimization: (*i*) SMD simulations were conducted on the last sampled frame of the NVT trajectory of VEGFR2/SUN (fast dissociating ligand) and VEGFR2/SOR (slow dissociating ligand) complexes by setting the constant velocity v to 0.0001, 0.0002, 0.00025 and 0.0005 Å/ps while keeping fixed the force constant k to 15 kcal/mol*$Å^2$; (*ii*) SMD simulations were conducted on both the complexes by setting the spring constant value k at 5, 10, 15, 20, 30 and 40 kcal/mol*$Å^2$ while keeping fixed the constant velocity v at

0.0005 Å/ps, respectively; (*iii*) selection of the smallest spring constant value and velocity which allows for the conservation of the stiff-spring approximation for both the systems (v=0.0005 Å/ps, k=40 kcal/mol*$Å^2$).

The reaction coordinate was initially defined as the distance between the ligand center of mass and the Cα atom of Val899, buried in the ATP binding site, which was changed for 1 Å (k=15 kcal mol^{-1} $Å^{-2}$, v=0.0001, 0.0002, 0.00025 Å/ps) or 2 Å (k=15, 20, 25, 30, 35 and 40 kcal mol^{-1} $Å^{-2}$, v=0.0005 and 0.001 Å/ps). Further attempts to model the unbinding reaction coordinates were done for VEGFR2/SOR complex. For these experiments the reaction coordinate was defined as: 1) the distance between the center of mass of the ligand and the Cα atom of Glu885 lying in the αC-helix; 2) the distance between the center of mass of the ligand and the centroid of Cα atoms of A866-M869, L912-V916 (N-lobe β-sheets); 3) the distance between the center of mass of the ligand and the centroid of Cα atoms of S1009, A1013, F1091, L1095 (C-lobe α-helices); 4) the distance between the center of mass of the ligand and the centroid of the whole kinase; 5) the distance between the ligand center of mass and the sorafenib binding site center of mass. Visual inspection and RMSD calculation of the frames sampled for each reaction coordinate were monitored to spot kinase conformational change.

All these Amber experiments were run with Cineca Consortium Eurora CPUs (64 processors) and lasted for 12 hrs./processor when v=0.0005 Å/ps, k=40 kcal mol^{-1} $Å^{-2}$ and 25 hours for v=0.0001 Å/ps, k=15 kcal mol^{-1} $Å^{-2}$. CPUs rather than GPUs were always used to accomplish these experiments. This decision was based on Amber manual recommendations as the Amber GPU code is still very new and has not been tested as extensively as the CPU one. Therefore it has to be used with caution and in case of problems the same experiment should be repeated with CPUs. Even though SMD (Jarzysnki sampling, jar=1 i.e. nmropt=1) appears to be supported, the use of nmropt=1 should be avoided as "Currently there is no CUDA kernel for the nmropt features. Hence, simulations using nmropt=1 require additional synchronizations between GPU and CPU on every MD step. Depending on system size this setting can have a big impact on performance" (Amber manual v12).

(vi.) Potential of Mean Force Calculation (PMF, Φ)

Frame#3000 and frame#5000 of sorafenib were used as starting structure for two 36ns long independent SMD trajectories (3,600 frames sampled for each one). Frame#3000 of sunitinib was utilized to run a 24ns-long simulation (2,400

frames sampled) while frame #5000 was modeled till 8 Å of reaction coordinate displacement due to limited computational resources and timing restrictions. All the SMD runs were conducted imposing a pulling velocity (*v*) of 0.0005 Å/ps and k= 40 kcal mol^{-1} $Å^{-2}$. The value of the exerted force (*f*) was printed out every 5 ps (*dt*) and the work done on the system during SMD was calculated by numerical integration (eq.1)[7-9]. The stiff spring approximation (eq. 2) [7-9] was satisfied with a spring constant of 40 kcal/mol*$Å^2$. As soon as computer power is available, all the other starting frames of sorafenib and sunitinib will be simulated in the same way to achieve statistically robust results.

PMF along the unbinding reaction coordinate was reconstructed from the SMD trajectories described above and using Jarzynski's equality (eq. 3)[7-9]. The cumulant expansion approach up to second order (eq. 4)[7-9] was used for reconstructing the PMF.

$$W_{(x(t))} = \int_0^{x_{(t)}} F(t)dx(t) \quad (1)$$

$$F(\lambda) \cong \phi(\lambda) \quad (2)$$

$$e^{-\beta\Delta F} = \langle e^{-\beta W} \rangle \quad (3)$$

$$\Delta F = \langle W \rangle - \sigma^2 \frac{W}{2K_B T} \quad \text{where} \quad \sigma^2 W = \langle W^2 \rangle - \langle W \rangle^2 \quad \text{and} \quad K_B T = 0.6186 \quad (4)$$

Results

(i) Calibration of pulling velocity and force constant

The first set of experiments were carried out on the last frame of the NVT trajectory of each ligand and were devoted to the optimization of the constant velocity while keeping fixed the force constant. At this purpose, the force constant k was fixed at 15 kcal mol^{-1} $Å^{-2}$ while the constant velocity was set to 0.0001, 0.0002, 0.00025, 0.0005, 0.001 and 0.01 Å/ps. At the end of each experiment, the end-to-end distance of each frame was plotted *vs.* the actual reaction coordinate. As shown in **Figure 6** and **7**, even for the slowest simulation (v=0.0001 Å/ps, 10ns simulation, 1200 frames sampled) the force constant seems to be too loose to guarantee the stiff spring approximation satisfaction, in particular for the tighter bound ligand sorafenib (**Figure 7**). Further reduction of the constant velocity was not tried as the timing required

by these simulations would require longer time than that performed with v=0.0001 Å/ps (> 23 hours, 64 processors). Results of the experiments performed with higher pulling velocity (v= 0.001 and 0.01 Å/ps, respectively) are reported in **Appendix 1**.

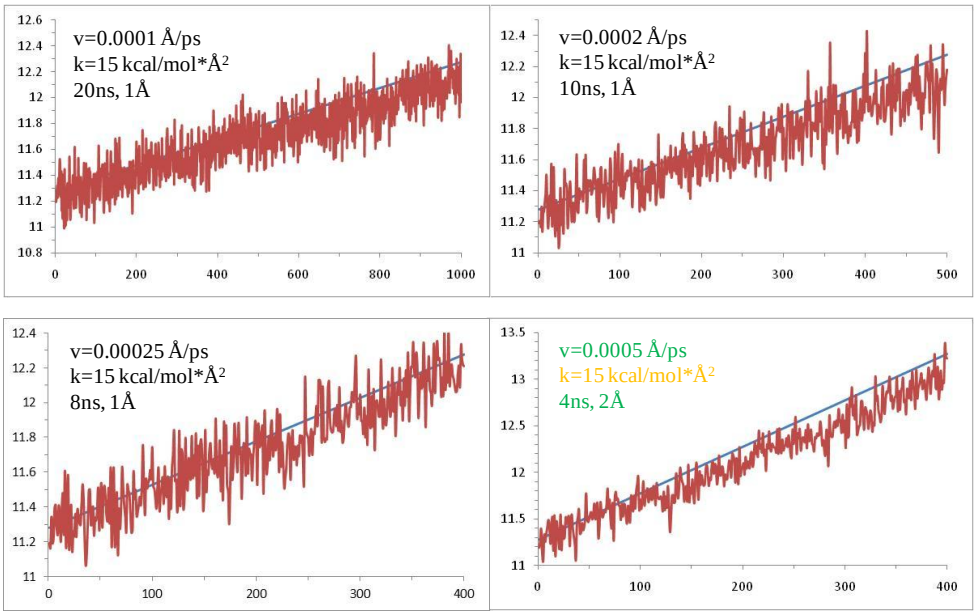


Figure 6. VEGFR2/SUN complex. Plot of the end-to-end distance (red line) and constrained distance (blue line) (Y-axis) measured in each frame (X-axis) in VEGFR2/SUN complex simulations performed with k=15 kcal mol^{-1} $Å^{-2}$ and v= 0.0001, 0.0002, 0.00025, 0.0005 Å/ps, respectively.

Hence, the constant velocity was fixed to v = 0.0005 Å/ps while the force constant k was set to 15, 20, 25, 30, 35 and 40 kcal mol^{-1} $Å^{-2}$ (4ns & 400 frames sampled/run). The simulations performed for the VEGFR2/SUN complex (**Figure 8**) showed that the stiff spring approximation is satisfied by fixing k= 30 kcal mol^{-1} $Å^{-2}$. However, for simulations of the tighter binding inhibitor sorafenib VEGFR2/SOR complex (**Figure 9**) the force constant had to be further increased and the stiff spring approximation is fulfilled when k=40 kcal mol^{-1} $Å^{-2}$. This simulation lasted for 12 hours/processor (64 processors in total). No significant changes in the conformation of the enzyme were observed during these calibration experiments. Results of additional tests performed with higher force constant values are reported in **Appendix 1**.

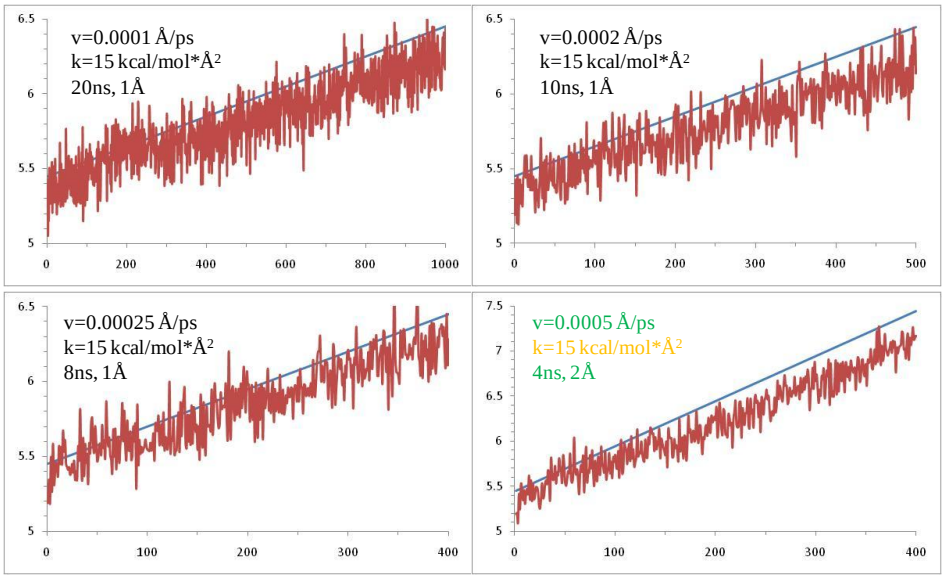


Figure 7. VEGFR2/SOR complex. Plot of the end-to-end distance (red line) vs. the constrained distance (blue line, Y-axis) measured in each frame (X-axis) in VEGFR2/SOR complex simulations performed with k=15 kcal mol^{-1} $Å^{-2}$ and v= 0.0001, 0.0002, 0.00025, 0.0005 Å/ps, respectively.

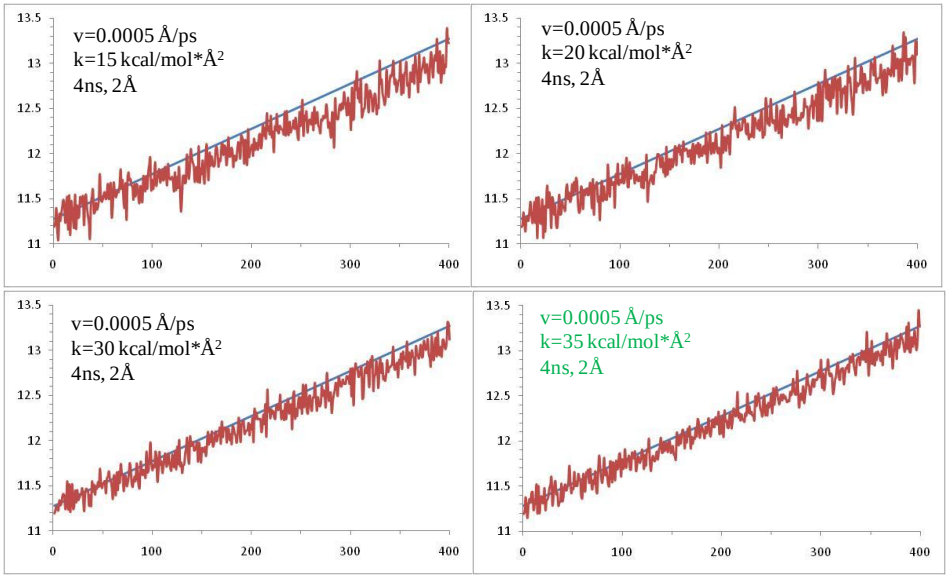


Figure 8. VEGFR2/SUN complex. Plot of the end-to-end distance (red line) vs. the constrained distance (blue line, Y-axis) measured in each frame (X-axis) in VEGFR2/SUN complex simulations performed with v= 0.0005 Å/ps and

k=15, 25, 30 and 35 kcal mol^{-1} $Å^{-2}$ In light of these results, v=0.0005 Å/ps and k=40 kcal/mol*$Å^2$ parameters were used for subsequent studies as they represent a trade-off between accuracy and timing.

