MedChemComm

Accepted Manuscript

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

An overview of recent Molecular Dynamics applications as medicinal chemistry tool for undruggable sites challenge.

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Ugo Perricone,*[a†] Maria Rita Gulotta [ab†], Jessica Lombino[ab], Barbara Parrino[b], Stella Cascioferro[b], Patrizia Diana[b], Girolamo Cirrincione[b] and Alessandro Padova [a]

Molecular Dynamics (MD) has become increasingly popular due to the development of hardware and software solutions and improvement in algorithms, that allowed researchers to scale up calculations in order to speed up them. MD simulations are usually used to address protein folding issues or protein-ligand complex stability through energy profile analysis over time. In recent years, the development of new tools able to deeply explore Potential Energy Surface (PES) allowed researchers to focus on the dynamic nature of binding recognition process and binding-induced protein conformational change. Moreover, modern approaches have demonstrated to be effective and reliable in calculating some kinetic and thermodynamic parameters behind the host-guest recognition process. Starting from all of these considerations, several efforts have been made in order to integrate MD within the virtual screening process in drug discovery. Knowledge retrieved from MD can be, in fact, exploited as a starting point to build pharmacophores or docking constraints in the early stage of the screening campaign as well as to define key features, in order to unravel hidden binding modes and help the optimisation of the molecular structure of a lead compound. Based on these outcomes, researchers are nowadays using MD as an invaluable tool to discover and target previously considered undruggable binding sites, including protein-protein interactions and allosteric sites on protein surface. As a matter of fact, the use of MD has been recognised as vital in the discovery of selective protein-protein interaction modulators. The use of a dynamic overview on how the host-guest recognition occurs and of the relative conformational modifications induced, allow researchers to optimise small molecules and small peptides capable to tightly interact within the cleft between the two proteins.

In this review we point to present the most recent applications of MD as integrated tool to be used in the rational design of small molecules or small peptides able to modulate undruggable targets, such as allosteric sites and protein-protein interactions.

Introduction

The rational design of new selective chemical entities represents without any doubt the most important issue in Medicinal Chemistry, often referred to as rational drug design or simply rational design. *In silico* rational drug design is based on an early hypothesis of a desired effect due to the modulation of a specific biological target, structurally complementary to the designed molecule. This process is then characterized by the search of molecules that are directed to a specific target whose biological role within the pathology is known[1]. The main goal of a drug discovery process is the identification of a small molecule responsible for the modulation of a specific target. The hardest challenge for the medicinal chemist involved in this process is therefore the quest of the optimal affinity, selectivity, and stability of the designed molecules[2].

If on one side the first step of rational drug design is the target validation intended as its real and clear involvement in the biochemical process associated to the pathology, on the other hand there is a not negligible aspect to be considered for the target: its druggability. This term has been extensively adopted to explain, in different contexts, some properties of proteins, ligands or genes. In Medicinal Chemistry, the definition refers to the target capability of binding to a small molecule for the modulation of its activity[3], and sometimes the term druggability is substituted by the synonyms “ligandability” or “bindability”[4,5]. In the last years, there has been a huge rise in the interest for undruggable sites to be deeply studied in order to find a strategic manner to target them anyhow. In 2017, a very interesting perspective on undruggable sites has been published on *Nature Reviews,* reporting the experts opinions on undruggable sites of targets involved in cancer[6]. In this work, Dang, Premkumar, Shokat and Soucek explain their own point of view on the definition of undruggable site. The most curious aspect of this work is the different point of view presented by

[a] *Computational and Medicinal Chemistry Group, Fondazione Ri.MED, Via Bandiera 11, 90133 Palermo, Italy*
[b] *Dipartimento di Scienze e Tecnologie Biologiche Chimiche e Farmaceutiche (STEBICEF), Università degli Studi di Palermo, Via Archirafi 32, 90123 Palermo, Italy*
† These authors contributed equally to this work.

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

the different authors on the same issue. Dang mainly focuses on the protein-protein interaction (PPI) as one of the main undruggable targets because of the lack of well-defined and deep pockets[7,8]. Reddy states that the use of term undruggable is fairly exaggerate and it would be more correct to define those sites as difficult to drug. Actually, analysing the most common studied undruggable sites, during the years several small molecules has been identified as binders, with several candidates reaching the clinical trials (e.g. Bcl-2 family members and transcription factors as STAT3, MDM2 and others). Another important concept concerning undruggable sites is expressed by Shokat in the same manuscript. He points out that there are two main aspects that must be considered in order to classify a binding site as really undruggable. One is related to the chemical intractability of the target and the other one focuses on the need to have sufficient data demonstrating the clinical meaningfulness of a modulator towards that target. Laura Soucek, also contributes to the same topic expressing an interesting differentiation between those targets that have not been yet "drugged" because of structural difficulties, and those difficult to drug because not disease-specific for example in normal and cancer cells. In the latter case the intervention could produce severe side effects in normal cells expressing the same target. In recent years several methods to assess target druggability have been created with the most used mainly based on protein surface descriptors as curvature and lipophilicity. These features are considered to be fundamental for the recognition and consequent binding of small molecules[9–11]. In the last years, massive advancements have been completed in the basic understanding of the biological properties and biochemical role of the undruggable sites especially thanks to the rise in the available structural insights provided both by X-ray crystallography and Nuclear Magnetic Resonance (NMR). Nevertheless, working on undruggable targets is not always possible using only experimental techniques because of the difficulties related to time, costs and not appropriate methods available. For example, the use of X-ray structures provides a static picture of the protein, without any kind of information about how the target can structurally evolve in the presence of a modulator. On the other side, NMR presents some other limitations related to the target location, size and characteristics (not always it is possible to examine the whole protein-ligand complex with NMR techniques)[12]. Opportunely, at the same time, the computational methods reveal their usefulness, offering a valid and supportive alternative to the classical experimental methods[13–17].

Recently, the use of computational methods in drug discovery and development has expanded in popularity and implementation thanks to the good and reliable results showed. The implementation of computational methods within the drug design process, commonly called computer-aided drug design (CADD), gained its main success, in the early years of use, due to its capability of increasing the hit rate of novel drug compounds when compared to the classical high-throughput screening (HTS) approach. The main application of these methods is still the rapid use of virtual screening campaigns to cut the time-to-market for the discovery of new chemical entities. For this purpose, the approach mostly used is the structure-based one, where the three-dimensional structure of the biological target, obtained through methods such as X-ray crystallography, NMR spectroscopy or homology modelling, is used to evaluate the binding capability of a small molecule library[18,19].

In the last years the advances in software and hardware performance allowed researchers to adopt Molecular Dynamics (MD) with great success[20] to address drug discovery issues such as protein-ligand interaction stability[21–24], binding kinetics[25–29] and binding process[30–32]. The understanding of molecular motions is basically the main issue related to molecular recognition and represents the evolution of the old idea of "Lock-and-key model" where a frozen receptor can house a small molecule without mutating its conformation[33]. The dynamic nature of the receptor has been, in fact, largely demonstrated and conformational changes have been related to ligand binding[34,35]. This dynamicity of protein conformation, related to the binding process of small molecules, has represented the key to unveil some strategies to target undruggable sites.

On the light of these considerations, different approaches have been developed based on MD with the aim to explore protein flexibility and discover otherwise accessible hidden binding sites. One of the first methods to unveil undruggable sites through classical MD simulations was issued by Seco et al.[36] In their work, an explicit restrained MD simulation was applied in order to evaluate the binding propensities of a probe addressed to the protein surface. Starting from the molecular trajectory, free energy calculations were performed to assess the molecular recognition process[37]. More recently a similar approach has been adopted by Bakan et al.[38] to demonstrate that the approach of small molecules to proteins produce global

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

and local structural rearrangements of the protein that represent the starting point for increasing target druggability[38]. Nevertheless, some computational techniques adopted for such a kind of study, not always reproduce the native milieu in which proteins are normally plunged into the cells. It is actually important to maintain the maximum reliable conditions in order to avoid protein denaturation or wrong folding changes[39]. Indeed, the dynamic nature of proteins is particularly crucial for some targets exhibiting active, inactive and intermediate configuration alternation (e.g. GPCRs) and it should be taken into consideration when evaluating binding mode of small molecules, as demonstrated by Ferruz et al. in their recent publication[40]. In their work, authors used a combination of MD together with Markov state models (MSMs) to analyse binding mode and interaction pattern of a dopamine D3 receptor antagonist (PF-4363467). The use of aggregated MD through MSMs allowed researchers to unveil an otherwise hidden cryptic pocket, created by the positional shift of highly conserved residue F346[6.52]. The discovery of this cryptic pocket and the pose of the ligand observed by researchers could not be observed using canonical static docking approach.

The aim of this review is to present how MD applications have been used in recent years to treat undruggable sites, in order to unravel targetable pockets using small molecules. Particularly, we have focused on the application of MD to two specific undruggable sites of main interest for medicinal chemists: Allosteric sites and protein-protein interactions (PPI). The next sections of this review will be dedicated to each of these two issues. Inside each part of the work, we will present the different approaches adopted by researchers in order to deepen the binding sites exploration and the design of selective modulators.

Allosteric sites

Proteins functions are strictly correlated to their flexibility and conformational transition causing cavities shape modifications and exposure. This dynamic process is fundamental for the recognition of chemical or biological guests useful to a particular biological process. As a matter of fact, biochemistry interest has always been focused on the conformational changes of proteins related to their biological role with specific regard to the possible cooperative functionalisation between different positions on the protein surface. Such a phenomenon is known as allostery and typically occurs when the protein binding process to a guest molecule transmits some conformational changes to some other proximal or distal different sites on the protein surface[41,42]. Some recent observations unveiled that allostery can be facilitated by dynamic and intrinsically disordered proteins offering new insights in the understanding of allosteric functional regulation[43].The structural rearrangements associated to protein activity modulation can be referred to small side chains conformational changes as well as important modifications, within the quaternary structure, of spatial protein motifs distribution. Following the classical model of allostery proposed in the last years, the real definition should refer to the latter case[44]. Molecular modeling, and particularly molecular simulations can indeed help the understanding of such functional motifs rearrangement and movements within the protein. Considering the wide collection of molecular simulations approaches, it is imperative to underline that not every model is appropriate to completely explore the protein conformational space. Some classical simulations can in fact be limited to the exploration of only certain energy landscape portions because of their sampling capabilities[44]. When using MD it is difficult to correctly and widely sample the useful landscape for structure transition and it becomes essential to use biased MD techniques[45–47] or other modified approaches such as supervised MD[48]. Using biased methods allow researchers to explore wide landscape of protein motions focusing on the energetic aspect of the allostery phenomenon. Moreover, the use of classical MD simulations not always is capable to provide useful information about conformational changes, unless long simulations are used[49], but in the latter case it must be also considered that the longest the simulation is, the highest the approximation becomes[50]. Practically, the best way to catch all the useful conformational sampling information should be the use of enhanced sampling techniques[51,52]. One possible alternative to long simulations could be the use of multiple shorter simulations then analysed with Markov State Models to catch all the quantitative parameters to analyse[53–55]. In the latter, all the process must be well monitored in order to explore the structure and rebuild the whole process. In this section we will present the most recent approaches based on MD used to deepen allosteric regulation of proteins function related to their biological role. We will hereby present the most recent advances grouped based on the approach used (classical MD or Biased MD).

