European Journal of Medicinal Chemistry 91 (2015) 4–14

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Review article

Molecular dynamics in drug design

Hongtao Zhao, Amedeo Caflisch*

Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland

ARTICLE INFO

Article history:
Received 10 June 2014
Received in revised form
31 July 2014
Accepted 3 August 2014
Available online 4 August 2014

Keywords:
Atomistic simulation
In silico screening
High-throughput docking
Bromodomains
Tyrosine kinases
Proteases

ABSTRACT

Molecular dynamics (MD) simulations are useful tools for structure-based drug design. We review recent publications in which explicit solvent MD was used at the initial or final stages of high-throughput docking campaigns. In some cases, MD simulations of the protein target have been carried out before docking to generate a conformer of the protein which differs from the available crystal structure(s). Furthermore, MD runs have been performed after docking to assess the predicted binding modes of the top ranking compounds as final filter in silico or to guide chemical synthesis for hit optimization. We present examples of in silico discoveries of tyrosine kinase inhibitors and bromodomain antagonists whose binding mode was predicted by automated docking and further corroborated by MD simulations with final validation by X-ray crystallography.

1. Introduction

While atomistic MD simulations are still computationally expensive for docking large libraries of compounds, their application for hit discovery and optimization is increasing steadily. Here we review in silico screening campaigns in which MD played a key role in the identification of small molecules that bind to protein targets. The focus is on explicit solvent MD simulations carried out in our research group to prepare a structure for docking or to assess predicted binding modes. The high-throughput campaigns performed in our group are summarized in Table 1, and representative hits are shown in Fig. 1 [1–15]. MD simulations were employed at the protein-preparation stage or final scoring step in the campaigns that led to the identification of the enzyme inhibitors **5**, **11**–**17** [5,11–14] and the bromodomain ligands **18**–**19** [15]. Several of the hits identified in silico, e.g., compounds **7** and **10**–**11**, have been advanced into series of potent and selective tyrosine kinase inhibitors which are promising pre-clinical candidates [11,16,17].

2. MD as a tool for mapping molecular fragments to binding sites

2.1. Pioneering studies and recent developments

Almost 30 years ago, Peter J. Goodford pioneered the use of molecular fragments to map protein binding sites by calculating the interaction energy on a grid around the protein surface [18]. In 1991, Andrew Miranker and Martin Karplus proposed energy minimization as a simple and efficient method for generating functionality maps, that is optimal positions and orientations of functional groups in the binding site of a protein target [19]. The method was called multiple copies simultaneous search (MCSS) because during minimization the interaction energy between multiple replicas of a molecular fragment was switched off so that each replica feels only the force field of the protein. Some of the biophysical methods for fragment-based lead identification, such as structure–activity relationship by nuclear magnetic resonance spectroscopy [20], are similar in principle to the MCSS approach. It is interesting to note that computational methods for generating functionality maps [18,19,21,22] have preceded experimental methods that report on the binding of small molecules. The experimental techniques include X-ray crystallography [23], nuclear magnetic resonance spectroscopy [24], surface plasmon resonance [25], mass spectrometry [26,27], substrate activity screening (where the fragments are substrates later converted into inhibitors [28–30]), and tethering [31,32].

Recently, MD simulations have been employed for determining the binding modes of small aliphatic and aromatic molecules into the oncoprotein BCL-6 [33] and isopropyl alcohol into five different proteins [34]. These studies were published 18 years after the minimization-based MCSS protocol of Miranker and Karplus, which in principle allowed also for binding site flexibility by a combination of MD and energy minimization. The MD protocol developed recently in MacKerell's group is called SILCS (Site-Identification by

* Corresponding author.
E-mail address: caflisch@bioc.uzh.ch (A. Caflisch).

http://dx.doi.org/10.1016/j.ejmech.2014.08.004
0223-5234/

Table 1
Structure-based virtual screening campaigns performed at the Department of Biochemistry of the University of Zurich during the decade 2005–2014.

	Protein	Representative hits[a]			Hit rate[b] (%)	Scoring method	Ref.	PDB entry
		No.	Affinity (μM)	LE				
Proteases	β-secretase	**1**	3.0	0.19	17	LIECE[c]	[1]	NA
	β-secretase	**2**	7.1	0.19	10	LIECE	[2]	NA
	Plasmepsin	**3**	2.0	0.25	32	Consensus	[3]	NA
	NS3 protease	**4**	40	0.33	5	LIECE	[4]	NA
	NS3 protease	**5**	2.8	0.34	40	Filtering	[5]	NA
	Cathepsin B	**6**	4.8	0.29	3	Consensus	[6]	NA
Kinases	EphB4	**7**	1.5	0.32	19	LIECE	[7]	4GK3 [d]
	CDK2	**8**	7.8	0.32	3	LIECE	[8]	NA
	EphB4	**9**	2.0	0.31	13	Filtering	[9]	NA
	EphB4	**10**	0.3	0.35	44	FFES[e]	[10]	4P4C
	EphB4	**11**	5.2	0.30	67	FFES	[11]	4G2F
	Abl1, EphA3	**12**	3.9	0.22	25	FFES	[12]	NA
	ZAP70	**13**	21	0.26	31	FFES	[13]	NA
	ZAP70	**14**	14	0.25				NA
	JAK2	**15**	0.1	0.37				NA
	SYK	**16**	23	0.21	9	FFES	[14]	NA
	SYK	**17**	18	0.22				NA
Bromodomains	BRD4	**18**	7.0	0.37	17	FFES	[15]	4PCE
	BRD4	**19**	7.5	0.37				4PCI
	CREBBP	–	5.0	0.37	12	SEED [47]	[f]	4TQN