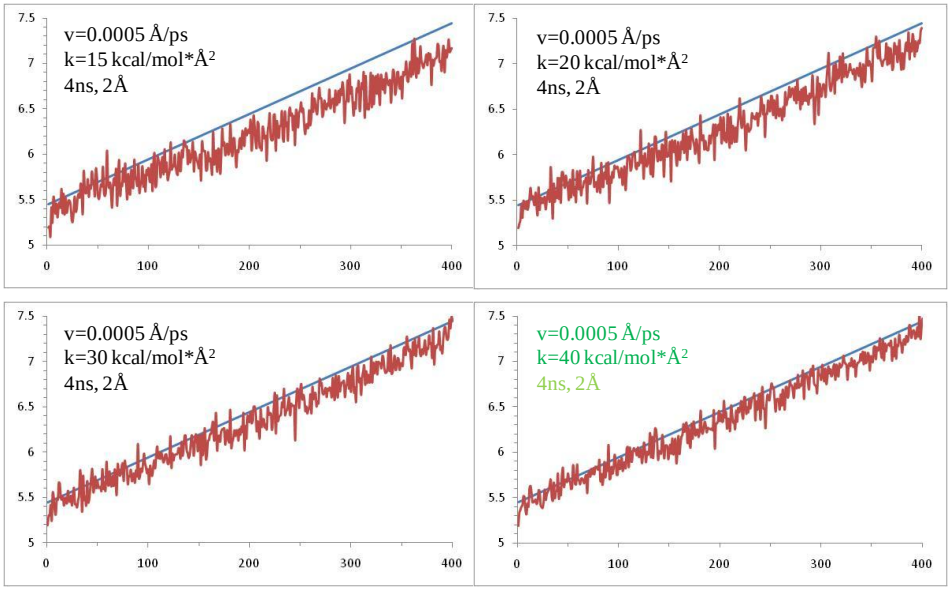


Figure 9. VEGFR2/SOR complex. Plot of the end-to-end distance (red line) vs. the constrained distance (blue line, Y-axis) measured in each frame (X-axis) in VEGFR2/SOR complex simulations performed with v= 0.0005 Å/ps and k=15, 20, 30 and 40 kcal/mol*$Å^2$.

(i) Selection of the reaction coordinate

Subsequent SMD simulations were devoted to the identification of the most appropriate reaction coordinate for unbinding experiments. Unfortunately, Random Accelerated Molecular Dynamics (RAMD)[10] is not available within Amber v11 and v12. Hence, attempts to randomly identify ligand way-outs could not be tried and the selection of the unbinding coordinates was based on visual inspection of the crystal complexes available.

In the first experiment, the reaction coordinate was defined as the distance between the ligand center of mass and the Cα atom of Val899 (V98 Amber pdb numbering) which lies approximately in the middle of the ligand binding site.

SMD simulations of frame#3000 of VEGRF2/SOR were performed by changing the reaction coordinate between 0 and 4 Å in 2 Å/step. At the end of the 8ns-long simulation (800 frames sampled), even though the stiff spring approximation is fairly well satisfied (**Figure 10**, top right), a remarkable change was observed in the conformation of the loop hosting Val899, higher than the crystal structure resolution (**Figure 10,** top left; the blue arrow highlights the conformational change). This conformational change is also reflected by the increase in the RMSD of the protein (**Figure 10,** bottom left) even though no rupture point was identified in the force plot (**Figure 10**, bottom right). This is due to the fact that the ligand adopts a pretty similar orientation as that observed in the X-ray structure. It is likely that the conformational change observed for the loop hosting Val899 is an artifact of the simulation performed.

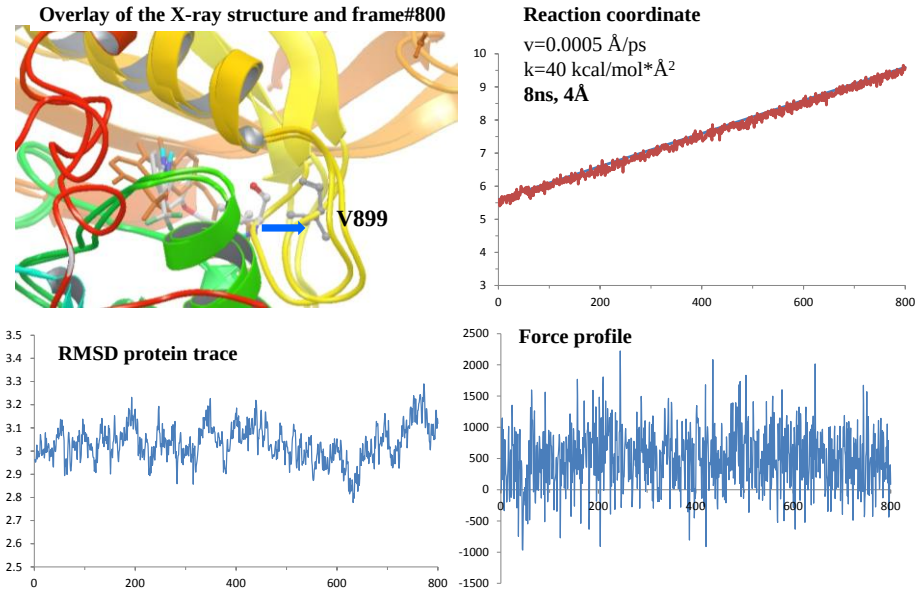


Figure 10. VEGRF2/SOR complex. Top left: Overlay of the VEGFR2/SOR X-ray structure with respect to the last frame sampled after 8ns simulation. **Top right:** Plot of the end-to-end distance (blue line) vs. the actual value (red line) of the reaction coordinate. **Bottom left**: Plot of the protein RMSD calculated by using protein trace. **Bottom right:** Plot of the force (pN) in the frames sampled.

In light of these results, the reaction coordinate was changed and defined as the distance between the center of mass of the ligand and the Cα atom of Glu885 lying in the αC-helix. This residue is involved in a salt-bridge interaction with Lys868 and lies approximately in the middle of the inhibitor binding sites.

Frame#3000 of both the systems under study were then simulated by changing the new reaction coordinate between 0 and 6 Å in 2 Å/step. The pulling velocity and the force constant were set as in the previous experiments (v=0.0005 Å/ps and k=40 kcal/mol*$Å^2$). In **Figure 11** are reported the results of the simulations performed for VEGFR2/SOR complex. As shown, the stiff spring approximation is well satisfied (**Figure 11**, top right). However, visual inspection of the frames sampled between 4 and 6 Å and subsequent RMSD calculation for protein trace highlighted a remarkable movement of the αC-helix while the loop hosting Val899 fluctuated within the crystal structure resolution (**Figure 11**, bottom left). At the same time, ligand did not significantly change its orientation and conformation with respect to those observed in the crystal structure (**Figure 11,** top left). Consequently, no rupture point was identified in the force plot (**Figure 11**, bottom right). Similar results were obtained for VEGR2/SUN complex (**Appendix 2, Figure A2_S1**). Hence, this data suggests that the αC-helix displacement is most likely an artifact of the SMD simulations.

Further experiments using alternative reaction coordinate definition gave rise to a significant change in the enzyme conformation and sorafenib in particular was not able to find any exit door. On the contrary, modeling studies of sunitinib appeared to be overall less problematic in this respect. Details of these trials are reported in **Appendix 2 (Figures A2_S2, A2_S3, A2_S4, A2_S5 and A2_S6).**

When the reaction coordinate was defined as the distance between the ligand centre of mass and the centroid of sorafenib binding site, defined as the enzyme residues lying within 8 Å from each ligand atom, this ligand was able to find a way-out from VEGFR2 binding site.

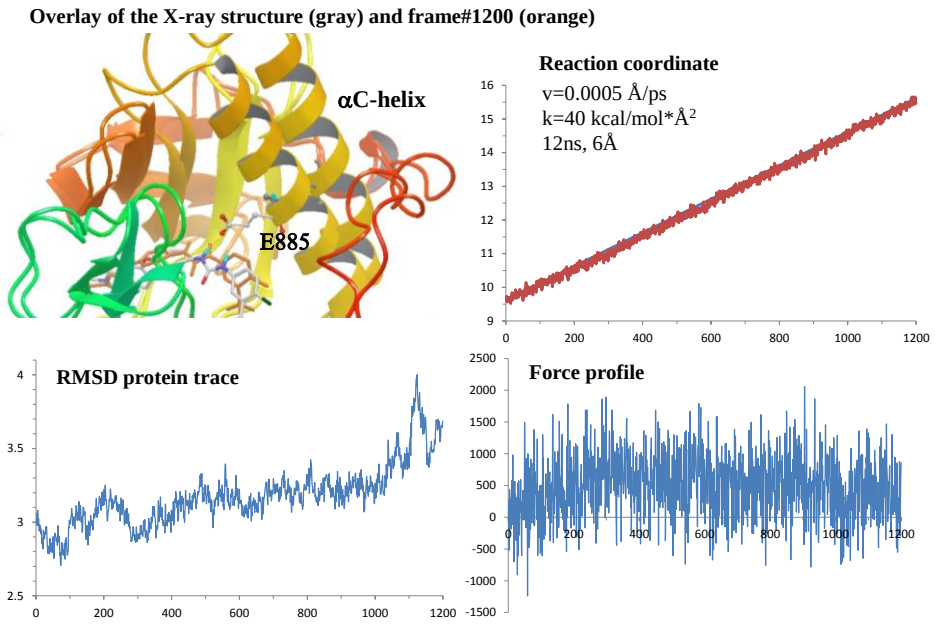


Figure 11. VEGRF2/SOR complex. Top left: Overlay of the VEGFR2/SOR X-ray structure with respect to the last frame sampled after 12ns-long simulation. **Top right:** Plot of the end-to-end distance (blue line) vs. the actual value (red line) of the reaction coordinate. **Bottom left**: Plot of the protein RMSD calculated by using protein trace. **Bottom right:** Plot of the force (pN) in the frames sampled.

In these sets of experiments this tightly bound inhibitor was able to escape from the enzyme binding site (**Figure 12**, top left) by changing the orientation of the αC-helix but without disrupting the conformation of the whole kinase (**Figure 12**, bottom left). At the same time the distance constraint was fairly well satisfied (**Figure 12**, top right) and the force profile (**Figure 12**, bottom right) highlights the presence of a rupture point between 2ns and 4ns (between frames 400 and 600). Quite interestingly, the ligand is leaving the binding site by moving towards the JM domain rather than from the ATP binding site entrance, as this cavity is wide-open and solvent exposed.