Use of molecular probes in Molecular Dynamics for novel allosteric binding sites discovery

Some of the below reported works present noteworthy approaches to study structural evolution of allosteric binding

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

site cavities and conformational shifting in proteins when targeted with probe molecules, that can act as modification inducers.

One of the first assays in this field has been published by Bakan et al. in 2012[38]. In their work, authors start from the definition of druggability as something related to affinity of small molecules towards binding sites available on the protein surface. The use of molecular probes designed on approved drugs scaffolds allowed researchers to effectively evaluate binding affinity and druggability for some challenging targets and especially for some hidden allosteric sites. The use of approved drugs with known experimental binding data constituted a method validation to check if theoretical binding affinities were correctly estimated. In this work, two main aspects were deepened, firstly the analysis of putative binding mode for several probe ligands towards different proteins, and secondly the consequent identification of the most druggable sites. Starting from collected data as geometry and energy parameters of the recognition and binding process, researchers could correctly evaluate the binding affinities. This methodology has been applied to several targets as protein tyrosine phosphatase 1B (PTP1B), lymphocyte function-associated antigen 1, vertebrate kinesin-5 (Eg5), and p38 mitogen-activated protein kinase (MAPK). One of the most interesting aspects sharpened by this method is the possibility to unveil putative interaction spots on the protein surface, otherwise hidden. Remarkably, authors used probes with different physicochemical properties to harvest consistently more reliable predictions, not biased on a specific chemotype or physicochemical profile of ligand, therefore widening the type of putative binding sites that can be "gathered" within the protein. The use of mixed probes highlighted some interesting aspects of the binding process for the analysed targets, especially on the Eg5 and p38 MAPK. On the other side, using charged probes let authors to discover some important ligand features useful to bind phosphatase PTP1B catalytic sites.

Benzene probes have been also used in a similar manner by Tan et al. in a very recent work[56]. In their study, authors charted putative oncoprotein MDM2 binding sites using a Ligand-Mapping Molecular Dynamics (LMMD) simulation technique. In particular, this method allowed the discovery of two new sites both in the N-terminal domain of the protein. The first one situated between Tyr100 and Tyr104 was mapped in both apo and holo form of the protein, revealing an interaction region which in the X-ray crystal structure seemed inaccessible to any ligand. This region had been already demonstrated to be essential for the nutlin binding at the p53 cleft[57]. The second presumed site, in a region nearby the p53-Pro27 interaction site, was only identified during simulation of the apo protein. In the holo protein simulation, this region resulted occupied by the C-terminal of the p53 binding partner. The efficacy of the method adopted was then demonstrated by biophysical assays that proved the real existence of one of the predicted binding sites close to the consensus p53-binding cleft in MDM2, underlying how molecular simulations can be invaluable for the rational design of new drugs. Based on the simulations outcomes a series of hydrocarbon stapled peptides were proposed to target the binding site identified during simulations (Figure 1). The use of these peptides improved the activity of already known MDM2 ligands because of a cooperative action of the new discovered allosteric site with the canonical orthosteric site. Moreover, the structural knowledge of these new binding sites opened up to novel strategies of peptides optimisation, that can improve the binding mode of other ligands targeted for the MDM2/p53 interaction.

A noteworthy application of molecular probes has been similarly presented by Luo et al.[58]. In their work, authors focused on a challenging target such as the two-pore domain potassium channel (K2P). This is a main actor in the membrane potential maintenance presenting a unique structural feature of extracellular cap formed by the E1 and E2 helices, not observed in other ion channels. Luo et al. adopted different computational chemistry techniques, mutagenesis, and electrophysiology experiments in order to characterise binding mode of N-(4-cholorphenyl)-N-(2-(3,4-dihydrosioquinolin-2(1H)-yl)-2-oxoethyl) methanesulfonamide (TKDC), used as probe, to study the binding process to the extracellular cap of K2P channel, an allosteric and difficult site to be targeted. The study underlined important differences in the different binding mode of ligands on the different potassium channels (TREK, TRAAK etc.). The first stage of computational approach was conducted using molecular docking on the previously identified putative binding site. Following the docking outcomes, some mutagenesis experiments within the binding site were conducted in order to confirm the residues responsible for the

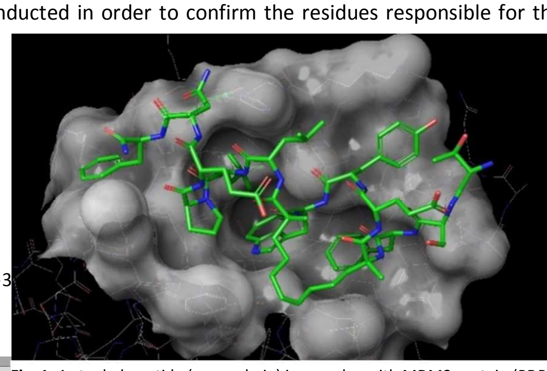


Fig. 1: A stapled peptide (green chain) in complex with MDM2 protein (PDB ID: 4UE1)

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

binding mode retrieved from docking experiments in the different potassium channel models. Starting from the binding mode hypothesis, a chemical probe derived from TKDC was used for molecular dynamics to explore the extracellular binding site of the different channel models and further validate the proposed binding mode previously observed. These outcomes have been then exploited to supervise virtual screening of other molecules and find out new allosteric inhibitors of K2P (Figure 2A).

Biased Molecular Dynamics for allosteric modulators design

In the last years several MD applications have been used to better understand the allosteric sites role within the proteins. One of the most attractive approaches is based on Supervised Molecular Dynamics (SuMD). In this technique, ligand–receptor recognition is well-investigated in a relatively short time period (ns). The method relies on the use of a specific algorithm capable to focus on the binding process between protein and ligand speeding up the recognition trajectory. Such an approach allows to deepen some crucial aspects of the binding event with special focus on the meta-binding sites or allosteric sites[59]. In their work, Deganutti et al. applied SuMD to GPCR A_3 adenosine receptor to study the effect of a positive allosteric modulator (LUF6000, Figure 2B)[48]. This work represents one of the first applications of MD for studying allosteric recognition mechanism. Indeed, it effectively revealed its importance in the elucidation of what experimentally observed, opening up to two different hypotheses of binding and subsequent regulation of LUF6000 on A_3 adenosine receptor. In their experiment, the authors placed the allosteric modulator about 60 Å away from the orthosteric binding site, occupied by the natural agonist adenosine.

The binding process pathway obtained from the simulation enlightened two possible ways through which the ligand may act on the receptor. In the first hypothesis, LUF6000 produced conformational changes inside the protein empowering the adenosine to strengthen interactions previously formed in the binding pocket. The other mechanism proposed was the formation of a ternary complex LUF6000-receptor-adenosine where the role of LUF6000 was to act as a sort of cap stabilising the binding of the other two interacting partners. Another biased MD technique widely used for allosteric modulators discovery is MetaDynamics (MetaD). The great intuition behind this method relies into the use of an additive potential applied to the system analysed in order to overcome some energy barriers of the Potential Energy Surface (PES) allowing the complete exploration of the energy landscape in the protein conformation shift.

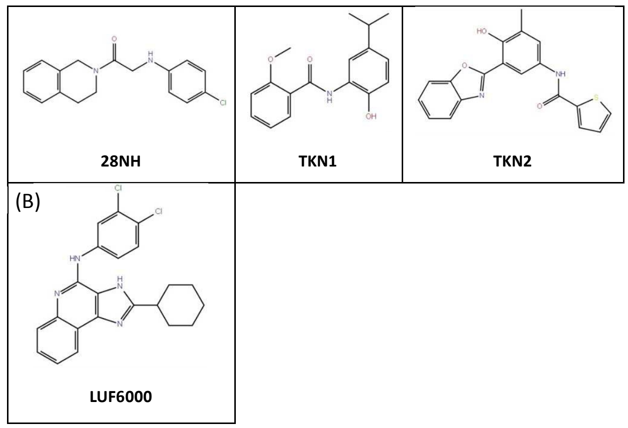


Fig. 2: **(A)** Allosteric inhibitors of K2P - 28NH, allosteric inhibitor of TREK channels; TKN1 and 2, allosteric inhibitors of TRAAK. **(B)** Positive allosteric modulator, LUF6000.

MetaD was firstly applied by Laio and Parrinello and it is usually adopted for molecular simulations in order to expand protein conformational changes exploration[60]. As a matter of fact, the use of this technique is particularly indicated for allosteric modulation studies because of its capability to deeply explore all the conformational changes in a target. In 2015, Grazioso et al. applied MetaD together with Essential Dynamics to Alpha7 nicotinic receptor to give a mechanistic hypothesis of the allosteric modulation[61]. In detail, the application of these techniques allowed researchers to explore the free energy landscapes related to the open and closed states of the protein loop C. In this study, the effect of different modulators (including an agonist, a positive allosteric modulator, and a newly reported ago-allosteric modulator) on the conformational change of the protein was investigated. Every ligand considered showed a unique particular free energy profile, and most importantly, the possible interaction between orthosteric site in the loop C and M helices within the protein structure. This specific interaction was evidenced by the ago-allosteric modulator GAT107. In fact, when bound to the allosteric binding site, GAT107 induced a loop C rearrangement typical of a full agonist, thus providing a possible explication of the experimentally demonstrated ago-allosteric properties of GAT107. The results obtained from the computational approach were in perfect agreement with what observed in the

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

experimental assays and represented an outstanding advancement in the nicotinic receptor biology knowledge.

In 2017, Gomez-Gutierrez et al. used accelerated MD for 6 μs to identify putative allosteric modulators for p38α MAP kinase[62]. The method[63] is based on an enhanced-sampling algorithm for sampling conformational space by reducing energy barriers, thus modifying the potential energy profile. Fixed a defined energy level, the algorithm does not affect energy-profile points above this zone, while it rises up wells that are below the fixed threshold energy level. As a result, the energy profile barriers are reduced allowing a wider exploration of the entire potential energy surface otherwise not easily scouted. In their work, Gomez-Gutierrez et al. let p38α MAP kinase to undergo a 6 μs accelerated dynamics collecting all the structures during the trajectory and clustering all the ensemble. Clusters were created using Principal Component Analysis (PCA) first and Cluster Analysis after, in order to collect the most representative structures during the simulation. The collected structures represented the starting point for a hot spot analysis conducted on FTMap. Through this approach, it was possible to confirm all the canonical well-known sites of the protein such as DFG pocket, lipid binding pocket, DEF site and others. Moreover, this study unveiled new allosteric binding sites named NP1 and NP10. These two contacts areas in particular caused protein structural rearrangement involving elements responsible for the protein activation (e.g. the activation loop, the catalytic loop, the glycine-rich loop and others).

Classical Molecular Dynamics for allosteric modulators design

Classical MD has been widely applied for studying allosteric modulation of biological targets. One interesting application of molecular simulation to comprehend the structural mechanism associated to signalling pathways of lymphocyte function has been recently published by Abdullahi et al.[64]. In their work, the authors applied molecular modeling to examine lymphocyte function activation, through an allosteric shape shifting of the lymphocyte-associated antigen, when it is bound to a specific modulator (ICAM Binding Enancher-667 – IBE-667). During the simulations, several parameters were collected in order to deepen and understand the succession of events leading to the active conformation of protein. Particular attention was given to variations in residual distances, dihedral angles, triCα angles to evaluate how the structural rearrangement was related to these variations. The conformational change between the inactive and active state of the target was characterised by significant fluctuation in residues positions and in energy stability of the complex. The shape shifting was strictly accompanied by a α7 helix movement driven by metal-ion dependent adhesion active site (MIDAS domain), both synergistically cooperating to the activation of LFA-1 Integrin responsible for lymphocyte function. The strength behind this work was the capability of demonstrating a synergistically interplay between MIDAS domain region and downward α7 helix motion necessary for the biological activity.