[a] Compound with highest measured affinity obtained directly from high-throughput docking and scoring. The chemical structures of these compounds are shown in Fig. 1 while the binding modes of compounds **18** and **19** are shown in Fig. 6(C) and (D), respectively. The ligand efficiency (LE) is the measured affinity divided by the number of non-hydrogen atoms and has units of kcal/mol per heavy atom.
[b] Percentage of active compounds i.e., compounds with measured affinity <100 μM divided by the number of molecules tested experimentally. The number of compounds tested ranges from 5 (in the 2nd campaign for NSGH protease) to 88 (in the 2nd campaign for β-secretase).
[c] LIECE: linear interaction energy model with continuum electrostatics [66].
[d] Hit optimization by chemical synthesis of derivatives resulted in low nM inhibitors [16,17] (see also PDB entries 4GK2 and 4GK4).
[e] FFES: force-field energy with electrostatic solvation evaluated by numerical solution of the finite-difference Poisson equation using the continuum dielectric approximation [10].
[f] Min Xu et al. unpublished results.

Ligand Competitive Saturation), and as in MCSS the attractive interactions between fragments are switched off. This simulation stratagem makes possible the use of a very high concentration even for hydrophobic fragments, which would otherwise aggregate in the simulation box [35]. We note *en passant* that this is an interesting example in which a simulation protocol allows one to study a molecular system under conditions that are not accessible by experiments.

2.2. Thermodynamics and kinetics of small molecule binding to proteins from MD simulations

We have proposed MD as a tool to analyze the free-energy surface and pathways of (un)binding of small molecules from/to proteins [36,37]. Because of the available crystal structures and measured binding affinities we have applied MD to the peptidyl-prolyl cis-trans isomerase called FKBP (the FK506 binding protein) and six ligands which have between four (dimethylsulphoxide) and eleven (5-diethylamino-2-pentanone) non-hydrogen atoms. Their affinity for FKBP is in the high μM to low mM range as measured by tryptophan fluorescence quenching assay [38]. For each ligand, a conformational space network [39] of the binding process was generated (Fig. 2). In a first step, the relative position and orientation saved along multiple trajectories were clustered according to a set of intermolecular distances. The clusters were considered as nodes of a network, and the direct transitions between these clusters observed during MD were the links of the network. Interestingly, the network analysis revealed multiple binding modes characterized by distinct intermolecular hydrogen bonds and hydrophobic contacts (Fig. 2). Moreover, the unbinding kinetics showed single-exponential time dependence which indicates that the barrier for full dissociation is significantly higher than the barriers between different binding modes. It is instructive to compare experimental and simulation approaches. The aforementioned biophysical techniques for the analysis of fragment binding to proteins have limitations in temporal and/or spatial resolution. In contrast, the MD simulations of (un)binding generate a complete picture of the free-energy surface and (un)binding pathways at atomic level of detail [36,37,40].

Bromodomains are α-helical bundles of approximately 110 residues, which bind acetylated lysine side chains mainly on histone tails [41]. Some of the 61 human bromodomains have been involved in cancer and inflammation. We have carried out MD simulations of two bromodomains (BAZ2B and CREBBP) to assess the structural stability of the six water molecules that seem to be conserved at the bottom of the acetyl-lysine binding site in most crystal structures of bromodomains [42]. The MD runs revealed that the occupancy of the structured water molecules is influenced by the flexibility of the loop connecting helices Z and A (Fig. 3). Additional simulations in the presence of high concentration of cosolvent (i.e., 440 mM of dimethylsulfoxide, methanol, or ethanol) revealed that some of the structured water molecules can be displaced transiently [42]. This observation is consistent with two recently disclosed crystal structures of the fifth bromodomain of human Poly-bromodomain containing protein 1 (PB1) in complex with hydroxyphenyl-propenone ligands (PDB codes 4Q0N and 4Q0O) which show that the phenyl ring of the ligand can replace some of the water molecules at the bottom of the binding pocket.

Two main observations emerge from our MD studies of molecular fragments and cosolvent binding to FKBP and bromodomains: (1) the presence of metastable states corresponding to multiple binding poses, and (2) the importance of solvent molecules in molecular recognition. These two features are fully captured by atomistic, explicit solvent MD simulations while they are difficult to