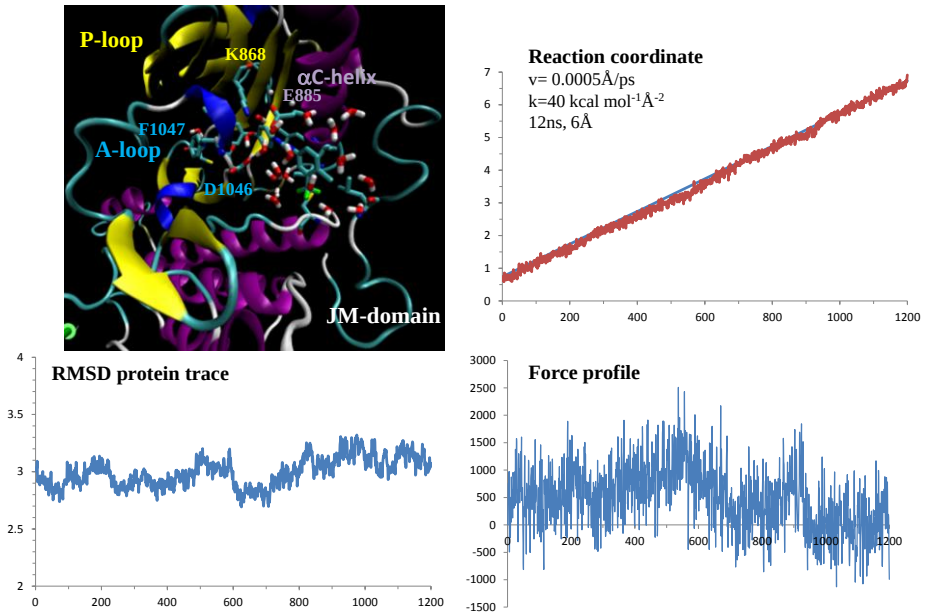


Figure 12. VEGFR2/SOR complex. **Top left:** Frame #1200 sampled after 12ns-long simulation (6Å reaction coordinate displacement). **Top right:** Plot of the end-to-end distance (blue line) vs. the actual value (red line) of the reaction coordinate. **Bottom left**: Plot of the protein RMSD calculated by using protein trace. **Bottom right:** Plot of the force (pN) in the frames sampled.

The simulation performed with VEGFR2/SUN complex under the same conditions did not show any change in the conformation of the enzyme (**Figure 13,** bottom left) and the stiff spring approximation is well satisfied (**Figure 13**, top right). Differently from VEGFR2/SOR complex this ligand is leaving the enzyme cavity from the ATP binding site entrance (**Figure 13**, top left) and no remarkable change in the force profile was noticed along the simulation performed (**Figure 13**, bottom right).

Following these encouraging results, the reaction coordinates was explored up to 18 Å of total displacement (36ns-long simulations).

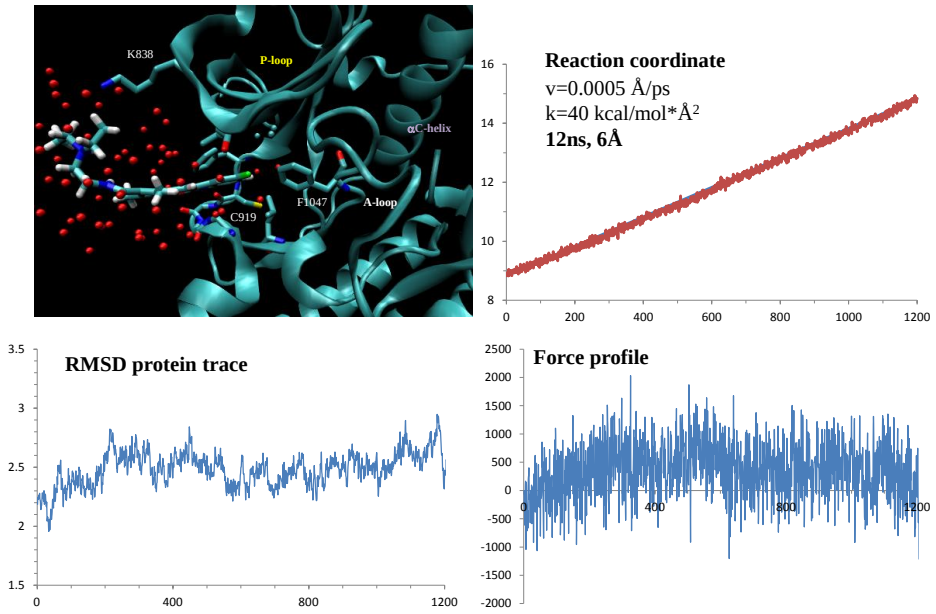


Figure 13. VEGFR2/SUN complex. **Top left:** Frame #1200 sampled after 12ns-long simulation (6Å reaction coordinate displacement). **Top right:** Plot of the end-to-end distance (blue line) vs. the actual value (red line) of the reaction coordinate. **Bottom left**: Plot of the protein RMSD calculated by using protein trace. **Bottom right:** Plot of the force (pN) in the frames sampled.

(iii) PMF along the Unbinding Reaction Coordinates

In **Figure 14** (right) is reported the preliminary PMF of sorafenib calculated by applying the Jarzynski's equality (eq. 3) and the cumulative expansion up to the second order (eq. 4) to the simulations performed with frame#3000 and frame#5000. The end-to-end *vs.* the constrained distance shown in **Figure 14** (left) proves that the stiff-spring approximation is satisfied under our simulation conditions by using a very small pulling velocity v = 0.0005 Å/ps and a force constant of 40 kcal/mol*Å^2. At the same time, the RMSD of the protein trace (**Figure 15**) reveals that the overall conformation of the protein is not dramatically changed by ligand unbinding with respect to that of the X-ray structure.

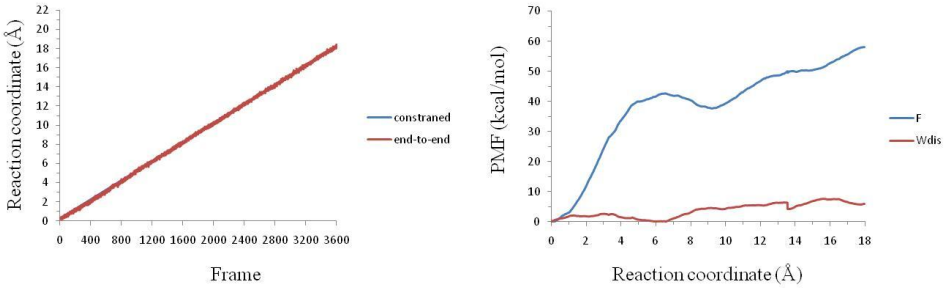


Figure 14. VEGFR2/SOR complex. Left: Plot of the end-to-end (red line) vs. the constrained distance (blue line) measured in each frame sampled in the 36ns long simulation starting from frame #5000 of sorafenib. Similar results were obtained with frame#3000. **Right:** PMF and dissipated work (W_{diss}) of the simulations performed for frame#3000 and frame#5000 of sorafenib.

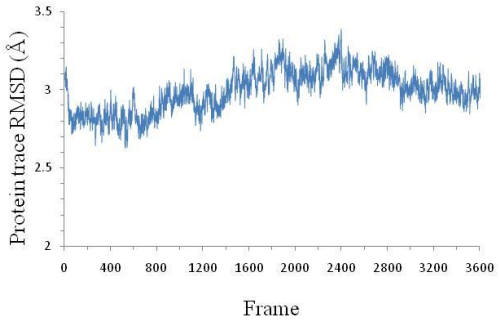


Figure 15. VEGFR2/SOR complex. RMS deviation of protein trace calculated for each frame sampled in the 40ns long simulation of frame#5000 of sorafenib. Similar results were obtained with frame #3000.

In **Figure 16** is reported the plot of the average force profile of sorafenib (bottom) and selected frames taken from the simulation exhibiting the lower work value (frame #3000).

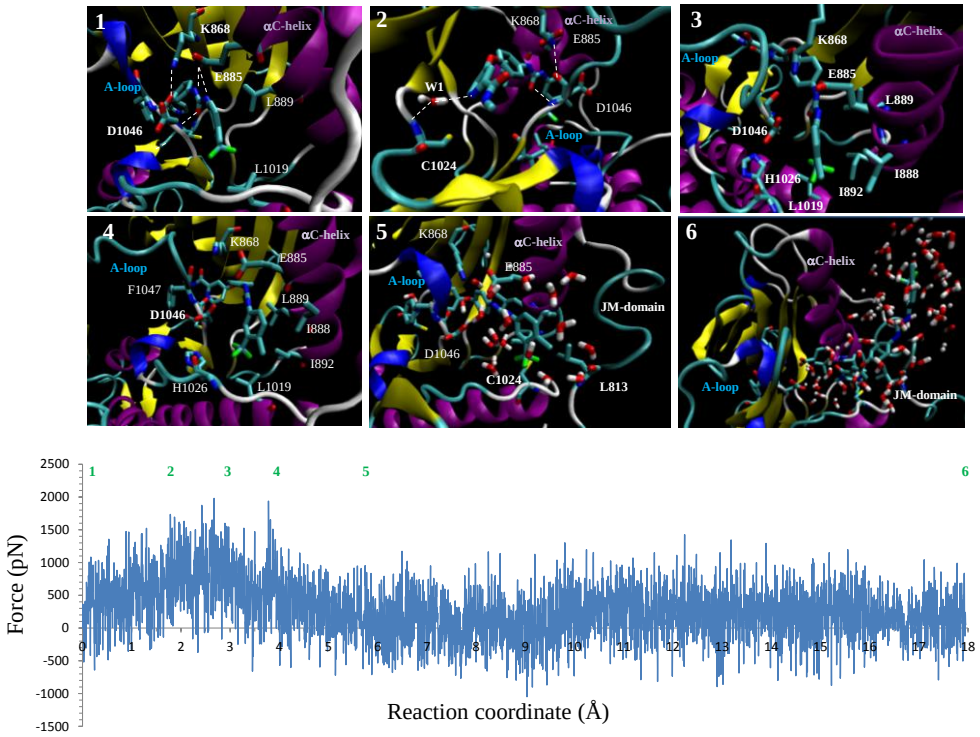


Figure 16. VEGFR2/SOR complex. Top. Representatives frames of the simulation exhibiting the lowest work value (frame #3000). **Bottom.** Average force profile of sorafenib (simulations performed starting from frame#3000 and frame#5000) as a function of the reaction coordinate.

Snapshot #1 was taken at the beginning of the simulation and shows that sorafenib urea moiety makes a HB-interaction with D1046 (DFG-out motif) and two charge-reinforced HB-interactions with E885, involved in a salt bridge with K868, which in turn is interacting with the carbonyl backbone of D1046. Hence, the ligand appears to be tightly bound by this rich network of interactions in a gorge of the binding site. Besides, the mCF_3-pCl-phenyl ring makes hydrophobic interactions with L889 and is projected in a widely-open solvent exposed region of the binding site. Finally, the pyridyl moiety makes a HB-interaction with the hinge residue C919. Between 0 and 2 Å of reaction coordinate displacement, there is an increase in the force value, which is due to the disruption of the direct hydrogen bond interaction between the ligand pyridine fragment and the ligand acetylamide NH with the NH backbone of the

hinge amino acid C919, which is assisted by a water molecule (snapshot #2). Between 2 and 4 Å of the reaction coordinate displacement the hydrogen bond interaction between the carbonyl group of the ligand urea and the NH backbone of D1046 is lost (snapshot #3). Besides, only one hydrogen bond interaction involving the NH urea and E885 is still present. At the same time the mCF_3-pCl phenyl fragment approaches some hydrophobic residues in the αC-helix (I888, I892) while the core ligand phenyl ring is moving close to L889. As a result, the orientation of the αC-helix is changed in order to facilitate ligand exit towards the JM domain (snapshot #4). Between 4 Å and 6 Å of the reaction coordinate displacement, one of the urea NH moiety is still involved in a HB-interaction with E885 and delays ligand exit (snapshot #5). Further to 6Å reaction coordinate displacement, the force profile is relatively flat and the ligand lies outside the ATP binding site, but still close to the JM domain (snapshot #6).

The simulation of frame #3000 of sunitinib described in the calibration paragraph was progressed till 12 Å of the reaction coordinate displacement while that of frame #5000 till 8Å due to limited computational resources. The stiff-spring approximation was well satisfied along all the simulation of frame #3000 (**Figure 17**, left) and the RMSD of the protein did not change significantly (**Figure 17**, right). Similar results were obtained for frame #5000.