MD, in the last years, has been widely applied to the identification of new scaffolds and new chemical entities, integrating the classical virtual screening techniques to prioritise hit molecules and elucidate their binding mode within the protein cavity. Some mitotic kinesin Eg5 allosteric modulators have been identified thanks to MD application by Makala and Ulaganathan in their work published in 2017[65]. Eg5 is a well-known target for cancer therapy, but all the discovered compounds so far were addressed to the canonical orthosteric binding site because of its ease of access for small molecules and the availability of structural data. In this work, the authors firstly docked some free available molecular libraries on the allosteric site (site2) of the target and ranked molecules prioritising them based on the docking score. The best 5 poses retrieved from the docking were then submitted to MD to evaluate conformational rearrangement and stability of the ligand-protein complex. The results obtained from this study suggested the pyridazine scaffold as an optimal starting point for further Eg5 allosteric modulators development.

The use of MD as strategy for unveiling undruggable sites in membrane ion channels was employed by Martin et al. and Guan et al.[66,67] to study respectively pentameric ligand-gated ion-channels and Calcium channels. In the first work, authors presented an extensive point of view on glutamate-gated chloride channel (GluCl) responsible for the intercellular synaptic communication. A biochemical mechanism of activation/deactivation for this target is not still well-recognised and this work was a first attempt to elucidate the shift from open to closed state of the channel. The experiment was conducted using the protein in its active state bound to L-glutamate and the positive allosteric modulator (PAM) Ivermectin. On this system a μs-long simulation was run in order to explore the structural relaxation upon PAM modulator ejection. Analysing the MD trajectory two different transition states were retrieved, and most importantly, it was clarified

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

that the structural global twisting observed is the unique responsible feature for the channel pore closing, acting on the M2-M3 loop at the interface between extracellular and transmembrane domain. Further simulations in equal restrained conditions showed the same structural rearrangement opening up to a pharmacological mechanism clarification of PAMs in this neurotransmitter receptor family. The rearranged structure observed during the dynamics simulation was comparable to the X-ray crystal structure of GluCl, thus enforcing the reliability of Molecular Dynamics-based method. Glutamate receptors (PDB ID: 4OR2, Figure 3) represent one of the most important targets involved in neurological diseases and, in certain cases, they show low subtype selectivity for orthosteric modulators. For this reason, it is necessary to address new investigation strategies such as the allosteric modulation. Starting from this hypothesis, Jiang et al. led a study on mGluR1 receptor in order to discover putative new negative allosteric modulators (NAM) derived from Chinese herbs components[68]. In this research work, authors started from the crystal structure of the seven-transmembrane domain of mGluR1 to detect the putative allosteric binding sites and run some pharmacophore screening. The authors combined the structure-based interaction pharmacophore with ligand-based approach in order to increase the specificity of virtual screening.

The ligand-based approach exploited different scaffolds of known NAMs, thus allowing to have pharmacophore models differentiated into molecular classes. Starting from these first results, the models were validated using a test set and the most reliable model was then used for the prospective screening campaigns. The best ranked compounds based on pharmacophore features matching were then docked into allosteric binding site of the protein in order to evaluate their binding pose. Later, the best poses were submitted to MD to evaluate the effective stability of the ligand-protein complex and the interactions stability of the selected molecules.

One of the most important outcomes from this study was the identification of some key protein residues (Leu757, Asn760, Trp798, Phe801, Tyr805 and Thr815) that revealed to be crucial for NAM selectivity and binding stabilisation.

Another comprehensive sight of the possible use of MD to discover otherwise unrevealed undruggable sites has been proposed by Pabon and Camacho in a very exhaustive work published in 2017[69]. In this study, the attention was focused on the apo form of anticancer target PD1, a particular hard-to-drug protein. The use of MD revealed a new hidden hydrophilic cavity around Asn66 and Ile126 residues participating to the ligand recognition pattern.

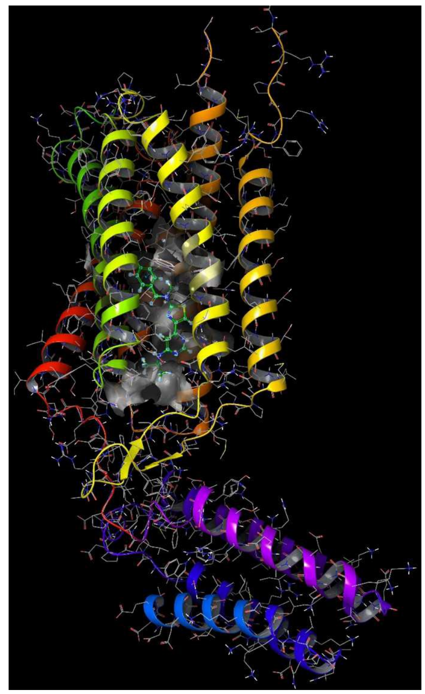


Fig. 3: Metabotropic Glutamate receptor 1 in complex with a negative allosteric modulator (green molecule).

The use of two PD1 ligands, L1 and L2, permitted to discover that, while the unbound PD1 presented a hard-to-target hydrophilic interface to host the ligand, the recognition of both L1 and L2 induced a complex conformational shift with the consequent opening up of a hydrophobic cavity otherwise unreachable by any small molecule. These outcomes opened up new strategies to rationally design selective compounds and suggest a possible efficient biophysical approach to the evaluation of the binding pathways as a mean for targeting undruggable proteins.

As recently shown in a paper published by Marko Novinec[70] computational investigations helped also the discovery of allosteric effectors active on cathepsin K and other related endopeptidases. In his work, Novinec firstly presented an interesting scenario about allosteric targeting as a progressively winning strategy in drug discovery, even though not yet well explored mainly because of lack of structural information. Papain-like cysteine peptidase cathepsin K allosteric modulation

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

was investigated through MD, together with other peptidases, to catch any putative conformational shift primarily important for the protein activation process. In this approach, MD-derived conformational space was plotted for different cathepsin endopeptidases (L, K, S and V), using Principal Component Analysis in order to explore possible conformation “clusters”. Proteins of the same family adopted similar conformations during the MD trajectories. At a later stage, the author proceeded with a deeper analysis of cathepsin K to show how some known allosteric modulators, NSC13345 and NSC94914, affect conformational changes. During the simulations, these effectors were able to influence the active site conformational shift, affecting the region nearby sites S1 and S2 pockets. This portion is the tightest part of the active site cleft and it is responsible for specificity in ligand binding process. Using molecular docking on structures retrieved from MD, it was supposed that allosteric modulation starts stabilizing pre-existing conformational state, prior to influence the binding of substrate into the orthosteric site. The comparison of this results to other of related enzymes confirmed this as a possible mechanism for allosteric modulators, thus the usefulness of MD in unveiling binding site otherwise difficult to be recognised and targeted.

Another interesting application of MD to unveil allostery phenomenon has been published in 2017, by Latallo et al. In this work the authors used simulations to predict allosteric mutations responsible for increasing antibiotic resistance mediated by beta-lactamase[71]. This research topic starts from the evidence that allosteric mutations are really difficult to be predicted prospectively. In this work, allosteric mutants of CTX-M9 enzyme have been used for MD simulations examining a wide range of antibiotics. Experimentally, mutated isoforms of CTX-M9 showed a rise in their catalytic activity and efficiency. When the same mutants were studied “statically”, starting from crystal structures, no differences were noted in comparison with the wild-type form of the enzyme. Based on this outcome, researchers concluded that the activity increase could be related to conformational changes within the structure, not observable with canonical static *in silico* screening, allowing the enzyme to enhance its activity. The use of machine learning techniques, applied to MD trajectories allowed to discover the most important allosteric mutations influencing the conformational rearrangement of the catalytic site. This study highlighted how conformation shift was important for the rising of the catalytic activity of the enzyme in developing drug resistance, not affecting the minimum-free energy. There are different theories about the conformational change propagation from the allosteric sites to the core catalytic pocket[72]. In their work, Latallo et al., showed that such mutations substantially did not affect the catalytic site conformation in the apo form of the protein, leaving the general structure conformation unaltered. Thanks to the application of machine learning techniques to trajectories derived conformations, the authors found out that the catalytic activity variation could be connected to a particular set of protein residues involved in coordinating catalytic water within the substrate binding site or directly involved in the substrate positioning. These outcomes open up some interesting hints in the rational design of new antibiotics.

Molecular Dynamics simulations were also applied to address the problem of allosteric modulation in $A2_A$ adenosine receptor. Recently an interesting work has been produced by Caliman et al.[73] as a follow-up of a previous one where the same authors had already explored the conformational analysis of Apo $A2_A$ adenosine receptor through MD[74]. In the previous work, Molecular Dynamics analysis allowed to identify different non-orthosteric sites on the active conformations, on both two intermediate conformations observed, and on the inactive conformations of the protein. In this more recent work, authors started from different structures retrieved from the previous dynamics simulations and 20 different X-ray structures of the same protein family, in order to map the previously retrieved allosteric sites using the fragment-based approach accessible by FTMap software[73]. This software, usually used to identify available non-orthosteric sites, revealed to be especially helpful for transmembrane proteins and especially for compounds that cannot target the extracellular part of the protein but go on the lipid bilayer of the membrane. The use of MD combined with FTMap allowed to mainly identify five allosteric binding sites that are present in both active, intermediate and inactive protein conformations. We here report a table of the five sites identified as was reported in the original paper (Table 1). Such a kind of study represents an invaluable help for the designing of new allosteric modulators for difficult targets where selectivity is difficult to be reached. As in the previous case, the Bartuzi research group extensively adopted MD on a GPCR receptor to simulate the activation and the interplay between allosteric sites in the human μ-opioid receptor (MOR)[75,76].

