Journal of Molecular Graphics and Modelling 45 (2013) 192–201

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Exploring the effect of PARP-1 flexibility in docking studies

Albert A. Antolin[a,1], Andrea Carotti[b,1], Roberto Nuti[b,1], Aydie Hakkaya[b], Emidio Camaioni[b], Jordi Mestres[a], Roberto Pellicciari[b,c], Antonio Macchiarulo[b,*]

[a] Chemogenomics Laboratory, Research Program in Biomedical Informatics (GRIB), IMIM Hospital del Mar Research Institute and Universitat Pompeu Fabra, Doctor Aiguader 88, 08003 Barcelona, Catalonia, Spain
[b] Dipartimento di Chimica e Tecnologia del Farmaco, Università di Perugia, via del Liceo 1, 06123 Perugia, Italy
[c] TES Pharma S.r.l., via Palmiro Togliatti 22bis, 06073 Corciano, Perugia, Italy

ARTICLE INFO

Article history:
Accepted 6 August 2013
Available online 3 September 2013

Keywords:
Replica exchange molecular dynamics
PARP
Molecular docking
Virtual screening

ABSTRACT

Poly(ADP-ribose)polymerase-1 (PARP-1) is an enzyme belonging to the ADP-ribosyltransferase family. A large body of works has validated PARP-1 as an attractive drug target for different therapeutic areas, including cancers and ischemia. Accordingly, sampling the conformational space of the enzyme is pivotal to understand its functions and improve structure-based drug discovery approaches.

In the first part of this study we apply replica exchange molecular dynamic (REMD) simulations to sample the conformational space of the catalytic domain of PARP-1 in the ligand-bound and unbound forms. In the second part, we assess how and to what extend the emerging enzyme flexibility affects the performance of docking experiments of a library of PARP-1 inhibitors. This study pinpoints a putative key role of conformational shifts of Leu324, Tyr325 and Lys242 in opening an additional binding site pocket that affects the binding of ligands to the catalytic cleft of PARP-1. Furthermore, it highlights the improvement of the enrichment factor of active ligands obtained in docking experiments when using conformations generated with REMD simulations of ligand-bound PARP-1.

1. Introduction

Poly(ADP-ribose)polymerase-1 (PARP-1) is the founder member of the mammalian PARP superfamily [1–3]. Upon activation by DNA damage, PARP-1 catalyzes the attachment of poly ADP-ribose units from NAD+ to target proteins, with ensuing modifications of their functions [4]. In such a way, the enzyme is involved in genome repairing, chromatin remodeling, and cell survival or death pathways [5–8].

The scientific interest in PARP-1 is continuously growing since the development of PARP inhibitors is being pursued as therapeutic opportunity in a range of diseases including stroke, cardiac ischemia, cancer, inflammation and diabetes [9–12].

Despite part of these efforts have proved successful with some PARP-1 inhibitors entering in clinical trials and others being available as chemical tools [13], recent data point out that these compounds generally lack binding specificity at members of the PARP family [14]. As a consequence, academic groups and pharmaceutical companies have engaged in the quest of novel and more selective PARP inhibitors [15,16]. Taking advantage from recent progress in the structure determination of PARP domains [2], some of these efforts rely on the use of structure-based virtual screening approaches to disclose novel chemical classes of enzyme inhibitors [17–20], though the effect of PARP flexibility on the performance of docking experiments has still remained elusive.

As a continuation of our previous efforts in the design, synthesis and biological appraisals of PARP-1 inhibitors [21–25], in this work we investigate the conformational space of the catalytic domain of PARP-1 and how and to what extend the emerging binding site flexibility may affect docking experiments of a library of PARP-1 inhibitors. To this end, the present work is composed of two parts. In the first part, replica exchange molecular dynamics (REMD) are used to sample the conformational space of the catalytic domain of PARP-1 in the ligand-bound and unbound forms. Using REMD, many replicas of the system are simulated at the same time at different temperatures and there is an exchange of coordinates and velocities between a pair of adjacent temperature simulations, according to a Metropolis criterion [26–29]. Therefore, system conformations at higher temperature can swap to lower temperature, increasing the rate of conformational sampling at that temperature. The efficiency of REMD has been recently proved in sampling ligand induced fit conformations of target [30], and in generating binding site conformations with different geometries for virtual screening experiments [31].

Abbreviations: PARP, poly(ADP-ribose) polymerase; REMD, replica exchange molecular dynamics; PCA, principal components analysis; RMSD, root mean square deviation; EF, enrichment factor.

* Corresponding author. Tel.: +39 075 585 5160; fax: +39 075 585 5114.
E-mail address: antonio@chimfarm.unipg.it (A. Macchiarulo).

[1] These authors contributed equally to the work.

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.08.006

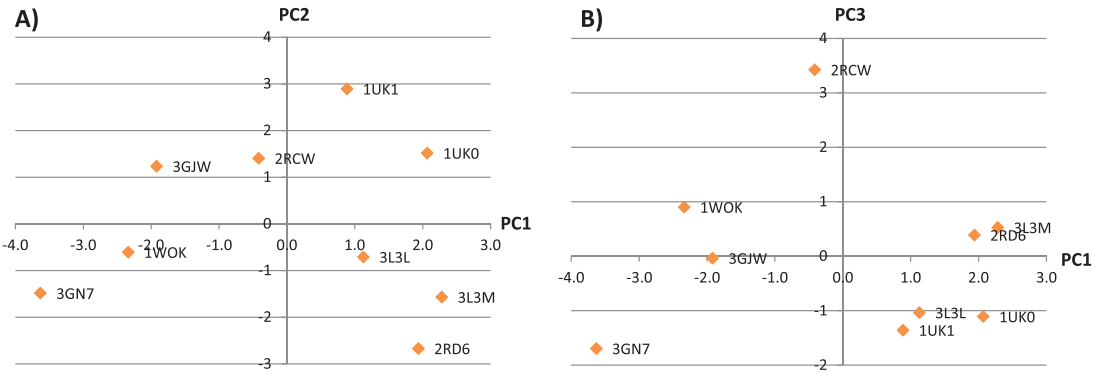

Fig. 1. Plots of PC1 and PC2 components (A), and PC1 and PC3 components (B) of PCA-1.