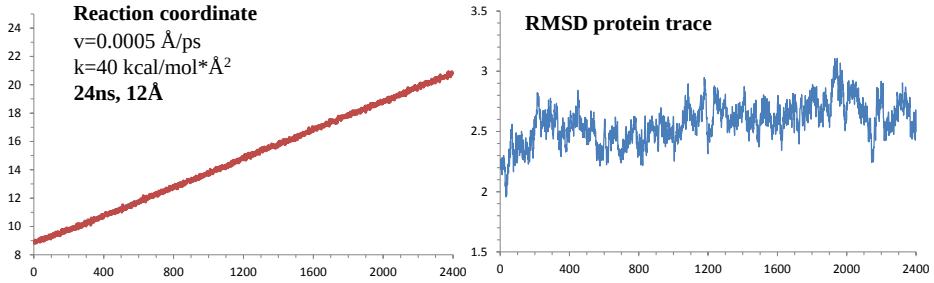


Figure 17. VEGFR2/SUN complex. Left. Plot of the end-to-end (red line) vs. the constrained distance (blue line) measured in each frame sampled in the 24ns long simulation starting from frame #3000 of sunitinib. **Right:** RMS deviation of protein trace calculated for each frame sampled in the 24ns long simulation of frame#3000 of sunitinib.

The preliminary force profile obtained for frame #3000 (**Figure 17,** bottom) suggests that, differently from sorafenib, sunitinib unbinding pathway does not require the disruption of any strong interaction.

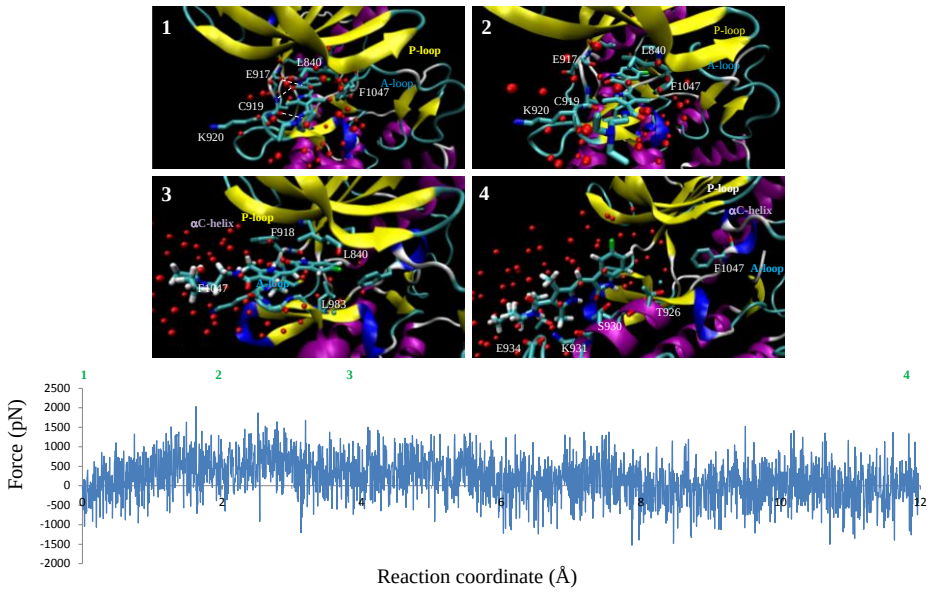


Figure 17. VEGFR2/SUN complex. Top: Representatives frames of the simulation exhibiting the lowest work value (frame #3000). **Bottom:** Plot of the preliminary force profile of sunitinib as a function of the reaction coordinate (simulation performed starting from frame#3000).

At the beginning of the simulation (snapshot #1) sunitinib makes two hydrogen bond interactions with the hinge residue C919 and E917. Besides, the aromatic ring of the ligand makes hydrophobic contacts with F1047 (DFG-out motif) and L840 in the P-loop. Between 0 and 2 Å displacement of the reaction coordinate, the hydrogen bond with E917 is broken while that with the hinge residue become looser and acts as a pivot delaying ligand exit (snapshot #2). At the same time interactions with F1047 become less strong. Between 2Å and 4Å of the reaction coordinate displacement, the hinge hydrogen bond interaction is weaker (snapshot #3) and the ligand is almost entirely exposed to the solvent. Between 4Å and 12Å of the reaction coordinate, the ligand has left the ATP binding site. However, it still makes contacts with some solvent exposed amino acids E934, K931, S930 and T926 in a α-helix (snapshot #4).

A preliminary comparison of sorafenib and sunitinib PMFs is reported in **Figure 18**. Due to limited computational resources available for the simulation of frame #5000 of sunitinib, the comparison has been done till 8 Å of reaction coordinate displacement only.

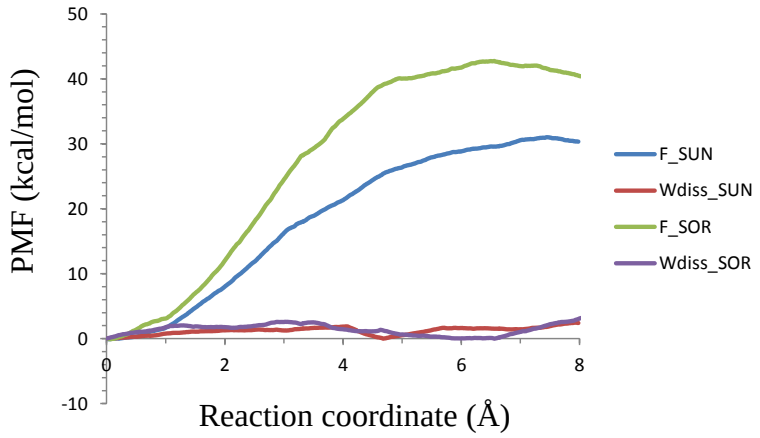


Figure 18. PMF profile of sorafenib (green) and sunitinib (blue) and their dissipated works.

As shown, the PMF of sorafenib (green line) is higher than that of sunitinib (blue line). In particular, after 8 Å of the reaction coordinate displacement, the PMF of sorafenib is approximately 10 kcal/mol higher than that of sunitinib (sorafenib $\Phi_{average}$ = 40.5 kcal/mol *vs.* sunitinib $\Phi_{average}$ = 30.3 kcal/mol). This result is qualitatively in agreement with the experimental k_{off} and residence time values of these derivatives (sorafenib k_{off} = 3.3 *10^{-5} (s^{-1}) and $t_{1/2}$ = 370 min; sunitinib k_{off} > 2.3 *10^{-3} (s^{-1}) and $t_{1/2}$ < 5 min)[4b]. At the same time, their dissipated works is comparable: after 8 Å of reaction coordinate displacement the dissipated work of sorafenib (W_{diss} = 3.02 kcal/mol, purple line) is very similar to that of sunitinib (W_{diss} = 2.41 kcal/mol, red line). These parameters are related to the strength of friction and represent the irreversible work due to non-equilibrium condition. More robust conclusions will be drawn once these simulations as well as those of the other ligand frames will be completed.

Discussion

The SMD simulations of sorafenib and sunitinib were performed without forcing these ligands to follow a predefined direction. By increasing the distance between sorafenib binding site centroid and the ligand center of mass measured in each starting frame, these inhibitors have found their own way to escape. However, it cannot be ruled out that these ligands could potentially follow other unbinding pathways.

Overall the identification of this reaction coordinate following a "trial and error" approach, inspired by the knowledge of the enzyme-ligand structure, proved to be quite difficult and computer resource demanding. The criteria used to discard inappropriate solutions were based on the measurement of protein RMSD trace. Hence, significant changes in the enzyme conformation upon ligand unbinding leading to high RMSD values were judged as artifacts and discarded. However, it cannot be ruled out that ligand unbinding might be associated with significant protein conformational change. To prove that, other techniques such as metadynamics simulations should be used to model simultaneously ligand unbinding and enzyme conformational changes. As an alternative, the use of random approaches such as RAMD, currently not available in Amber software, would have likely facilitated the identification of unbiased exit pathways in a shorter amount of time.

Besides, SMD simulations required the calibration of the pulling force and the force constant as they both strictly depend on the system under study. The selection of the optimal parameter combination, v=0.0005 Å/ps and k=40 kcal/mol*$Å^2$, was the result of a trade-off between the fulfillment of the stiff spring approximation and the computer power as well as simulation timing. In particular, lower pulling velocities could not be utilized due to resource constraints. Even though the force constant used is quite strong, k=40 kcal/mol*$Å^2$ represents the minimum value required for sorafenib to satisfy the stiff spring approximation, likely reflecting its tight binding to this enzyme (sub-nanomolar pIC_{50} value)[4b].

Preliminary results of the SMD simulations performed reveal that sunitinib leaves the ATP binding site through the cavity entrance. Forces involved are low and there are no significant changes in the enzyme conformation upon ligand unbinding. The main ligand-enzyme interaction is represented by the hydrogen bond with the hinge residue, which acts as pivot delaying ligand exit. Unfortunately, only two simulations have been performed for this ligand due to

limited computational resources. On the contrary, sorafenib exits the ATP binding site opposite to the cavity entrance and moves towards the JM domain. The hydrogen bond interaction established with the hinge residue is lost in the first 2 Å displacement of the reaction coordinate. The breakage of such interaction is assisted by a water molecule which bridges the ligand to this amino acid. Besides, the ligand urea moiety makes hydrogen bonds with E885 and D1046 and lies in a narrow region of the binding site where a network of protein-ligand hydrogen bond interactions is established. Hence, it is likely that the presence of this tight network of hydrogen bond interactions and the limited size of this subpocket delay ligand exit and force the αC-helix to change its spatial orientation. Besides, the ligand phenyl ring decorated with hydrophobic EWG substituents establishes lipophilic contact with some hydrophobic amino acids in the αC-helix and moves close to them along unbinding. After 18 Å of reaction coordinate displacement, sorafenib has left the ATP binding site but it lies nearby the JM domain. In light of that, further exploration of the reaction coordinate is required in order to have the ligand fully solvated.

Finally, a preliminary comparison of sorafenib PMF with that of sunitinib showed that the PMF of sorafenib is higher than that of sunitinib. This result is qualitatively in agreement with the experimental k_{off} and residence time values of these derivatives[4b]. At the same time, their dissipated works which represents the irreversible work due to non-equilibrium condition is comparable. More robust conclusions will be drawn once the simulations of all ligand frames will be completed.

Conclusions

VEGFR2 is a clinically validated target in the oncology field. Medicinal chemistry efforts aimed at finding potent inhibitors of this receptor tyrosine kinase resulted in the identification of sunitinib and sorafenib, both on the market to treat Renal Cell Carcinoma. These ligands are characterized by a different selectivity kinase profile, likely determining their distinct clinical profile. Besides, while sunitinib quickly dissociates from VEGFR2, sorafenib exhibits quite a long residence time at this enzyme, which might impact on its duration of action in vivo.

Herein, a preliminary detailed atomistic description of the unbinding process of sorafenib and sunitinib was unveiled by using SMD simulations. While

sunitinib exits the ATP binding site from the cavity entrance, sorafenib moves towards the JM domain, causing a change in the orientation of the αC-helix. Analysis of the force profile reveals that the rupture requires first of all the water-assisted breakage of its interaction with the hinge amino acid and subsequently the disruption of the hydrogen bond of the carbonyl urea moiety with D1046 (DFG-out motif). At the same time the hydrogen bond interactions established between the ligand urea moiety and E885 in the αC-helix likely delay ligand exit. Besides, the formation of hydrophobic interactions of the mCF3,pCl-phenyl ring with hydrophobic residues in αC-helix further slows down the ligand unbinding process. More robust conclusions will be drawn once the simulations of the other ligand frames have been performed, followed by calculation and comparison of their PMFs and dissipated work along all the reaction coordinate.

In light of these results, SMD methodology proved to be a valuable tool to unveil detailed information of the unbinding pathways of both these ligands. However, despite its relatively simple theoretical background, its application to real case studies has highlighted a few critical aspects. First of all, the identification of the reaction coordinate might not be straightforward. The use of a "trial and error" approach, even if inspired by the knowledge of the enzyme-ligand structure, proved to be time consuming and computer resource demanding. The SMD simulations of sorafenib and sunitinib were performed without forcing these ligands to follow a predefined path as no directionality was imposed while the distance of the centroids was evenly increased. However, it cannot be ruled out that these ligands could potentially follow other unbinding pathways. The use of random methods such as RAMD might prove valuable to speed up the identification of unbiased exit pathways.

Another critical aspect is represented by the calibration of the pulling velocity and the force constant. This process proved to be computationally intensive in this case study. The selection of the most appropriate combination of these parameters was basically driven by the fulfillment of the stiff spring approximation as well as by the constraints of computer resource available. In particular, the identification of the appropriate force constant was very critical for the slow-dissociating ligand. The high value identified (k=40 kcal mol^{-1} $Å^{-2}$) was required likely to overcome the strong interactions established by this ligand with VEGFR2.