In the first work, the authors started from evidences on negative modulators of μ-opioid receptor, a GPCR receptor, for

View Article Online
DOI: 10.1039/C8MD00166A

identification of allosteric sites and modulation mechanisms. In detail, starting from homology modeling, known compounds were docked into the allosteric sites and then conformations and complex stability was evaluated using MD. In particular, Salvinorin A, a negative allosteric MOR modulator, was used to evaluate key residues in the binding process using a site-directed

Site	Location	Regions	Residues
1	Intracellular crevice	TM3/TM4/TM5	$I^{3.40}$, $F^{3.41}$, $L^{3.44}$, $A^{3.45}$, $D^{3.49}$ (TM3); $I^{4.45}$, $I^{4.48}$, $C^{4.49}$ (TM4); Y112 (ICL1); $C^{5.46}$, $P^{5.50}$ (TM5)
2	G protein coupling site	TM2/TM3/TM6/TM7	N39 (ICL1); $T^{2.39}$, $N^{2.40}$ (TM2); $D^{3.49}$, $R^{3.50}$ (TM3); $H^{6.32}$, $S^{6.36}$, $F^{6.44}$ (TM6); $Y^{7.53}$ (TM7); I292 (C-Term)
3	The Lipid interface	TM5/TM6	$P^{5.50}$, $M^{5.54}$ (TM5); $V^{6.41}$, $F^{6.44}$, $W^{6.48}$ (TM6)
4	C-terminus cleft	TM1/TM7	$L^{1.45}$, $G^{1.49}$, $L^{1.52}$ (TM1); $V^{7.47}$, $P^{7.50}$, $F^{7.51}$ (TM7)
5	Extracellular cleft	TM3/TM4	$C^{3.30}$, $F^{3.31}$, $V^{3.34}$ (TM3); $L^{4.58}$, $G^{4.57}$, $F^{4.54}$ (TM4)

Table 1: data taken from original paper[73].

mutagenesis. According to what reported in literature as experimentally proved, the residues Ile316 (7.39), Tyr320 (7.43), Gln115 (2.60), Tyr312(7.35), Tyr313 (7.36), and Tyr119 (2.64) demonstrated to be essential in the recognition and binding of allosteric modulators in κ-opioid receptor. From the results obtained, the authors suspected possible overlapping regions between the Salvinorin A interacting portion (TM I-TM II-TM VI interface) and orthosteric binding site. The main mechanism of action underlined for Salvinorin A consists in interfering with orthosteric ligands. The same authors went deeply in order to better understand possible interplay between two allosteric sites and agonist binding process.

This time, 200 ns replica molecular simulations were run using a positive modulator, BMS986122 (BMS), to evaluate its influence in the binding process of the agonist (R)-Methadone (RME) and Na^+. The simulations were able to differentiate for BMS and Na^+ orientations within the TM VII. On the different trajectories, PCA was applied to investigate on possible clusters of common conformations and TM rearrangements. From the simulations analysis it was deducted that the agonist binding process was negatively influenced by the presence of the sodium ion interacting with conserved Asp2.50, as it was been already proposed both experimentally[77] and computationally[78,79]. Focusing on the BMS, it was noted that its role was the stabilisation of the target–agonist interaction, whereas the allosteric binding of the Na^+ was instead disrupted. These data were in complete agreement to experimental findings, that is a BMS-induced increase in full agonists affinity towards MOR[80]. These results moreover highlighted a possible binding mode of the RME involving Asp3.32 with a consequent rearrangement of TM VII position. This conformational change seems to be driven by an influence on Trp7.35 in the binding pocket, causing the rotation of TM VII. This hypothesis was then confirmed by adding BMS, that stabilised the RME binding through a direct interaction with Trp7.35.

Furthermore, during the revision process we found a "just published" work by Meng et al.[81], where a great approach based on the use of MD/MSMs approach combined with multi-source seeding strategy was used to explore different possible conformations and transition states of Abl Tyrosine kinase. The use of multi-source seeding strategy consists in using different source protein conformations (X-ray, homologs, "piecewise-mixing" conformations from different crystal structures, previous MSMs retrieved conformations). The use of such an approach allowed researchers to widely explore the conformational space of Abl Tyrosine kinase, including the myristoyl-binding pocket situated at the C-terminus of the protein. This portion was identified as an allosteric site responsible to modulate conformational transitions between active and inactive state of the protein. From MSMs outcomes, researchers found out a specific conformational state of the allosteric site, promoting the DFG-out conformation and maintaining the protein into the inactive state. These findings represent a great value in the design of possible Abl allosteric inhibitors.

Protein-Protein Interactions as undruggable targets

In the last decades, PPI drew attention of the scientific research community across academia and industry, because of the increasing knowledge about their relevant role in cells for signalling and regulation of cellular life-cycle and vital functions (e.g. cellular growth, differentiation and apoptosis). The PPI targets space is consistently larger than the classical protein targets one, with putative relevant PPIs comprised between 130.000 to 650.000[82–85]. Such a huge amount of potential targets is associated with biological implications in several diseases, for example cancer and neurodegenerative disorders.

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

Unfortunately, to date less than 0,01% of PPIs belonging to human interactome presents approved modulators[86]. Protein-protein interactions are particularly complex targets, usually labelled as undruggable, and thus there are few discovery programs as these are considered high-risk failure therapies[82]. Actually, cells are crowded environments where proteins behave as promiscuous macromolecules, i.e. able to take part to interaction networks, binding more than one partner, and in this way making difficult achieving specificity during a drug design process. Hence, targeting bi-molecular complexes requires an interdisciplinary approach to identify binding determinants at PPI interfaces and overcome issues tightly linked to the intrinsic nature and structural features of protein-protein interactions[87,88]. Indeed, PPI interfaces are often shallow and lack deep grooves able to accommodate a ligand and recognise its shape in a complementary manner. While in a classical receptor-ligand interaction the first one presents a well-defined pocket with clear complementary binding recognition motifs, a protein-protein interaction occurs when both protein partners establish few high-affinity contacts through the so-called "hot spots" residues, exploiting complementary regions as well as several weaker interactions. However, hot spots are residues widely deployed within protein surfaces, and thus sequentially not connected among them within the same protein, creating a discontinuous epitope[87,89]. Another consideration is related to the size of PPI interface, that is wider (on average 1500 to 3000Å^2) than receptor-ligand contacts areas of classical targets (about 300 to 1000Å^2) (Table 2)[87]. Furthermore, if a protein target takes part into the interaction with another protein, mainly providing protrusions and not just pockets for accommodation, managing the design of a small molecule able to cover a protrusion of the protein target appears rather unlikely[82].

Despite the above difficulties and the shortage of structural information about protein-protein complexes, PPIs are becoming more accepted and popular targets[90], thanks to promising computational techniques, such as the MD. Here, we provide an overview of some case studies where Molecular Dynamics techniques proved its usefulness over the drug design and discovery process for thoroughly studying protein-protein interactions, but above all facing the intrinsic difficulties of this type of targets.

Molecular Dynamics simulations for the identification of potential pockets and binding hot spots

To date, many theoretical and computational tools have been developed to map a potential protein-protein binding site, some examples are AnchorQuery™ [91], and FTMap[92]. The limitation of these techniques relies on the static structures to be analysed. In this context, MD can represent a valuable tool. Actually, it has been used by numerous research groups to identify hot spots responsible for the interaction between two proteins[93].

Host-guest interaction pocket attributes		
	Protein-Ligand	**Protein-Protein**
Shape	Deep	Shallow, flat
Size	~300 to 1000Å^2	1150-1200Å^2 small interfaces [97,98] 1200-2000Å^2 medium interfaces 2000-4660Å^2 large interfaces [97,99,100]
Types of interaction	Electrostatic interactions, hydrogen bonds, hydrophobic contacts, π–stacking	Hydrophobic contacts (for protein-protein complexes formation), and electrostatic interactions (for PPI stabilization)

Table 2: data from original paper[88].

The hot spots within a PPI interface represent less than 50% of contact area between proteins and they are defined as those amino acid residues that replaced with alanine, by alanine scanning mutagenesis, provoke a decrease of binding free energy of at least 2 kcal/mol[94,95]. The hot spots more frequently found into PPI interfaces are the amino acids Tyr, Trp e Arg[95,96]. All these three residues take part in hydrophobic interactions, since the main driven force for protein-protein complexes formation is precisely the hydrophobicity[88].

In a work published in 2014, Sing Tan et al. showed how MD simulations present a remarkable potential for detecting hydrophobic hot spots and for ligand-mapping. Indeed, these computational techniques allowed to unveil cryptic binding sites on specific proteins (RAD-51 and MDM2) surfaces by MD simulations of 5 ns and 20 ns. The authors suggest to use shorter simulations (5 ns) for mapping protein X-ray crystal structures and simulations of 20 ns for NMR-resolved structures to ensure a careful exploration of protein cavity.

While shorter simulations allowed to identify pockets buried by amino acid side chains and protein backbone in MDM2 and RAD-51, respectively, those of 20 ns were able to unveil the hot spot Leu26, that is the most buried of the three MDM2 residues responsible for p53-binding. Furthermore, in the same paper, the authors focused on the capacity of MD techniques to unveil

View Article Online
DOI: 10.1039/C8MD00166A

hydrophobic regions able to bind hydrocarbon-stapled peptides at interfaces of proteins, such as MCL-1 and BH3 α–helices of Bcl-2 family proteins[93]. The hydrocarbon-stapled peptides are peptides folded as α-helix, which present all-hydrocarbon "braces" (staples), that make them suitable pharmacological candidates to disrupt protein-protein interactions[101]. For this purpose, small molecules probes (benzene, propane and isopropanol) were used to mimic most of protein residues interactions between and hydrocarbon-stapled peptides. In particular, exploiting the ligand-mapping MD simulations it was possible to detect a new binding site, previously unexplored, and design a peptide inhibitor ($SAHB_{8\text{-}12}$) of the hydrophobic interactions between the two protein partners[93].

Another example of MD application to PPI modulators design, is reported in the work of Saez and co-workers published in 2015. They carried out MD studies concerning the investigation of hot spots involved in the interaction of the acid-sensing ion channel 1a (ASIC1a) and its selective inhibitor Psalmotoxin-1 (PcTx1), a peptide extracted from spider venom. ASIC1a is a member of degenerin/epithelial sodium channel family[102,103], involved in several diseases including chronic pain[104] and ischaemic stroke[105]. To date, the most potent and selective inhibitor discovered for this ion channel is precisely the peptide Psalmotoxin-1, that binds and fixes the desensitized state of ASIC1a[106,107]. Saez et al. analysed the interaction of these two protein partners to elucidate the nature of the binding contacts and to identify the crucial hot spots[108]. Starting from a crystal structure resolved by Dawson et al. in 2012[109], the authors observed that the two protein partners established 57 intermolecular contacts, which authors defined as "pairwise interactions < 5 Å". The bulk of this network was reduced by MD simulations of 30 ns carried out with GROMACS 3.3.3 (GROMOS54a7 force field). The interaction network rate was analysed out of 600 frames for each simulation, setting a cut-off of 5 Å for non-bonded interactions and 2,5 Å for hydrogen bonds. The MD trajectory analysis halved pairwise interactions to almost 31 intermolecular contacts. These outcomes were further examined in depth by alanine scanning mutagenesis, which led to the identification of a smaller amount of hot spots, thus paving the way to novel selective compounds design for ASIC1a[108].

Furthermore, in a recent article, Biswas et al. presented a study on TRAF6/Basigin interaction implicated in melanoma metastasis. Basigin (BSG) protein is able to stimulate the overexpression of matrix metalloproteinases (MMPs), which contribute to cancer development, and to interact with tumour necrosis factor receptor-associated factor 6 (TRAF6), promoting the invasiveness of melanoma cells[110–113]. Biswas and collaborators performed MD simulations of individual proteins and complex through GROMACS 5.0.5[114] package with CHARMm force field. Simulations times were 70 ns (BSG), 50 ns (TRAF-6), and 120 ns (BSG-TRAF6 complex). MD results allowed to observe a conformational change of BSG transmembrane region, participating to the PPI, which as a consequence of TRAF6 binding acquired a helical conformation. Besides, these simulations provided information about the interacting hot spots of TRAF6, recognizing residues more contributing to binding free energies MD proved to be a useful instrument to recognise residue contacts previously not identified between protein partners, as compelling aid for drug design and development.