Despite these critical aspects and the significant amount of computational resources required, the application of SMD to study the unbinding pathways of

these ligands proved to be very informative. The robustness of the results reported here will be improved by increasing the statistics of the calculations.

Abbreviations

VEGFR2: Vascular Endothelium Growth Factor Receptor 2; RAMD: Random Accelerated Molecular Dynamics; SMD: Steered Molecular Dynamics; RMSD: Root Mean Square Deviation

References

(1) Mc Tigue M., Murray B.W., Chen J.H., Deng Y.-L., Solowiej J., Kania R.S. Molecular conformations, interactions, and properties associated with drug efficiency and clinical performance among VEGFR TK inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 18281.

(2) Rini B.I., Escudier B., Tomczak P., Kaprin A., Szczylik C., Hutson T.E., Michaelson D., Gorbunova V.A., Gore M.E, Rusakov I.G., Negrier S., Ou Y.-C., Castellano D., Yeong Lim H., Uemura H., Tarazi J., Cella D., Chen C., Rosbrook B., Kim S., Motzer R.J. Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial *Lancet* **2011**, *378*, 1931.

(3) Miyamoto N., Sakai N., Hirayama T., Miwa K., Oguro Y., Oki H.,Okada K., Takagi T., Iwata H., Awazu Y., Yamasaki S., Takeuchi T., Miki H., Hori A., Imamura S. Discovery of N-[5-({2-[(cyclopropylcarbonyl)amino]imidazo[1,2-b]pyridazin-6-yl}oxy)-2-methylphenyl]-1,3-dimethyl-1H pyrazole-5-carboxamide (TAK-593), a highly potent VEGFR2 kinase inhibitor. *Bioorg. Med. Chem. Lett.* **2013**, *21*, 2333.

(4) a) Iwata H., Imamura S., Hori A., Hixon M.S., Kimura H., Miki H. Biochemical characterization of a novel type-II VEGFR2 kinase inhibitor: Comparison of binding to non-phosphorylated and phosphorylated VEGFR2. *Bioorg.Med.Che.Lett.* **2011**, *19*, 5342. b) Iwata H., Imamura S., Hori A., Hixon M.S., Kimura H., and Miki H. Biochemical Characterization of TAK-593, a Novel VEGFR/PDGFR Inhibitor with a Two-Step Slow Binding Mechanism. *Biochemistry* **2011**, *50*, 738

(5) Schrödinger, Mestro version 9.1 Schrödinger, LLC, New York 2009 http://www.schrodinger.com/

(6) The Amber Molecular Dynamics Package ambermd.org http://ambermd.org/

(7) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.

(8) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.* **2003**, *119*, 3559.

(9) Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946.

(10) Ludemann S.K., Lounnas V., Wade R.C. How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms *J. Mol. Biol.* **2000**, *303*, 797.

Appendix 1

Additional calibration experiments aimed at finding the optimal combination of pulling velocity and force constant.

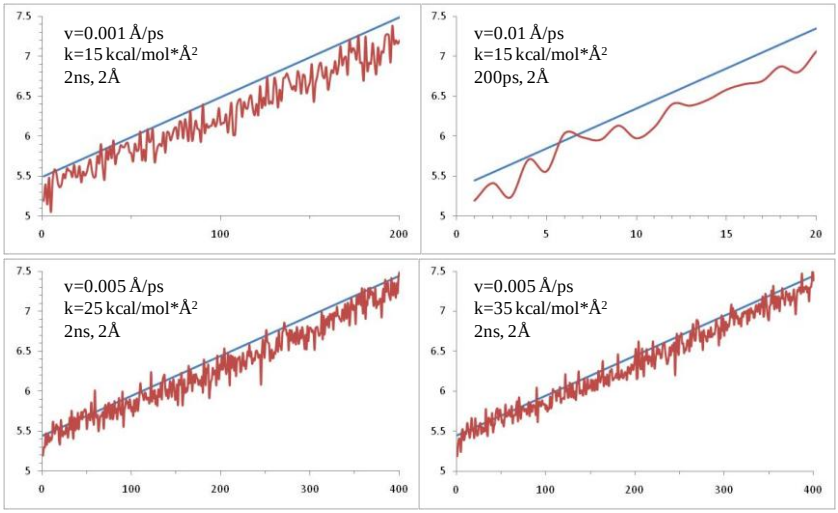


Figure A1_S1. VEGFR2/SOR simulations. Top. Plot of the end-to-end distance vs. the actual value in complex simulations performed with k=15 kcal mol^{-1} $Å^{-2}$ and v= 0.001 (left) and v=0.01 Å/ps (right). **Bottom.** Plot of the end-to-end distance vs. the actual value in the simulations performed with v= 0.0005 Å/ps and k=25 (left) and 35 kcal mol^{-1} $Å^{-2}$ (right).

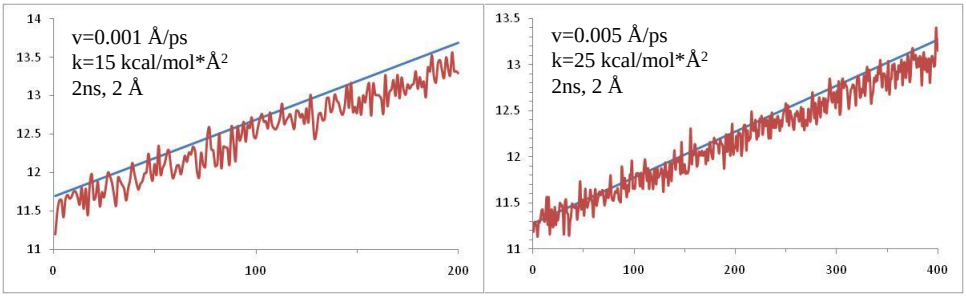


Figure A1_S2. VEGFR2/SUN simulations. Plot of the end-to-end distance vs. the actual value in the simulations performed with k=15 kcal mol^{-1} $Å^{-2}$ and v= 0.001 (left); k=25 kcal mol^{-1} $Å^{-2}$ and v= 0.005 Å/ps (right).

Appendix 2

SMD simulations performed to identify the unbinding reaction coordinate.

2.1 VEGFR2/SUN complex. Reaction coordinate defined as the distance between the ligand center of mass and the Cα of Glu885.

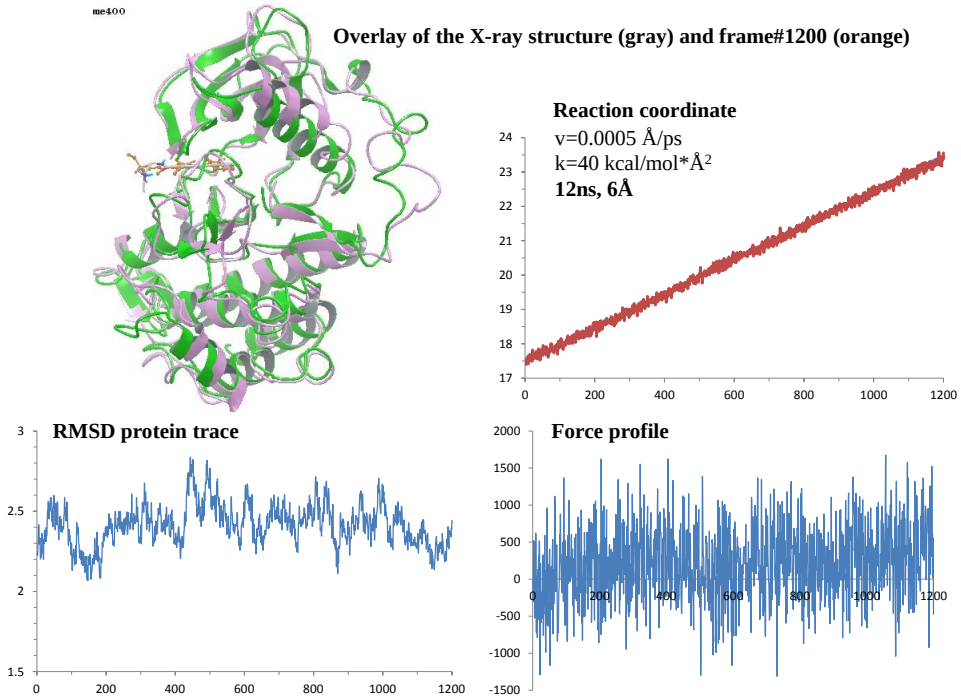


Figure A2_S1. VEGFR2/SUN complex. Plot of the end-to-end distance vs. the actual value of the reaction coordinate (top right), force profile (bottom right), RMSD of protein trace (bottom left). Overlay of frame #1200 (protein atoms in magenta; ligand in orange) on the X-ray structure (protein atoms green, ligand in gray)

2.2 VEGFR2/SUN complex. Reaction coordinate defined as the distance between the ligand center of mass and the centroid of A866-M869, L912-V916 Cα atoms (N-lobe β–sheets).

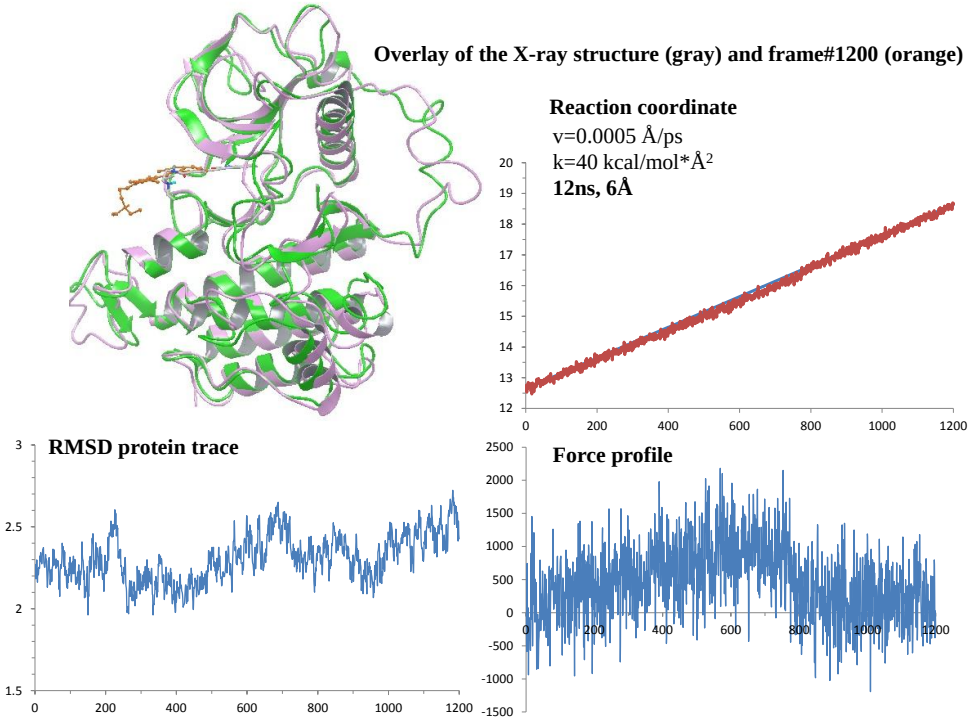


Figure A2_S2. VEGRF2/SOR complex. Plot of the end-to-end distance vs. the actual value of the reaction coordinate (top right), force profile (bottom right), RMSD of protein trace (bottom left). Overlay of frame #1200 (protein atoms in magenta; ligand in orange) on the X-ray structure (protein atoms green, ligand in gray)

2.3 VEGFR2/SOR complex. Reaction coordinate defined as the distance between the ligand center of mass and the centroid of A866-M869, L912-V916 Cα atoms (N-lobe β–sheets).

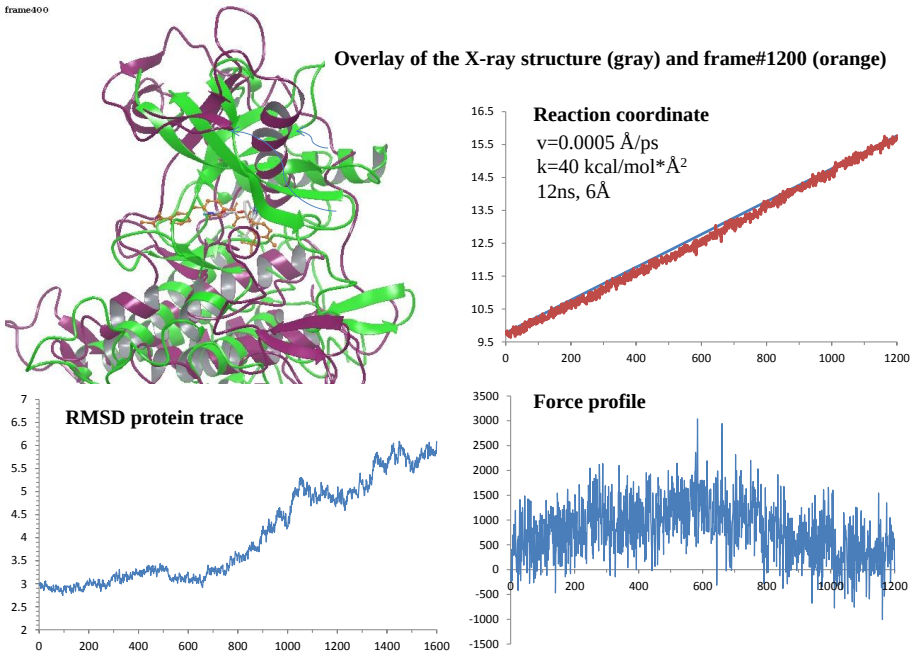


Figure A2_S3. VEGFR2/SOR complex. Plot of the end-to-end distance vs. the actual value of the reaction coordinate (top right), force profile (bottom right), RMSD of protein trace (bottom left). Overlay of frame #1200 (protein atoms in magenta; ligand in orange) on the X-ray structure (protein atoms green, ligand in gray)

2.4 VEGFR2/SOR complex. Reaction coordinate defined as the distance between the ligand center of mass and the centroid of S1009, A1013, F1091, L1095 Cα atoms (C-lobe α–helices).

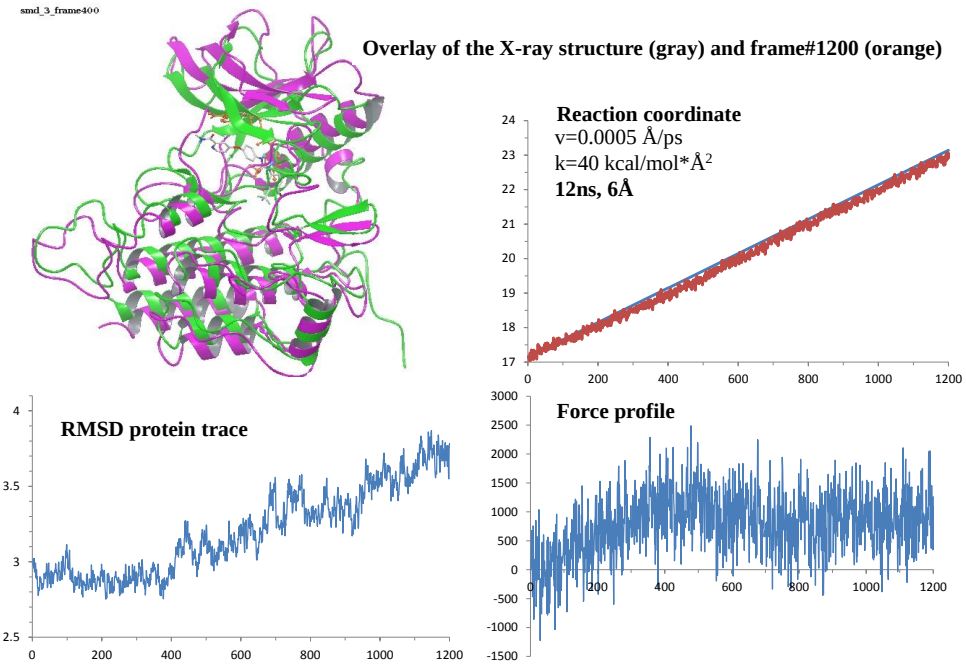


Figure S4. VEGFR2/SOR complex. Plot of the end-to-end distance vs. the actual value of the reaction coordinate (top right), force profile (bottom right), RMSD of protein trace (bottom left). Overlay of frame #1200 (protein atoms in magenta; ligand in orange) on the X-ray structure (protein atoms green, ligand in gray).

2.5 VEGFR2/SOR complex. Reaction coordinate defined as the distance between the ligand center of mass and the centroid of the kinase.

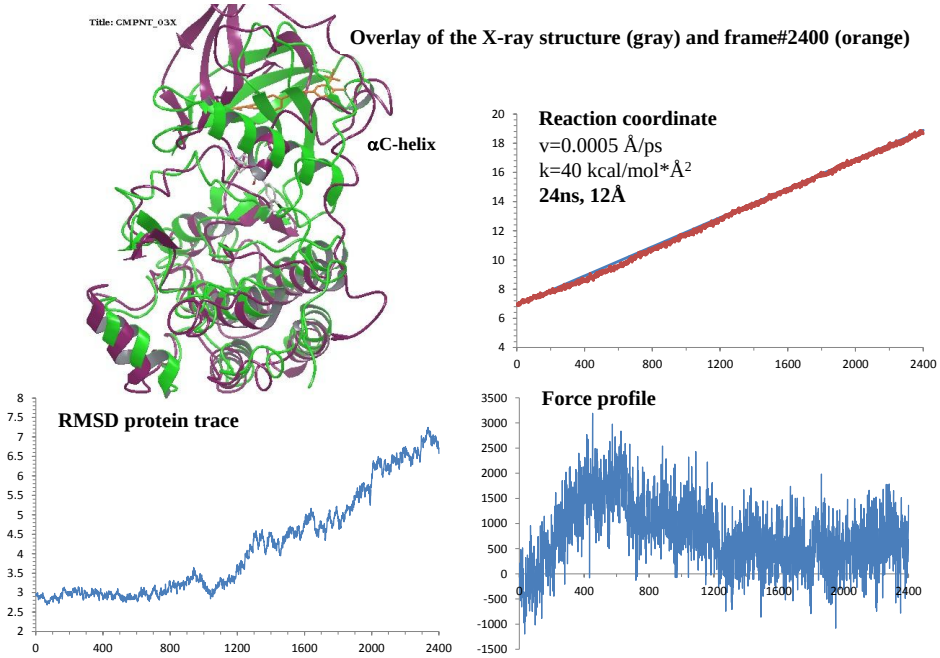


Figure A2_S5. VEGFR2/SOR complex. Plot of the end-to-end distance vs. the actual value of the reaction coordinate (top right), force profile (bottom right), RMSD of protein trace (bottom left). Overlay of frame #2400 (protein atoms in magenta; ligand in orange) on the X-ray structure (protein atoms green, ligand in gray)

2.6 VEGFR2/SUN complex. Reaction coordinate defined as the distance between the ligand center of mass and the centroid of the kinase.

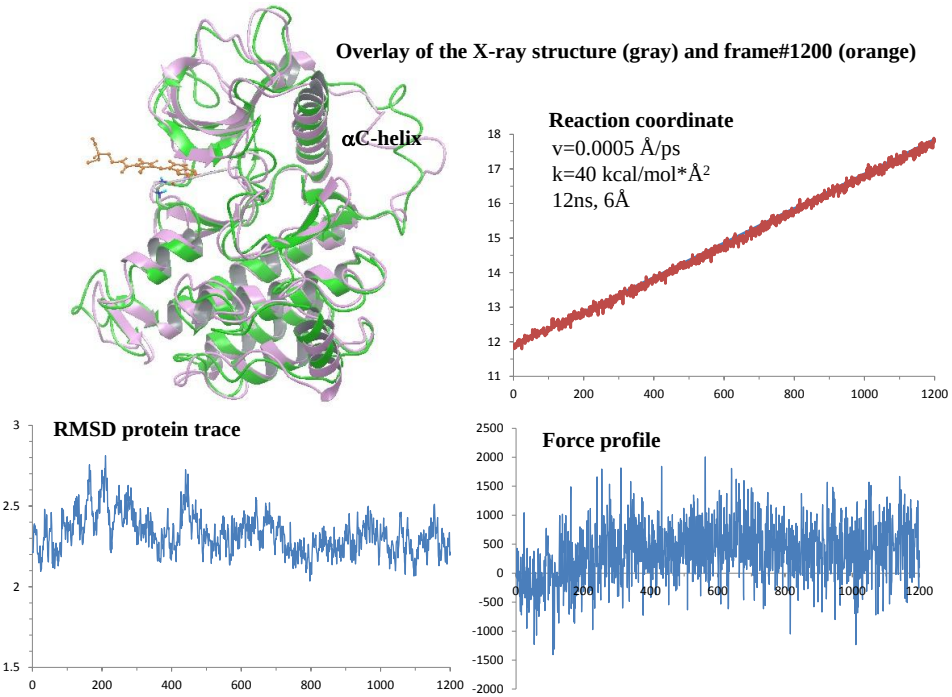


Figure A2_S6. VEGFR2/SUN complex. Plot of the end-to-end distance vs. the actual value of the reaction coordinate (top right), force profile (bottom right), RMSD of protein trace (bottom left). Overlay of frame #2400 (protein atoms in magenta; ligand in orange) on the X-ray structure (protein atoms green, ligand in gray).

CONCLUSIONS

Nowadays several computational chemistry techniques are routinely used to effectively support all phases of the drug discovery process and to accelerate the identification of novel chemical entities to address unmet medical needs. In this PhD thesis, several methodologies have been utilized to cope with specific discovery needs.

In the 1st chapter dealing with bitter taste receptors, chemoinformatics techniques have been extensively used to identify relationships between structural and pharmacophoric features of bitter tastants with their biological profiles. Besides, homology models of three receptor subtypes most abundantly expressed in human airways smooth muscle cells have been built and validated in light of the few experimental data available to elucidate potential binding modes of tastants. As the majority of bitter taste activators are weakly binders of their receptors, the results of these chemoinformatics and modeling studies could be exploited to build ligand focused sets for screening aiming at finding novel potent and selective activators.

In the chapter #2 dealing with modeling studies of glucocorticoid receptor and its ligands, steered molecular dynamics simulations have unveiled the unbinding pathways of prototypical steroid agonists both from the native and a clinically–relevant mutated glucocorticoid receptor. Besides, a robust validated simulation protocol, qualitatively relating ligands PMFs with their k_{off} values, can be utilized to prioritize the synthesis of novel steroid template GR agonists.

Similarly, in chapter #3 steered molecular dynamics simulations has been used to elucidate the unbinding pathways of two clinically relevant ATP-competitive VEGFR2 inhibitors (sorafenib and sunitinib) from this tyrosine kinase. The preliminary results obtained showed that the fast-dissociating ligand sunitinib leaves the ATP binding site from the binding site entrance. On the contrary, the prototypical slow-dissociating inhibitor sorafenib exits the cavity by moving opposite to the binding site entrance and approaches the JM domain. Its unbinding pathway is associated with a change in the conformational orientation of the flexible αC-helix. More simulations are required to improve the robustness of these preliminary conclusions.

Steered molecular dynamics is a methodology which belongs to the family of enhanced sampling techniques that are concerned with dynamics behaviors of ligand –protein interactions. It is well recognized that binding of a compound to

a target protein is a dynamic process involving molecular recognition and complex stability measured as the residence time of the drug molecule on its target. Increase in the residence time of a ligand to its receptor in vitro might expand its duration of action in vivo and result in a lower dose usage. Hence, the availability of a robust computational method able to successfully rank compounds according to their PMFs would be highly desirable to accelerate the lead optimization process. Historically, SMD had been implemented to model protein folding, ligand unbinding as well as movement of ions through channels embedded in membranes. Recently, application of this technique to address drug design issues has been described in the literature. It should be recognized that the widespread usage of this technique observed in recent years is related to the availability of powerful computational engines (GPUs) and infrastructure such as that provided by CINECA Consortium. As these case studies have highlighted, this technique provided detailed information about the unbinding pathways of clinically relevant ligands for their targets, which could be exploited to design novel drugs.