Another interesting work published in 2017 by Xue and collaborators concerns the use of a recent MD technique, the Steered Molecular Dynamics (SMD). The aim of this work aimed at obtaining information about an interaction between two chaperones, Hsp70 (also called Hsc70) and Hsp40 (or auxilin)[115]. These two proteins take part in a cellular ATP-consuming network, which ensures the correct proteins folding, membrane translocation and protein degradation[115–117]. In order to extensively analyse Hsp70 nucleotide-binding domain (NBD) and Hsp40 J domain interaction, the SMD turned out to be useful. SMD is a methodology which involves an unequilibrated system and consists in applying external forces to the system under consideration forcing the protein complex dissociation[115,118–121]. During the simulations, binding energy changes were registered and values were reported into a curve against the simulation time. In an early phase, starting from a PDB file (1Q2G) containing Hsc70-auxilin complex, Xue et al. equilibrated the system performing a classical MD simulation of 20 ns at pH 7.0 and 300 K (using GROMACS 4.5 program[115,122], in constant NPT and periodic boundary conditions). RMSD (Root-Mean-Square Deviation) was calculated and allowed to extrapolate the stabilised structure at 7 ns of simulation as a starting point for SMD studies. Later, the equilibrated system was submitted to a SMD simulation of 2 ns, applying a spring constant of 300 kJ mol^{-1} nm^{-2}. This value was chosen as more suitable for the complex under consideration, given that it produced reliable and detectable rupture forces responsible for Hsc70-auxilin complex dissociation. Binding energies values (kJ mol^{-1} nm^{-2}) against simulation time (ps) were reported into two

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

curves according to different types of interaction involved (electrostatic interactions and van der Waals interactions) and the sums of each points of these two curves were plotted into another curve. This plot suggested that, in the early phases of SMD simulation, electrostatic interactions out-numbered van der Waals (VdW) interactions, on the contrary, in the second part of the simulation VdW interactions dominated. At the end of SMD simulation, the binding energies of the complex achieved the value zero, revealing a complete protein complex dissociation. Furthermore, in order to deepen which residues and how much these ones affect electrostatic and VdW interactions within the complex, Xue and co-workers also reported the residues involved in binding domains of proteins and the related binding energies. In this way, it was possible to identify key residues in the J domain of auxilin and in NBD of Hsc70. In view of the above, a host-guest complex dissociation process using SMD is not to be intended as the opposite of binding process between two interacting partners. Instead, this technique represents an aid to deepen the character of established interactions and the related effect within a given complex[115]. A very recent and modern approach based on MD, together with MSMs has been presented by Martinez-Rossell et al.[123]. In their work, authors adopted a MD-driven approach for fragment screening on CXCL12, a "hard-to-drug" chemokine because of its partial shallowness. The interesting and innovative point of view presented by authors consists in using combined methods for targeting CXCL12 and not its receptor CXCR4 as already done by others[124]. In their work, Martinez-Rossell et al. used a series of short MD simulations on the target to explore possible conformations. On the obtained conformations docking was run using fragment libraries. Then, different cycles of *short MD-MSM-adaptive respawn*, based on the use of the target and solvated ligand, were run to produce putative binding poses to be analysed. One of the strengths of the method used is based on the short simulations as starting points for an adaptive sampling scheme, consisting of their concatenation using MSM approach in order to enhance the sampling capability. The use of stochastic method such as MSM allows the reduction of the single timescale simulations and at the same time open up the possibility of deeply study a wide landscape, allowing to create aggregate simulations of 29 to 45 µs. Furthermore, the use of MSM in this work permitted to observe unbound and different bound states for the ligands screened, together with probabilities of complex creation and free energies associated. Besides, this approach allowed researchers to evaluate kinetics parameters (k_{on} and k_{off}) for the unbound/bound state of the ligands screened.

Conformational shifts analysis as starting point for PPI modulators design

The structures obtained by X-ray crystallographic do not permit to deeply explore each protein cavity, especially those transient, which could be responsible for protein-protein interactions. Furthermore, X-ray crystal structures do not account for flexibility and adaptation induced by both proteins that often result in complementarity between two protein partners[125]. In this case, starting from PDB files or homology models, MD methodology provides a dynamic trajectory over time of the positions and the velocities of all atoms in a system, allowing to investigate protein surface, reproducing conformational changes occurring into cellular environment, and detecting potential shallow pockets able to accommodate ligands[88].

In 2015, Cau et al. reported a study on g14-3-3 protein of the protozoan parasite *Giardia duodenalis*, which colonises the upper regions of small intestine in mammals, causing severe consequences to host health[126,127]. At present, there are no vaccines available and the number of useful drugs is rather limited, presenting also refractory cases[127,128]. As a consequence, targeting the interaction between g14-3-3 protein, responsible of triggering invasive activity of the parasite, and host protein partners (phosphorylated Ser and Thr proteins – pSer/pThr proteins) represents a priority solution to address this unmet medical need. G14-3-3 protein is able to explicate its activity only upon phosphorylation on Thr214 residue, producing a conversion from "open" conformation of apo form to "close" one of phosphorylated-form[127,129,130]. In early phases, Cau and co-workers conducted SGLD (Self-Guided Langevin Dynamics) simulations using SANDER on nine tripeptides belonging to the phosphorylation region. At a later time, classical MD studies were performed with PMEMD (Particle Mesh Ewald Molecular Dynamics) module of AMBER12 with AMBER force field ff12SB in explicit water solvent on wild-type g14-3-3 protein, pThr214-g14-3-3 protein (g14-3-3 protein with phosphorylated Thr214) and T214E-g14-3-3 protein (with phosphomimetic T214E mutation). The authors concluded that in the closed protein conformation (i.e. phosphorylated form) the structural rearrangement at the expense of α8-α9 flexible loop, containing Thr214, is stabilised not so much by interaction between the loop and neighbouring residues, but rather by a steric hindrance of side chains, which provokes a dihedral

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

angles restraint, allowing protein to interact with its partners. These results were fundamental to permit authors to investigate chemical and physico-chemical properties of the interaction region through X-ray crystallographic studies, and finally identify the key hot spot residues of g14-3-3 protein[127].

In an article released in December 2017, Vin Chan et al. reported a MD study on murine double minute 4 protein (MDMX) and tumour suppressor protein p53 interaction. MDMX is a regulation factor of protein p53, which under cellular stress conditions undergoes phosphorylation on Tyrosine 99 (Tyr99 or Y99) residue and releases protein p53, resulting in cell cycle arrest and apoptosis[131–133]. In their work MD is applied to integrate a study previously carried out by Zuckerman et al. [134]. This study suggested that the release of protein p53 from MDMX was caused by a steric bump produced by phosphate group of pTyr99 in p53-binding site. This assumption didn't account the negative feedback mechanism, which occurs upon Tyr99 phosphorylation, whereby another phosphorylation on Tyr55 brings MDMX back to a conformation able to rebind protein p53 and inhibit its activity. Through MD simulations on MDMX-pTyr99 and MDMX-pTyr99-pTyr55, Vin Chan and co-workers suggested that in addition to a steric clash there is a MDMX region, the N-terminal lid, which takes part in protein p53 release through a salt bridge formation between pTyr99 (phosphorylated Tyr99) and Arg18 (or R18) of the lid. Indeed, this salt bridge stabilises the lid in a "close state", preventing MDMX-p53 interaction [133]. Such results were validated by mutagenesis studies using a glutamate residue, as a phosphomimetic[133,135,136]. At the same time, MD simulations on MDMX-pTyr99-pTyr55 unveiled the formation of electrostatic interactions between phosphate group of pTyr55 and three residues of the N-terminus of lid (Met1, Thr2, and Ser3), resulting in a shift of the lid towards an "open state" (away from p53-binding site). The starting structure for MD simulations was built with homology modelling using a PDB file of MDMX (missing N-terminal lid) as template. 200 ns MD simulations were performed as NPT ensemble, using PMEMD module of AMBER14 with ff99SB force field[137,138]. These studies turned out to be a valuable support, because they highlighted likely interactions not yet experimentally discovered, paving the way to further studies for deepening MD information and efficiently designing potential drugs[133].

Unlike classical targets, such us membrane receptors or enzymes, the complex nature of PPIs impacts on designing of modulators, and their chemical and physicochemical features. In most cases, protein-protein interactions do not have natural ligands or known active compounds to be exploited as templates for drug design, rendering hard the hit identification phase of PPI modulators (PPIMs) complementary to receptor binding pocket in terms of shape and chemical attributes[87,88]. Besides, classical virtual screening of drug-like compound libraries against a protein-protein binding site not always is able to provide reliable results both for structure- and fragment-based approaches. Successful PPI modulators, in fact, usually have molecular weights two or three times larger than traditional drugs, and hence they have wider sizes[86,139]. In addition, due to shallowness and quite broad solvent exposure of PPI binding clefts, generally hits show low affinities for protein-protein interactions, with a K_d of 0.1-5 mM[86,140,141]. Generally, PPIMs are classified according to their mechanism of action into disruptor or stabiliser modulators. The PPI disruptors are able to compete in binding one of the two protein partners (orthosteric disruptors) or destabilise a protein-protein interaction through an interaction with a distal or proximal site on protein surface (allosteric disruptors), eliciting a decrease in PP affinity. On the contrary, PPI stabilisers increase protein complex binding affinity and stability either acting directly at interaction interface (orthosteric stabilisation) or binding to a remote site of the protein and causing an increasing of PP affinity[142] (Figure 4). MD has become a valuable tool to validate the stability of a protein-protein-modulator complex and to deepen and unveil PPIMs binding modes. An example of complex stability validation can be found in a recent paper. In this article, Gupta and collaborators[143] showed how MD came to support other techniques, confirming virtual screening results. In particular, the target under consideration is a NH (3)-dependent nicotinamide adenine dinucleotide synthetase protein (GEM_3202 or NadE[144]) involved in a protein-protein interactions network responsible for infective activity of the opportunistic pathogen *Burkholderia cepacia* complex[143,145–148]. Gupta et al. exploited MD simulations to check the stability of the interaction between drug target and two potential hits (ZINC83103551 and ZINC38008121) identified through virtual screening. MD simulations were performed in triplicate on a validated homology model of the protein, using GROMACS 4.5.5 package.

The radius of gyration (Rg) fluctuations were calculated for both NadE-ZINC83103551 and NadE-ZINC38008121 interactions and revealed both complexes showed high structural solidity, obtaining on average 1.87-1.95 nm. Moreover, the measured

View Article Online
DOI: 10.1039/C8MD00166A

RMSD values of protein with ligands were low (~0.41-0.45nm for NadE-ZINC83103551 and ~0.15nm for NadE-ZINC38008121), confirming the high stability of the complexes[143]. In December 2017, another interesting work on PPI modulators design was published by Nath Jha et al.[149].

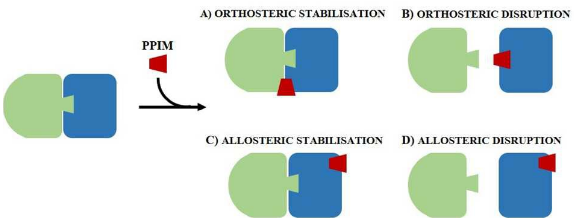


Fig. 4: PPI modulators mechanism of action.