However, at the same time, SMD experiments reported in this thesis have also highlighted some critical aspects of this methodology. First of all, the identification of a potential unbinding reaction coordinate was relatively easy for GR ligands following crystal structure visual inspection, but was not at all straightforward for VEGFR2 ones. The use of a "trial and error" approach, even if inspired by the knowledge of the enzyme-ligand structure, proved to be time consuming and computer resource demanding for VEGFR2. The use of random methods such as RAMD might represent a valuable alternative approach to speed up the identification of unbiased exit pathways. Therefore, the availability of automated tools able to accomplish this task would be highly desirable.

Another critical aspect of SMD is represented by the calibration of the pulling velocity and the force constant. This process has to be accomplished prior to any SMD simulation as they strictly depend on the system under study. This task proved to be computationally intensive both for VEGFR2 and GR. The selection of the most appropriate combination of these parameters was basically the result of a trade-off between the fulfillment of the stiff spring approximation and the computer resources available. In particular, the identification of the most appropriate force constant was very critical for the slow-dissociating ligand of VEGFR2 sorafenib and at the end of the day the high value identified ($k=40$ kcal mol^{-1} $Å^{-2}$) was the minimum required for this tightly-bound ligand. Hence, the availability of automated routines to accomplish the combined

optimization of the force constant and the pulling velocity would be highly desirable to speed up the calibration process. At the same time, conformational changes of the protein under study should be carefully monitored.

Despite these critical aspects and the huge amount of computational resources required to achieve statistically significant results, the application of SMD to GR and VEGFR2 cases proved to be very informative for drug design purposes. Considering the atomistic details unveiled by this technique, it is likely that SMD as well as other enhanced sampling methods will become more popular in the computational chemistry community in the years to come.

ACKNOWLEDGEMENTS

I would like to thank Chiesi Farmaceutici, in particular Dr. Maurizio Delcanale, for supporting this PhD study, CINECA Consortium for Computational resources and support, Prof. Andrea Cavalli and Dr. Agostino Bruno for helpful discussions.

POSTER GORDON CONFERENCE 2011

Chiesi
People and ideas for innovation in healthcare

Elucidation of bitter taste receptor agonists binding modes

[1]Anna Maria Capelli, [2]Gabriele Costantino

[1]Computational Chemistry, Chiesi Farmaceutici S.p.A., via S. Leonardo 96, Parma, Italy;
[2]Dipartimento Farmaceutico, via G.P. Usberti 27/A, Universita' degli Studi di Parma, Parma Italy

Introduction

In mammalians the sense of taste is mediated by receptors located in the oral cavity. Each of the five taste qualities (sour, salty, sweet , umami and bitter) recognizes specific food components in their diet (Lindermann 1996).
A family of genes has been identified and demonstrated to encode bitter taste receptors TAS2Rs (or T2Rs) (Adler et al. 2000, Chandrashekar et al. 2000, Matsunami et al. 2000). Vertebrate species differ in the number of TAS2Rs genes, ranging from three and four in chicken and Fugu fish, approximately 25 and 35 members in humans and mice, to 49 in frog (Shi et al. 2006).
The human bitter taste receptor gene family belong to the Frizzled/Taste2 subfamily of the G-protein coupled receptors (GPCR) (Kyle 2007, Lagerstroem 2008). They are expressed on the tongue in bitter taste receptor cells co-expressing specific signal transduction components like G_α gustducin and able to detect stimuli of only one taste quality (Chandrasheker et al. 2006).
Recently these receptors have been identified in isolated human airway smooth muscle cells (ASM) which also express the cognate G protein gustducin. TAS2R agonists such as saccharin, chloroquine and denatonium evoked increased intracellular calcium in ASM eliciting relaxation of isolated ASM suggesting that this novel pathway might be exploited to identify novel compounds to treat asthma and chronic obstructive pulmonary disease (COPD).
Cognate bitter compounds have been identified for 20 out of 25 human TAS2Rs. Recently, the screening of a diverse set of approximately 100 bitter taste substances in a recombinant system expressing 25 hTAS2Rs (Meyerhof et al. 2010) provided valuable information about structural and pharmacophoric requirements to bind these receptors.
Here we report the development of receptor models of human TAS2R10, TAS2R14 and TAS2R31, the most abundantly bitter taste receptors expressed in ASM, which were used to carry out docking studies with strychnine, flufenamic acid and aristolochic acid, the most potent and selective ligands for these receptor subtypes, respectively (**Figure 1**). Results of the docking studies will be presented in detail and discussed in light of available site-directed mutagenesis experiments (Brockhoff et al. 2010) and literature SAR data.

Methods

Human bitter taste receptor sequences were retrieved from UniProtKB ( ) and aligned with the use of BLASTP 2.2.25 (Altschul et al. 1997). According to the new numbering system, hTAS2R31 corresponds to hTAS2R44 described in Meyerhof et al. 2010. The biological fingerprint of each hTAS2R was built by using potency data generated for 104 bitter tastants (Meyerhof et al. Chem. Senses 2010; Chandrasheker et al. 2000, Bufe et al. 2002, Kim et al. 2003, Behrens et al. 2004, 2009; Kuhn et al. 2004, Pronin et al. 2004, 2007; Brockhoff et al. 2007; Sainz et al. 2007, Maehashi et al. 2008), which were converted into a bit string of 1 and 0 depending on the ligand response at each receptor i.e. 1 = activation, 0 = no activation. Therefore, each human receptor was described by a 104-bit string. Then, the biological similarity of each hTAS2R with respect to the others was calculated with the use of Tanimoto index and the biological fingerprints described above. Tanimoto similarity index is c/(a+b-c) where a is the number of bits on in receptor A, b is the number of bits on in receptor B, c is the number of bits in common between A and B. Two receptors exhibiting the same profile have Tanimoto = 1, while if nothing is in common, Tanimoto = 0. The structures of 97 compounds out of the 104 described by Meyerhof et al. 2010 were retrieved from PubChem database ( ). Bitter tastants cluster analysis was performed with LibraryMCS cluster algorithm and ChemAxon 1024 Chemical Fingerprints (ChemAxon 5.4). Human receptor models of hTAS2R10, hTAS2R14 and hTAS2R31 were built within MOE (CCG) by using human β2R as template (3P0G PDB code). Prior to run docking experiments, receptor models were optimized with the use of Protein preparation workflow available within Maestro (Schrodinger). Binding sites were identified with the use of Sitemap (Schrodinger) and utilized as starting points for Induced Fit Docking simulations (Schrodinger) to dock strychnine, flufenamic acid and aristolochic acid in hTAS2R10, hTAS2R14 and hTAS2R31 (**Figure 1**). Poses were visually inspected within Maestro. Finally, strychnine and flufenamic acid were docked in hTAS2R10 and hTAS2R14 after removal of their extracellular loops and with the use of Induced Fit docking (default parameters) for comparison. The center of the box was defined as the centroid of F107 and S274 for hTAS2R10; F108 and M270 for hTAS2R14.

Results and discussion

According to the sequence alignments obtained, TAS2Rs exhibit overall low identity among each other (**Figure 2**, top) as well as to GPCRs whose x-ray structures have been reported in the literature (**Figure 2**, bottom).

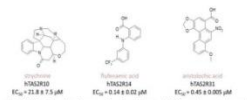
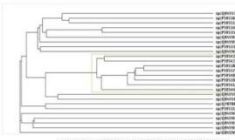


FIGURE 1. Structure of strychnine, flufenamic acid and aristolochic acid. EC_{50} were taken from Meyerhof et al. 2010.



% identity	T2R10_human	T2R14_human	T2R31_human
spad_bovin	23.69	30.98	31.25
adrb2_human	22.22	30	33.33
adrb1_meiga	18.2	30.36	21
cxcr4_human	23.53	22.92	21.38
drd1_human	18.37	37.5	21.65
axbr_human	18.54	30.07	29.07

FIGURE 2. Relationship of human TAS2Rs depicted in a dendrogram (above right) and identity with GPCRs for which x-ray structures are available (below right). Sequence identity was calculated with GPCRDB (12.01.2011) alignment engine available at . Accession dates: 28/02/2011 and 07/06/2011.

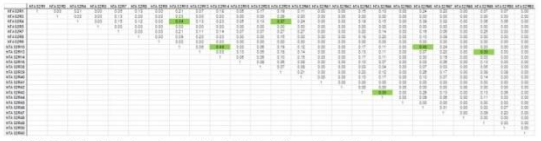
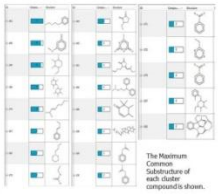


FIGURE 3. Similarity of bitter taste receptors calculated with their biological fingerprints and Tanimoto index.

Even though hTAS2R10 and hTAS2R14 exhibit low sequence identity, they show a certain degree of similarity calculated with the use of biological fingerprints and Tanimoto index, which reflects the fact that they share several common activators (15 out of the 97 agonists considered). On the contrary, hTAS2R31 (hTAS2R44) exhibits low biological fingerprints similarity with respect to hTAS2R10 and hTAS2R14 as well as to the other subtypes (**Figure 3**).



The Maximum Common Substructure of each cluster compounds shown.

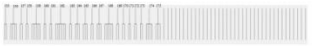


FIGURE 4. Cluster analysis of 97 bitter tastants described in Meyerhof et al. 2010. 21 clusters and 37 singletons were obtained.

According to the cluster analysis performed (**Figure 4**), bitter tastants look structurally different among each other suggesting that they can exhibit different binding modes. The 15 hTAS2R10 and hTAS2R14 common activators out of the 97 agonists considered are spread over 7 clusters and 6 singletons. At the same time these receptors can discriminate even among very similar bitter compounds. For example, cucurbitacin E and B belong to the same cluster (**Figure 4**, #158). While both these ligands are agonists of hTAS2R10, hTAS2R14 is activated by cucurbitacin B only.

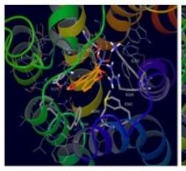


FIGURE 5. Aristolochic acid bound to hTAS2R31.

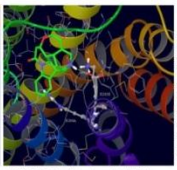


FIGURE 6. Strychnine bound to K265E and R268A hTAS2R31.

In **Figure 5** a docking solution of aristolochic acid bound to hTAS2R31 is shown. This result is in agreement with site directed mutagenesis studies reported in the literature (Brockhoff et al. 2010).

Strychnine is not able to activate native hTAS2R31 up to 0.03 mM concentration (Meyerhof et al. 2010). However, when K265 and R268 were mutated to glutamic acid and alanine, respectively, strychnine was able to activate this mutant at 30 μM (Brockhoff et al. 2010). In **Figure 6** strychnine bound to K265E R268A hTAS2R31 is shown.

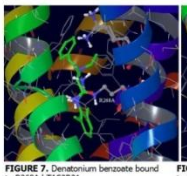


FIGURE 7. Denatonium benzoate bound to R268A hTAS2R31.

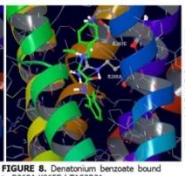


FIGURE 8. Denatonium benzoate bound to R268A K265E hTAS2R31.

Denatonium benzoate is not able to activate native hTAS2R31 up to 3 mM concentration (Meyerhof et al. 2010). However, when R268 was mutated to alanine, this agonist was able to activate this mutated receptor at 1 mM concentration (Brockhoff et al. 2010). When both R268 and K265 were mutated to alanine and glutamic acid, respectively, denatonium benzoate was still able to activate the receptor. In **Figure 7** a docking solution of denatonium bound to K265E hTAS2R31 is shown while in **Figure 8** is reported a pose of this ligand in the double mutated receptor.

Docking studies of strychnine in the model of hTAS2R10 suggest that this activator can either lie close to E16 (TM1) or make a hydrogen bond with S189 (TM2) (**Figure 9**). The latter binding mode projects this ligand close to W271 (TM7), which corresponds to K265 in hTAS2R31 and was also identified with the simulations performed with hTAS2R10 devoid of its extracellular loops. It is tentative to speculate that a common agonist binding site might be present for hTAS2R10 and hTAS2R31.
Docking studies of flufenamic acid in the model of hTAS2R14 show that the carboxylic group of this ligand can either bind R55 (TM1) or make a hydrogen bond interaction with T105 (TM3). The latter binding mode was also identified by performing docking experiments without the extracellular loops (**Figure 10**). Site-directed mutagenesis experiments should be undertaken to support these results.

FIGURE 9. Potential binding modes of strychnine in hTAS2R10.

FIGURE 10. A docking solution of flufenamic acid bound to hTAS2R14 after the removal of the ECLs

Conclusions and future perspective

Overall human bitter taste receptors exhibit low sequence identify among each other as well as for GPCRs whose X-ray structures are available. hTAS2R10 and hTAS2R14 show a certain degree of biological similarity as they share several common activators. However, cluster analysis has showed that their common tastants look structurally different, hence it is likely that they exhibit different binding modes at these receptors.
Docking studies of representative bitter tastants in hTAS2T10, hTAS2R14 and hTAS2R31 subtypes provide insights in their potential binding modes. While poses of aristolochic acid in complex with hTAS2R31 are consistent with site-directed mutagenesis studies reported in the literature, results of strychnine and flufenamic acid simulations in hTAS2R10 and hTAS2R14 models, respectively, need validation either by performing SDM experiments or indirectly by carrying out SAR expansion of these tool compounds. Data coming out from these studies would complement the information being generated on the function, genetics and cellular expression of these receptors.

References

Adler E., Hoon M.A, Mueller K.L., *Cell* **2000**, *100*, 693-702
Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. *Nucleic Acids Res.* **1997**, *25*, 3389-3402
Brockhoff A., Behrens M., Niv M.Y., Meyerhof W. *PNAS* **2010**, *107*, 11110-11115
Chandrashekar J., Mueller K.L., Hoon M.A., Adler E., Feng L., Guo W., Zuker C.S. *Cell* **2000**, *100*, 703-711
Chandrasheker J., Hoon M.A., Ryba N.J., Zuker C.S. *Nature* **2006**, *444*, 288
Deshpande D., Wang W. C. H., McIlmoyle E. L., Robinett K. S., Schillinger R. M., An S. S., Sham J. S. K., Liggett S. B. *Nature Medicine* **2010**, *16*, 1299-1304
Kyle Palmer R. *Molecular Interventions* **2007**, *7*, 87-98
Lagerstroem M.C., Schiöth H. B. *Nature Reviews Drug Discovery* **2008**, *7*, 339-358
Lindemann B. *Physiol Rev.* **1996**, *76*, 718-766
Meyerhof W., Batram C., Kuhn C., Brockhoff A., Chudoba E., Bufe B., Appendino G., Behrens M. *Chem Senses* **2010**, *35*, 157-170
Matsunami H., Montmayeur J.P., Buck L.B. *Nature* **2000**, *404*, 601

MOE 2010.12 Chemical Computing Group 1010 Sherbrooke St. W, Suite 910 Montreal, Quebec, Canada H1A 2R7
Shi P., Zhang J. *Mol.Biol.Evol.* **2006**, *23*, 292
Schrodinger LLC. 120 West 45th Street 17th Floor, Tower 45 New York, NY 10036-4041 Suite 2010

POSTER ACS 2012

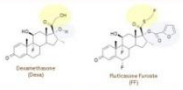
INSIGHT ON THE BINDING MODE OF GLUCOCORTICOID LIGANDS

A.M.Capelli[1,2], A.Bruno[2], G.Costantino[2]

[1]Chemistry Research and Drug Design Department, Chiesi Farmaceutici S.p.A., Largo F. Belloli, Parma, Italy
[2]Dipartimento Farmaceutico, Universita' degli Studi di Parma, via G. Usberti, Parma, Italy

Introduction

Glucocorticoids represent the main therapy to treat several chronic and acute inflammatory conditions like asthma, MS, RA, IBD. The biological action of glucocorticoids is mediated by the glucocorticoid nuclear receptor (GR). Upon ligand binding, GR translocates into the nucleus and binds to DNA at glucocorticoid response elements (GRE) in the promoter region of steroid-responsive genes. In addition, GR may also negatively modulate inflammatory responses through inhibitory interactions with pro-inflammatory transcription factors such as Activator protein (AP)-1 and Nuclear factor (NF)-kB [1].
A considerable amount of research has been done to increase the therapeutic index of corticosteroids by increasing their potency, duration of action and reduction of their systemic side effects. Fluticasone furoate (FF) is a topically active glucocorticosteroid [2], which shows higher receptor affinity and potency as well as longer duration of action with respect to other GR agonists such as dexamethasone (Dexa) and fluticasone propionate (FP) [3], a structural analogue of FF. Furthermore, FF is characterized by rapid receptor association with a slow dissociation from GR compared to Dexa and FP [4].



Objective of the study

In order to rationalize the duration of action of the current GR agonists significant efforts have been aimed at gathering detailed information of their binding and unbinding mode. Here we report our attempts to model the undocking process of Dexa and FF bound to GR with the use of Steered Molecular Dynamics (SMD) simulations and to relate these results to their binding kinetics.

Materials and methods

Construction of GR complexes. The 3D coordinates of Dexa bound to the ligand binding domain (LBD) of human GR and co-activator TIF2 were retrieved from the Protein Data Bank (pdb code: 1M2Z) [5]. Then, chain A (GR LBD) and chain B (TIF2) (monomeric receptor complex) were used in the subsequent simulations. Dexa was modelled starting from its crystallographic coorrdinates. Similarly, the GR LBD with bound FF were retrieved from the Protein Data Bank (pdb code: 3CLD) [2]. To introduce FF within the GR LBD crystallized with Dexa, the chain A of GR-FF/TIF2 X-ray structure was superimposed with that of GR-Dexa/TIF2 complex using Maestro (Schrodinger v9.1). Then the crystallized Dexa was replaced with FF.

Protein and ligand preparation. The monomeric receptor was prepared with the use of Leap (Amber v11). The geometry of Dexa and FF were optimized using ab initio QM (Jaguar, Schrodinger v9.1) and RESP charges were calculated. Ligand parameters were assigned with the general Amber force field (GAFF). The protein-ligand complexes were solvated with a cubic box (12 Å) of TIP3P water molecules using Leap. Na^+ and Cl^- were added to neutralize the net charge of the system and to mimic the physiological concentration (0.15M).

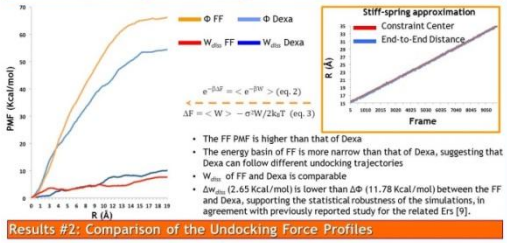
Steered Molecular Dynamics. All MD simulations were performed with AMBER v11. Both GR-Dexa/TIF2 and GR-FF/TIF2 systems were gradually minimized and then equilibrated at 310K for 5ns. The time length of each system simulation was 20ns and 1000 frames were collected. Along the last 10ns of equilibration (500 frames) frame #1, #125, #250, #375 and #500 were taken and used for constant-velocity SMD (20ns per frame).

Potential of Mean Force Calculation (Φ). The stiff spring approximation (eq. 1)[6] was satisfied with a spring constant of 15 kcal mol^{-1}Å^{-2}. During SMD the distance between the center of mass of the ligand and CA of R611 was gradually increased starting from 15 Å till 35 Å with a pulling velocity of 0.001 Å/ps. . PMF along the unbinding reaction coordinate was reconstructed from the SMD trajectories described above and using Jarzynski's equality (eq. 2)[7]. The cumulant expansion approach up to second order (eq. 3) [8] was used for reconstructing the PMF.

$$F(\lambda) \cong \Phi(\lambda) \quad (eq.\ 1) \qquad e^{-\beta\Delta F} = < e^{-\beta W} > \ (eq.\ 2)$$

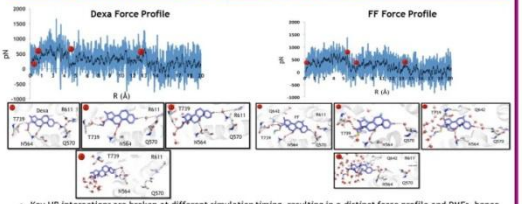
$$\Delta F = < W > -\sigma^2 W/2k_BT \quad (eq.\ 3) \qquad \sigma^2 W = <W^2> - <W>^2 \qquad k_BT = 0.6186$$

Results #1: PMF along the Unbinding Reaction Coordinate



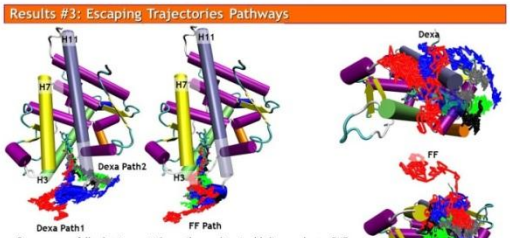
- The FF PMF is higher than that of Dexa
- The energy basin of FF is more narrow than that of Dexa, suggesting that Dexa can follow different undocking trajectories
- W_{diss} of FF and Dexa is comparable
- ΔW_{diss} (2.65 Kcal/mol) is lower than ΔΦ (11.78 Kcal/mol) between the FF and Dexa, supporting the statistical robustness of the simulations, in agreement with previously reported study for the related Ers [9].

Results #2: Comparison of the Undocking Force Profiles



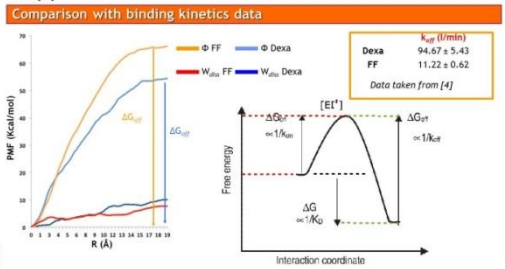
- Key HB interactions are broken at different simulation timing, resulting in a distinct force profile and PMFs, hence undocking rate
- The lipophilic furoate chain strongly affects the escaping trajectories of FF, causing a delay in the breaking of the key HB between R611, Q570 and the carbonyl oxygen at C3, N564 and the -OH group at C11.

Results #3: Escaping Trajectories Pathways



- Dexa escapes following two putative paths consistent with its own larger PMF basin (Results #1)
- FF escapes following a unique putative path, explaining the more narrow basin of its PMF
- The observed different behaviour of the drugs can be related to the presence of the bulkier furoate group in FF, in agreement with the reported mutagenesis data [10]

Comparison with binding kinetics data



	k_{off} (l/min)
Dexa	94.67 ± 5.43
FF	11.22 ± 0.62

Data taken from [4]

- The FF PMF is higher than that of Dexa, consistent with the different k_{off} values
- The applied protocol is able to qualitatively reproduce the differences in the experimental values of the dissociation rate constant

Conclusions

- Applying SMD and Jarzynski's equality we were able to compute the PMF for the Dexa and FF escaping route, with a significant statistical robustness
- Dexa and FF follow slightly different undocking pathways, which is reflected in the differtent shape of their PMF
- The observed PMF differences can be related to the presence of the bulkier furoate group, acting as an anchor and leading to the selection of a unique escaping trajectory
- Further experiments by using different glucocorticoids (e.g. Fluticasone Propionate, Mometasone Furoate) are needed to strenghten the ability of the protocol to discriminate among compounds with different k_{off}

Acknowledgements

The authors would like to thank the CINECA Consortium for the Computational resources and support, and Prof. Antonio Entrena Guadix of the Universidad de Granada for the precious support provided.

References. [1] Barnes, P.J. *J.Steroid Biochemistry & Mol.Biol.* 2010, *120*, 76. [2] Biggadike K. et al. *J.Med.Chem.* 2008, *51*, 3349. [3] Rossios C. et al. *European Journal of Pharmacology* 2011, 670, 244. [4] Valotis and Hogger, *Respiratory Research* 2007, *8*, 1465. [5] Bledsoe R.K. et al. *Cell* 2003, *110*, 93. [6] Jarzynski C. *Phys.Chem.Rev.* 1997, *78*, 2690. [7] Cavalli A. et al. *J.Am.Chem.Soc.* 2010, *132*, 7361. [8] Park S., Khalili-Araghi *J.Chem.Phys.* 2003, *119*, 3559. [9] Shen J., et al. *J. Phys. Chem. B*, 2009, *113*, 10436. [10] Bray P.J. and Cotton R.G.H. *Human Mutation*, 2003, *21*, 557.

Chiesi
People and ideas for innovation in healthcare