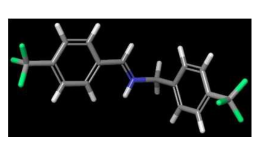
It concerns α-Synuclein (α-Syn) protein, responsible for dopaminergic neurons death in Parkinson's Disease. Recent studies demonstrated that the decrease of oligomeric α-Synuclein species and the acceleration of amyloid fibrils formation represent a potential entry point target strategy for novel drugs design[149–152]. According to this hypothesis, Nath Jha et al. conducted MD studies on three hexapeptides in presence of full-length α-Synuclein. The three peptides were designed on the basis of the hydrophobic region of α-Synuclein responsible for self-association and aggregation, non-amyloid-β component (NAC) region (α-Syn-71-82)[153].

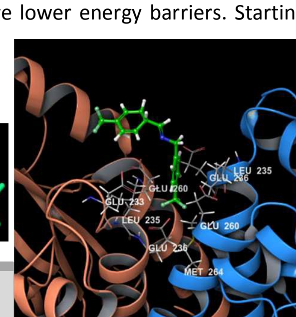
The rationally designed peptide sequences were VAQKTV (or peptide C, corresponding to amino acids 77-82 of NAC), VRQKTV (called R78, due to the mutation A78R), and VPQKTV (named P78, due to the mutation A78P). MD simulations for all three hexapeptides were performed in explicit solvent (TIP3P), keeping the ratio 1:5 (8 full-length α-Synuclein molecules and 40 hexapeptides) and a random orientation without interactions between proteins. After an early step of minimization and equilibration of 20 ns, MD simulations were performed for 100 ns at NPT conditions. The results suggested that the hexapeptide containing arginine (VRQKTV) showed major capacity to establish hydrogen bonds with α-Synuclein, with larger number of H-bond interactions and higher occupancy of contacts than the other two peptides. Furthermore, MD trajectories analysis brought to the identification of contact area between full-length α-Synuclein and peptide R78, i.e. the negatively charged C-terminus of α-Syn and positively charged Arg78 of peptide R78. These results paved the way to further study and design of novel specific peptide ligands able to stimulate α-Syn oligomers aggregation and prevent their cytotoxic activity[149]. Beyond the PPI disruptors and stabilisers, Fischer et al.[142] provided a further category of PPI modulators, that is the modulators of protein dynamics (or interfacial dynamic modulators). These, in fact, do not necessarily affect the binding affinity of a protein-protein complex, but they bind to clefts produced through homo- or hetero-oligomerisation, modifying the dynamic properties of the individual protein partners.

An example of the MD application to support the identification of this type of modulators was provided by Hammoudeh et al.[154] in an article released in 2014. The target considered was the dimeric form of bacterial enzyme dihydropteroate synthase (DHPS), implicated in a key step of folate biosynthetic pathway. The most commonly used drugs for this target are now beginning to show antimicrobial resistance phenomenon and for this reason, targeting DHPS became a health emergence. Hammoudeh et al. were capable to discover an allosteric PPI inhibitor (compound 11, Figure 5), which at low micromolar concentration reduces enzyme K_m and V_{max}. In this study, MD simulations were used to observe the fluctuations of loop1 and loop2, involved in the active site, in four different situations: without any binder, in presence of a natural substrate, with compound 11, and in presence of both natural ligand and inhibitor. MD simulations highlighted that compound 11 was able to bind the dimeric interface of DHPS and loop7 through its distal half.

This interaction causes a conformational change which is transmitted to the enzyme active site, thus reducing considerably loop1 and loop2 fluctuations, responsible for natural substrate binding.

An interesting work concerning thermodynamics and kinetics analysis of a protein-protein association was published by Plattner et al. in 2017[155]. The authors carried out a cutting-edge study using all-atom MD simulations and MSMs to explore states and properties, unlikely experimentally captured, of two proteins known to form a tightly bound complex, barnase and barstar. Plattner et al. performed a large number of aggregate MD simulations of overall 2 milliseconds, consisting in 1.7 milliseconds of individual trajectories and 0.30 milliseconds of multiple parallel adaptive MD runs. One of the main advantages of adaptive MD runs is that allow to speed up biological processes presenting high energy barriers decomposing them into smaller paths with relative lower energy barriers. Starting





View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

from the unbound state of the two proteins, hidden Markov model (HMM) was used to explore proteins states (early intermediates, late intermediates, pre-bound and bound states),

the related interacting contacts, the binding free energies and the transition rates, up to obtain the bound complex of the two protein partners. This latter was a complex form in equilibrium condition between "loosely bound" (5% population) and "tightly bound" states (95% population), with the last one consistent with the crystal structure by PDB database (PDB ID: 1BRS). Average heavy-atoms RMSD values between PDB crystal structure and "loosely bound" form was of 0.3 nm and 0.21 nm in comparison with the "tightly bound" form. This results are remarkably important for medicinal chemists, demonstrating a good reliability of Markov modelling method in reproducing information potentially consistent with experimental data, and exploring the conformational space of a finite number of biological molecules states and the related transition rates.

Conclusions and future perspectives

The main goal of Medicinal Chemistry is identifying chemical entities with optimal affinity, selectivity and safety for patients. Nowadays, computational techniques applied to drug discovery represent an invaluable tool for rationally designing novel chemical entities. Indeed, at present these methodologies have allowed researchers to discover small molecules or small peptides with good drug-like properties[2], speeding up the drug discovery process and decreasing research-associated costs. However, research activities in the field of Medicinal Chemistry have not yet been able to drug a wide range of biological targets, labelling them as undruggable, responsible for several diseases aetiology. Nevertheless, in the last years, there has been a huge rise in the interest for these pharmacological targets together with the increased use of Molecular Dynamics simulations to explore the related binding sites. Medicinal chemists were thus able to deeply investigate shallow and/or buried grooves, particularly those transient, impossible to be detected through the only observation of NMR and X-ray crystal data[12]. In this review, we have provided an overview of successful MD techniques mainly adopted in the last decade, to deepen undruggable targets-related drug discovery issues, such as protein-ligand/protein-protein interaction stability[21–24], binding kinetics[25–29] and interaction mode[30–32]. In particular, we have focused on the application of MD on two specific types of considerably important undruggable targets, allosteric sites and protein-protein interactions.

In this work, we have reported several case studies, whereby MD techniques resulted in the identification of binding cavities, contacts areas, or ligands binding modes. In conclusion, we would like to underline an important point of view as future perspective, recently presented by Pérez et al. in their latter *current opinion paper*[156]. In this work a scenario about the use of MD in synergism with machine learning (ML) approaches is presented. The combination of these two methods seems to be crucial in order to improve predictions accuracy and speed up the MD analysis process. ML/MD combined approach, together with a continuous increasing simulation timescale, as observed by Martinéz-Rossell et al.[157] in their recent overview, depict

Fig. 5: Compound 11 (on the left); Binding mode of compound 11 (green molecule) at dimeric DHPS interface (on the right).

without any doubt an encouraging landscape for researchers, paving the way for "drugging the undruggable targets".

Conflicts of interest

The authors declare no competing interest.

Notes and references

1 S. Mandal, M. Moudgil and S. K. Mandal, *Eur. J. Pharmacol.*, 2009, **625**, 90–100.
2 H. Shirai, C. Prades, R. Vita, P. Marcatili, B. Popovic, J. Xu, J. P. Overington, K. Hirayama, S. Soga, K. Tsunoyama, D. Clark, M. P. Lefranc and K. Ikeda, *Biochim. Biophys. Acta - Proteins Proteomics*, 2014, 1844, 2002–2015.
3 Y. Yuan, J. Pei and L. Lai, *Curr. Pharm. Des.*, 2013, **19**, 2326–33.
4 R. P. Sheridan, V. N. Maiorov, M. K. Holloway, W. D. Cornell and Y. D. Gao, *J. Chem. Inf. Model.*, 2010, **50**, 2029–2040.
5 F. N. B. Edfeldt, R. H. A. Folmer and A. L. Breeze, *Drug Discov. Today*, 2011, **16**, 284–287.
6 C. V. Dang, E. P. Reddy, K. M. Shokat and L. Soucek, *Nat. Rev. Cancer*, 2017, **17**, 502–508.
7 F. McCormick, *Clin. Cancer Res.*, 2015, **21**, 1797–801.
8 J. R. Whitfield, M.-E. Beaulieu and L. Soucek, *Front. cell Dev. Biol.*, 2017, **5**, 10.
9 M. Nayal and B. Honig, *Proteins Struct. Funct.*

View Article Online
DOI: 10.1039/C8MD00166A

Bioinforma., 2006, **63**, 892–906.
10 A. C. Cheng, R. G. Coleman, K. T. Smyth, Q. Cao, P. Soulard, D. R. Caffrey, A. C. Salzberg and E. S. Huang, *Nat. Biotechnol.*, 2007, **25**, 71–75.
11 * Philip J. Hajduk, and Jeffrey R. Huth and S. W. Fesik, , DOI:10.1021/JM049131R.
12 Y. Li and C. Kang, *Molecules*, , DOI:10.3390/molecules22091399.
13 S. P. Leelananda and S. Lindert, *Beilstein J. Org. Chem.*, 2016, **12**, 2694–2718.
14 I. M. Kapetanovic, *Chem. Biol. Interact.*, 2008, **171**, 165–176.
15 S.-S. Ou-Yang, J.-Y. Lu, X.-Q. Kong, Z.-J. Liang, C. Luo and H. Jiang, *Acta Pharmacol. Sin.*, 2012, **33**, 1131–40.
16 C. a Chang, R. Ai, M. Gutierrez and M. J. Marsella, *Methods*, 2012, **819**, 3–12.
17 G. Schneider and U. Fechner, *Nat. Rev. Drug Discov.*, 2005, **4**, 649–63.
18 H. Jhoti and A. R. Leach, *Structure-based drug discovery*, 2007.
19 M. Schneider, X. Fu and A. E. Keating, *Proteins*, 2009, **77**, 97–110.
20 M. De Vivo, M. Masetti, G. Bottegoni and A. Cavalli, *J. Med. Chem.*, 2016, **59**, 4035–4061.
21 M. Wieder, U. Perricone, T. Seidel, S. Boresch and T. Langer, *Monatshefte fur Chemie*, 2016, **147**, 553–563.
22 M. Wieder, U. Perricone, S. Boresch, T. Seidel and T. Langer, *Biochem. Biophys. Res. Commun.*, 2016, **470**, 685–689.
23 U. Perricone, M. Wieder, T. Seidel, T. Langer, A. Padova, A. M. Almerico and M. Tutone, *ChemMedChem*, , DOI:10.1002/cmdc.201600526.
24 M. Gorska-Ponikowska, U. Perricone, A. Kuban-Jankowska, G. Lo Bosco and G. Barone, *J. Cell. Physiol.*, , DOI:10.1002/jcp.25888.
25 S. Decherchi, A. Berteotti, G. Bottegoni, W. Rocchia and A. Cavalli, *Nat. Commun.*, 2015, **6**, 6155.
26 L. Mollica, S. Decherchi, S. R. Zia, R. Gaspari, A. Cavalli and W. Rocchia, *Sci. Rep.*, 2015, **5**, 11539.
27 A. Magistrato, J. Sgrignani, R. Krause and A. Cavalli, *J. Phys. Chem. Lett.*, 2017, **8**, 2036–2042.
28 M. Bernetti, M. Masetti, F. Pietrucci, M. Blackledge, M. R. Jensen, M. Recanatini, L. Mollica and A. Cavalli, *J. Phys. Chem. B*, 2017, **121**, 9572–9582.
29 R. Gaspari, C. Rechlin, A. Heine, G. Bottegoni, W. Rocchia, D. Schwarz, J. Bomke, H.-D. Gerber, G. Klebe and A. Cavalli, *J. Med. Chem.*, 2016, **59**, 4245–4256.
30 A. Cuzzolin, M. Sturlese, G. Deganutti, V. Salmaso, D. Sabbadin, A. Ciancetta and S. Moro, *J. Chem. Inf. Model.*, 2016, **56**, 687–705.
31 G. Deganutti, A. Welihinda and S. Moro, *ChemMedChem*, 2017, **12**, 1319–1326.
32 G. Deganutti and S. Moro, *Molecules*, 2017, **22**, 818.
33 E. Fischer, *Berichte der Dtsch. Chem. Gesellschaft*, 1894, **27**, 2985–2993.
34 B. Ma, M. Shatsky, H. J. Wolfson and R. Nussinov, *Protein Sci.*, 2002, **11**, 184–97.
35 S. J. Teague, *Nat. Rev. Drug Discov.*, 2003, **2**, 527–541.
36 J. Seco, F. J. Luque and X. Barril, *J. Med. Chem.*, 2009, **52**, 2363–2371.
37 K. W. Lexa and H. A. Carlson, 2012, **133**, 200–202.
38 A. Bakan, N. Nevins, A. S. Lakdawala and I. Bahar, *J. Chem. Theory Comput.*, 2012, **8**, 2435–2447.
39 R. Cuchillo, K. Pinto-Gil and J. Michel, *J. Chem. Theory Comput.*, 2015, **11**, 1292–1307.
40 N. Ferruz, S. Doerr, M. A. Vanase-Frawley, Y. Zou, X. Chen, E. S. Marr, R. T. Nelson, B. L. Kormos, T. T. Wager, X. Hou, A. Villalobos, S. Sciabola and G. De Fabritiis, *Sci. Rep.*, 2018, **8**, 1–10.
41 D. E. Koshland, G. Némethy and D. Filmer, *Biochemistry*, 1966, **5**, 365–385.
42 J. Monod, J. Wyman and J.-P. Changeux, *J. Mol. Biol.*, 1965, **12**, 88–118.
43 H. N. Motlagh, J. O. Wrabl, J. Li and V. J. Hilser, *Nature*, 2014, **508**, 331–339.
44 K. A. Henzler-Wildman, V. Thai, M. Lei, M. Ott, M. Wolf-Watz, T. Fenn, E. Pozharski, M. A. Wilson, G. A. Petsko, M. Karplus, C. G. Hübner and D. Kern, *Nature*, 2007, **450**, 838–844.
45 X. Xiao, X. Zeng, Y. Yuan, N. Gao, Y. Guo, X. Pu and M. Li, *Phys. Chem. Chem. Phys.*, 2015, **17**, 2512–22.
46 R. M. Wolf, *J. Comput. Aided. Mol. Des.*, 2015, **29**, 1025–1034.
47 A. Perdih, M. Kotnik, M. Hodoscek and T. Solmajer, *Proteins Struct. Funct. Bioinforma.*, 2007, **68**, 243–254.
48 G. Deganutti, A. Cuzzolin, A. Ciancetta and S. Moro, *Bioorg. Med. Chem.*, 2015, **23**, 4065–4071.
49 J. L. Klepeis, K. Lindorff-Larsen, R. O. Dror and D. E. Shaw, *Curr. Opin. Struct. Biol.*, 2009, **19**, 120–127.
50 M. Karplus and J. A. McCammon, *Nat. Struct. Biol.*, 2002, **9**, 646–652.
51 J. S. Anderson, S. M. Mustafi, G. Hernández and D. M. LeMaster, , DOI:10.1016/j.bpc.2014.06.004.
52 E. E. Boczek, L. G. Reefschläger, M. Dehling, T. J.

View Article Online
DOI: 10.1039/C8MD00166A

Journal Name

REVIEW ARTICLE

Struller, E. Häusler, A. Seidl, V. R. I. Kaila and J. Buchner, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E3189-98.

53 L.-T. Da, F. K. Sheong, D.-A. Silva and X. Huang, in *Advances in experimental medicine and biology*, 2014, vol. 805, pp. 29–66.

54 J. D. Chodera and F. Noé, *Curr. Opin. Struct. Biol.*, 2014, **25**, 135–144.

55 G. R. Bowman, X. Huang and V. S. Pande, *Methods*, 2009, **49**, 197–201.

56 Y. S. Tan, J. Reeks, C. J. Brown, D. Thean, F. J. Ferrer Gago, T. Y. Yuen, E. T. L. Goh, X. E. C. Lee, C. E. Jennings, T. L. Joseph, R. Lakshminarayanan, D. P. Lane, M. E. M. Noble and C. S. Verma, *J. Phys. Chem. Lett.*, 2016, **7**, 3452–3457.

57 L. Hernychova, P. Man, C. Verma, J. Nicholson, C.-A. Sharma, E. Ruckova, J. Y. Teo, K. Ball, B. Vojtesek and T. R. Hupp, *Proteomics*, 2013, **13**, 2512–2525.

58 Q. Luo, L. Chen, X. Cheng, Y. Ma, X. Li, B. Zhang, L. Li, S. Zhang, F. Guo, Y. Li and H. Yang, *Nat. Commun.*, 2017, **8**, 378.

59 D. Sabbadin and S. Moro, *J. Chem. Inf. Model.*, 2014, **54**, 372–376.

60 A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12562–6.

61 G. Grazioso, J. Sgrignani, R. Capelli, C. Matera, C. Dallanoce, M. De Amici and A. Cavalli, *J. Chem. Inf. Model.*, 2015, **55**, 2528–2539.

62 P. Gomez-Gutierrez, J. Rubio-Martinez and J. J. Perez, *J. Chem. Inf. Model.*, 2017, **57**, 2566–2574.

63 D. Hamelberg, J. Mongan and J. A. McCammon, *J. Chem. Phys.*, 2004, **120**, 11919–11929.

64 M. Abdullahi, F. A. Olotu and M. E. Soliman, *Biotechnol. Lett.*, , DOI:10.1007/s10529-017-2432-0.

65 H. Makala and V. Ulaganathan, *J. Recept. Signal Transduct.*, 2017, **0**, 1–8.

66 N. E. Martin, S. Malik, N. Calimet, J. P. Changeux and M. Cecchini, *PLoS Comput. Biol.*, 2017, **13**, 1–25.

67 R. Guan, H. Zhou, J. Li, S. Xiao, C. Pang, Y. Chen, X. Du, S. Ke, Q. Tang, J. Su, Y. Zhan and H. An, 2017, 1–14.

68 L. Jiang, X. Zhang, X. Chen, Y. He, L. Qiao, Y. Zhang, G. Li and Y. Xiang, *Molecules*, 2015, **20**, 12769–12786.

69 N. A. Pabon and C. J. Camacho, *Elife*, 2017, **6**, 1–24.

70 M. Novinec, *PLoS One*, 2017, **12**, 1–22.

71 M. J. Latallo, G. A. Cortina, S. Faham, R. K. Nakamoto and P. M. Kasson, *Chem. Sci.*, 2017, **8**, 6484–6492.

72 H. N. Motlagh, J. O. Wrabl, J. Li and V. J. Hilser, *Nature*, 2014, **508**, 331–339.

73 A. D. Caliman, Y. Miao and J. A. McCammon, *Chem. Biol. Drug Des.*, 2017, 5–16.

74 A. D. Caliman, S. E. Swift, Y. Wang, Y. Miao and J. A. McCammon, *Protein Sci.*, 2015, **24**, 1004–12.

75 D. Bartuzi, A. A. Kaczor and D. Matosiuk, *J. Chem. Inf. Model.*, 2015, **55**, 2421–2434.

76 D. Bartuzi, A. A. Kaczor and D. Matosiuk, *J. Chem. Inf. Model.*, 2016, **56**, 563–570.

77 W. Liu, E. Chun, A. A. Thompson, P. Chubukov, F. Xu, V. Katritch, G. W. Han, C. B. Roth, L. H. Heitman, A. P. IJzerman, V. Cherezov and R. C. Stevens, *Science*, 2012, **337**, 232–6.

78 Y. Shang, V. LeRouzic, S. Schneider, P. Bisignano, G. W. Pasternak and M. Filizola, *Biochemistry*, 2014, **53**, 5140–5149.

79 Y. Miao, A. D. Caliman and J. A. McCammon, *Biophys. J.*, 2015, **108**, 1796–1806.

80 K. E. Livingston and J. R. Traynor, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 18369–74.

81 Y. Meng, C. Gao, D. K. Clawson, S. Atwell, M. Russell, M. Vieth and B. Roux, *Predicting the Conformational Variability of Abl tyrosine kinase using Molecular Dynamics Simulations and Markov State Models*, 2018.

82 D. C. Fry, *Biopolymers*, 2006, **84**, 535–552.

83 K. Venkatesan, J. Rual, A. Vazquez, I. Lemmens, T. Hirozane-kishikawa, T. Hao, X. Xin, K. Goh, M. A. Yildirim, K. Heinzmann, F. Gebreab, J. M. Sahalie, C. Simon, A. De Smet, E. Dann, A. Vinayagam, H. Yu, D. Szeto, A. Dricot, N. Klitgord, R. R. Murray, C. Lin, M. Lalowski, J. Timm, K. Rau, C. Boone, P. Braun, E. Michael, F. P. Roth, D. E. Hill, J. Tavernier, E. E. Wanker, L. Barabási and M. Vidal, *Syst. Biol. (Stevenage).*, 2010, **6**, 83–90.

84 Q. C. Zhang, D. Petrey, L. Deng, L. Qiang, Y. Shi, C. A. Thu, B. Bisikirska, C. Lefebvre, D. Accili, T. Hunter, T. Maniatis, A. Califano and B. Honig, *Nature*, 2012, **490**, 556–560.

85 M. P. H. Stumpf, T. Thorne, E. de Silva, R. Stewart, H. J. An, M. Lappe and C. Wiuf, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 6959–64.

86 A. E. Modell, S. L. Blosser and P. S. Arora, *Trends Pharmacol. Sci.*, 2016, **37**, 702–713.

87 A. Metz, E. Ciglia and H. Gohlke, *Curr. Pharm. Des.*, 2012, **18**, 4630–4647.

88 C. Rakers, M. Bermudez, B. G. Keller, J. Mortier and G. Wolber, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2015,

View Article Online
DOI: 10.1039/C8MD00166A

5, 345–359.

89 E. Cukuroglu, H. B. Engin, A. Gursoy and O. Keskin, *Prog. Biophys. Mol. Biol.*, 2014, **116**, 165–173.

90 X. Zhang, S. Betzi, X. Morelli and P. Roche, *Future Med. Chem.*, 2014, **6**, 1291–1307.

91 D. R. Koes, A. Dömling and C. J. Camacho, *Protein Sci.*, 2018, **27**, 229–232.

92 D. Kozakov, L. E. Grove, D. R. Hall, T. Bohnuud, S. E. Mottarella, L. Luo, B. Xia, D. Beglov and S. Vajda, *Nat. Protoc.*, 2015, **10**, 733–755.

93 Y. S. Tan, D. R. Spring, C. Abell and C. S. Verma, *J. Chem. Theory Comput.*, 2015, **11**, 3199–3210.

94 T. Clackson and J. A. Wells, *Science*, 1995, **267**, 383–6.

95 R. Sable and S. Jois, *Molecules*, 2015, **20**, 11569–11603.

96 A. A. Bogan and K. S. Thorn, *J. Mol. Biol.*, 1998, **280**, 1–9.

97 I. S. Moreira, P. A. Fernandes and M. J. Ramos, *Proteins Struct. Funct. Bioinforma.*, 2007, **68**, 803–812.

98 L. Lo Conte, C. Chothia and J. Janin, *J. Mol. Biol.*, 1999, **285**, 2177–98.

99 J. Janin and C. Chothia, *J. Biol. Chem.*, 1990, **265**, 16027–30.

100 N. Horton and M. Lewis, *Protein Sci.*, 1992, **1**, 169–81.

101 L. D. Walensky and G. H. Bird, *J. Med. Chem.*, 2014, **57**, 6275–88.

102 S. Kellenberger and L. Schild, *Physiol. Rev.*, 2002, **82**, 735–767.

103 M. Izumi, M. Ikeuchi, Q. Ji and T. Tani, *J. Biomed. Sci.*, 2012, **19**, 77.

104 E. Deval, X. Gasull, J. Noël, M. Salinas, A. Baron, S. Diochot and E. Lingueglia, *Pharmacol. Ther.*, 2010, **128**, 549–558.

105 Z.-G. Xiong, X.-M. Zhu, X.-P. Chu, M. Minami, J. Hey, W.-L. Wei, J. F. MacDonald, J. A. Wemmie, M. P. Price, M. J. Welsh and R. P. Simon, *Cell*, 2004, **118**, 687–698.

106 P. Escoubas, J. R. De Weille, A. Lecoq, S. Diochot, R. Waldmann, G. Champigny, D. Moinier, A. Ménez and M. Lazdunski, *J. Biol. Chem.*, 2000, **275**, 25116–25121.

107 T. W. Sherwood, K. G. Lee, M. G. Gormley and C. C. Askwith, *J. Neurosci.*, 2011, **31**, 9723–9734.

108 N. J. Saez, E. Deplazes, B. Cristofori-Armstrong, I. R. Chassagnon, X. Lin, M. Mobli, A. E. Mark, L. D. Rash and G. F. King, *Br. J. Pharmacol.*, 2015, **172**, 4985–95.

109 R. J. P. Dawson, J. Benz, P. Stohler, T. Tetaz, C. Joseph, S. Huber, G. Schmid, D. Hügin, P. Pflimlin, G. Trube, M. G. Rudolph, M. Hennig and A. Ruf, *Nat. Commun.*, 2012, **3**, 936.

110 R. Biswas, S. Ghosh and A. Bagchi, *J. Mol. Recognit.*, 2017, **30**, e2643.

111 R. Li, L. Huang, H. Guo and B. P. Toole, *J. Cell. Physiol.*, 2001, **186**, 371–379.

112 C.-G. Liao, L. Yao, W. Xie, L. Liu, S.-D. Wu, N. Lu, J.-G. Huang, L.-M. Kong and H.-L. Zhang, *Cancer Cell Int.*, 2016, **16**, 28.

113 Z. Luo, X. Zhang, W. Zeng, J. Su, K. Yang, L. Lu, C. B. Lim, W. Tang, L. Wu, S. Zhao, X. Jia, C. Peng and X. Chen, *Oncotarget*, 2016, **7**, 7179–92.

114 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.

115 Y. L. Xue, L. Zhou, Y. Sun, H. Li, G. W. Jones and Y. Song, *J. Mol. Model.*, , DOI:10.1007/s00894-017-3453-2.

116 A. L. Fink, *Physiol. Rev.*, 1999, **79**, 425–449.

117 P. Walsh, D. Bursać, Y. C. Law, D. Cyr and T. Lithgow, *EMBO Rep.*, 2004, **5**, 567–71.

118 J. Kalikka and J. Akola, *Eur. Biophys. J.*, 2011, **40**, 181–194.

119 A. González, T. Perez-Acle, L. Pardo and X. Deupi, *PLoS One*, 2011, **6**, e23815.

120 L. Xu, N. Hasin, M. Shen, J. He, Y. Xue, X. Zhou, S. Perrett, Y. Song and G. W. Jones, *PLoS Comput. Biol.*, 2013, **9**, e1002896.

121 M. Shen, J. Guan, L. Xu, Y. Yu, J. He, G. W. Jones and Y. Song, *J. Biomol. Struct. Dyn.*, 2012, **30**, 652–661.

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

123 G. Martinez-Rosell, M. J. Harvey and G. De Fabritiis, *J. Chem. Inf. Model.*, 2018, acs.jcim.7b00625.

124 W.-T. Choi, Y. Yang, Y. Xu and J. An, *Curr. Top. Med. Chem.*, 2014, **14**, 1574–89.

125 M. M. R. Arkin and J. A. Wells, *Nat. Rev. Drug Discov.*, 2004, **3**, 301–317.

126 M. C. Halliez and A. G. Buret, *World J. Gastroenterol.*, 2013, **19**, 8974.

127 Y. Cau, A. Fiorillo, M. Mori, A. Ilari, M. Botta and M. Lalle, *J. Chem. Inf. Model.*, 2015, **55**, 2611–22.

128 M. Lalle, *Infect. Disord. Drug Targets*, 2010, **10**, 283–94.

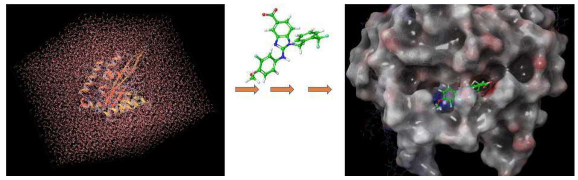
129 M. Lalle, C. Bavassano, F. Fratini, S. Cecchetti, P. Boisguerin, M. Crescenzi and E. Pozio, *Int. J. Parasitol.*, 2010, **40**, 201–213.

130 A. Fiorillo, D. di Marino, L. Bertuccini, A. Via, E. Pozio, S. Camerini, A. Ilari and M. Lalle, *PLoS One*, 2014, **9**, e92902.

View Article Online
DOI: 10.1039/C8MD00166A

131 D. Lane and A. Levine, *Cold Spring Harb. Perspect. Biol.*, 2010, **2**, 1–10.
132 A. Shvarts, W. T. Steegenga, N. Riteco, T. van Laar, P. Dekker, M. Bazuine, R. C. van Ham, W. van der Houven van Oordt, G. Hateboer, A. J. van der Eb and A. G. Jochemsen, *EMBO J.*, 1996, **15**, 5349–57.
133 J. V. Chan, D. X. P. Koh, Y. Liu, T. L. Joseph, D. P. Lane, C. S. Verma and Y. S. Tan, *Oncotarget*, 2017, **8**, 112825–112840.
134 V. Zuckerman, K. Lenos, G. M. Popowicz, I. Silberman, T. Grossman, J. C. Marine, T. A. Holak, A. G. Jochemsen and Y. Haupt, *J. Biol. Chem.*, 2009, **284**, 4031–4039.
135 N. Dissmeyer and A. Schnittger, Humana Press, 2011, pp. 93–138.
136 S. R. Stateva, V. Salas, G. Benaim, M. Menéndez, D. Solís and A. Villalobo, *PLoS One*, 2015, **10**, 1–20.
137 D. Case, V. Babin, J. Berryman, R. Betz, Q. Cai, D. Cerutti, T. Cheatham, T. Darden, R. Duke, H. Gohlke, A. Götz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.-S. Lee and P. A. Kollman, , DOI:10.13140/RG.2.2.17892.37766.
138 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins Struct. Funct. Bioinforma.*, 2006, **65**, 712–725.
139 O. Sperandio, C. H. Reynès, A.-C. Camproux and B. O. Villoutreix, *Drug Discov. Today*, 2010, **15**, 220–229.
140 E. Valkov, T. Sharpe, M. Marsh, S. Greive and M. Hyvönen, in *Topics in current chemistry*, 2011, vol. 317, pp. 145–179.
141 T. V. Magee, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 2461–2468.
142 G. Fischer, M. Rossmann and M. Hyvönen, *Curr. Opin. Biotechnol.*, 2015, **35**, 78–85.
143 M. Gupta, R. Chauhan, Y. Prasad, G. Wadhwa and C. K. Jain, *Comput. Biol. Chem.*, 2016, **65**, 80–90.
144 M. Gupta, R. Chauhan, Y. Prasad, G. Wadhwa and C. K. Jain, *Comput. Biol. Chem.*, 2016, **65**, 80–90.
145 I. K. Jordan, I. B. Rogozin, Y. I. Wolf and E. V. Koonin, *Genome Res.*, 2002, **12**, 962–968.
146 P. Bork, T. Dandekar, Y. Diaz-Lazcoz, F. Eisenhaber, M. Huynen and Y. Yuan, *J. Mol. Biol.*, 1998, **283**, 707–725.
147 M. Remm, C. E. V. Storm and E. L. L. Sonnhammer, *J. Mol. Biol.*, 2001, **314**, 1041–1052.
148 P. Goymer, *Nat. Rev. Genet.*, 2008, **9**, 650.
149 N. N. Jha, S. Ranganathan, R. Kumar, S. Mehra, R. Panigrahi, A. Navalkar, D. Ghosh, A. Kumar, R. Padinhateeri and S. K. Maji, *Biochemistry*, 2017, acs.biochem.7b01090.
150 K. A. Conway, S. J. Lee, J. C. Rochet, T. T. Ding, R. E. Williamson and P. T. Lansbury, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 571–6.
151 J. Li, V. N. Uversky and A. L. Fink, *Biochemistry*, 2001, **40**, 11604–13.
152 M.-B. Fares, N. Ait-Bouziad, I. Dikiy, M. K. Mbefo, A. Jovi i , A. Kiely, J. L. Holton, S.-J. Lee, A. D. Gitler, D. Eliezer and H. A. Lashuel, *Hum. Mol. Genet.*, 2014, **23**, 4491–4509.
153 B. I. Giasson, I. V. J. Murray, J. Q. Trojanowski and V. M.-Y. Lee, *J. Biol. Chem.*, 2001, **276**, 2380–2386.
154 D. I. Hammoudeh, M. Date, M.-K. Yun, W. Zhang, V. A. Boyd, V. Follis, E. Griffith, R. E. Lee, D. Bashford and S. W. White, , DOI:10.1021/cb500038g.
155 N. Plattner, S. Doerr, G. De Fabritiis and F. Noé, *Nat. Chem.*, 2017, **9**, 1005–1011.
156 J. Jiménez, M. Škalič, G. Martínez-Rosell and G. De Fabritiis, *J. Chem. Inf. Model.*, 2018, **58**, 287–296.
157 G. Martínez-Rosell, T. Giorgino, M. J. Harvey and G. de Fabritiis, *Curr. Top. Med. Chem.*, 2017, **17**, 2617–2625.

View Article Online
DOI: 10.1039/C8MD00166A



Molecular Dynamics has been demonstrated to be crucial for unveiling otherwise hidden binding sites especially for undruggable targets challenge