In the second part of the work, PARP-1 conformations with different properties of the catalytic cleft are selected and instrumental to perform ensemble docking experiments of a library of ligands including known PARP-1 inhibitors. The results are then discussed in terms of improvement of the enrichment factor when using ensemble of structures obtained from REMD simulations with respect to crystal structures.

2. Methods

2.1. Replica exchange molecular dynamic simulations

Crystal structures of PARP-1 were preliminary processed using Schrödinger's "Protein Preparation Wizard" (Schrödinger, LLC, New York, NY). In particular, the following entries of the protein databank [32] were retrieved and processed (PDB ID): 1UK0, 1UK1, 1WOK, 2RD6, 3GJW, 3GN7, 3L3M, 3L3L and 2RCW. Structures were processed by removing water molecules, adding hydrogen atoms, and assigning atom types. For all of the structures a brief minimization step was performed using the OPLS-2005 force field in order to remove steric clashes wherever present. The minimization process terminated when the root mean square deviation (RMSD) reached a maximum of 0.30 Å as calculated on the heavy-atoms.

In order to select the input structure for following REMD simulations, the binding site properties of the nine entries were analyzed as described in Section 2.2 using principal component analysis and binding site descriptors (PCA-1). The inspection of the loading plots of the first three components (Fig. 1) shows that eight of the nine crystal structures form two groups, with the remaining entry (PDB ID: 2RCW) lying in the middle of them. As a consequence, we deemed 2RCW an unbiased input structure for REMD simulations, insofar showing binding site properties in between those of the two groups of structures. REMD simulations were then performed using both the ligand-bound form (PDB ID: 2RCW) and the unbound form of the catalytic domain of PARP-1, removing the inhibitor co-crystallized in the entry 2RCW.

For each of the ligand-bound and unbound system, the protein was solvated in a truncated octahedron box with TIP3P water molecules extending 10 Å beyond the protein in all directions, and sodium chloride ions were added with ionic strength of 0.15 M, applying periodic boundary conditions. The final systems were composed of 38,944 atoms (PARP-1 ligand-bound system) and of 38,925 atoms (PARP-1 unbound system).

REMD simulations were performed using *Desmond* 2.4 software and the force field OPLS-2005. The simulation protocol included a starting relaxation step using *Desmond*, and a final production phase of 10 ns. In particular, the protocol comprised an initial minimization of the system over a maximum 2000 steps, with a convergence criterion of 50 kcal/mol/Å, and the presence of harmonic restrains on the solute atoms (force constant = 50.0 kcal/mol/Å^2); a second minimization without restrains; a third stage of 12 ps at 10 K with harmonic restrains on the solute heavy atoms (force constant = 50.0 kcal/mol/Å^2), using NVT ensemble and Berendsen thermostat [33]; a fourth 12 ps at 10 K, retaining the harmonic restrains, and using NPT ensemble and Berendsen thermostat and barostat; a fifth heating phase of 24 ps at 300 K, retaining the harmonic restrains, and using NPT ensemble and Berendsen thermostat and barostat; a final 24 ps at 300 K without harmonic restrains, using NVT ensemble and Nose-Hoover thermostat and Martyna–Tobias–Klein barostat [34,35]. The final production phase of REMD was run using a canonical NVT ensemble with a Nose-Hoover thermostat [35,36]. The selection of a canonical thermostat was done on the basis of the results reported by Rosta and coworkers showing that REMD simulations distort the configuration-space distributions when combined with thermostats that do not produce a canonical ensemble [37].

In this final step, REMD simulations were performed in a range of temperature of 300–315 K. Although this narrow range of temperature may reduce a thorough sampling of the conformational space of the enzyme, it was chosen to avoid protein unfolding at higher temperatures. Indeed, the catalytic domain of PARP-1 unfolded with a T_m of 48.6 ± 0.2 °C (321.6 ± 0.2 K) in buffer (means and standard deviations based on two replicates of triplicate experiments), and a T_m of 48.8 ± 0.1 °C (321.8 ± 0.1 K) in the presence of 1% DMSO [21 and E. Wahlberg and H. Schüler, personal communication].

The optimal number of replicas for each system was determined as a function of its size according to the algorithm proposed by Patriksson [38]. A number of 11 replicas were thus used for the ligand-bound and unbound forms of PARP-1, spanning from 300 K to 315 K (replica temperatures: 300 K, 301.641 K, 303.258 K, 304.91 K, 306.565 K, 308.212 K, 309.856 K, 311.504 K, 313.143 K, 314.826 K, 315 K) and resulting in 10 ns × 11 (replicas) of total simulation time. During REMD simulations, we used a time step of 2 fs. Atomic coordinates were saved every 4.8 ps, and REMD exchanges were attempted every 9 ps. The plot of the 11 replicas as a function of time is reported in Figs. s2 and s3 of the supporting information, showing the accepted exchanges from neighboring replicas (112 exchanges for ligand-bound PARP-1; 120 exchanges for the unbound PARP-1).

The trajectory collected at 300 K from REMD simulations was used for the following analysis, containing a total of 2084 conformations of ligand-bound and unbound PARP-1. Before analyzing the trajectory, ligand and water molecules were removed from each of the collected PARP-1 conformations. RMSD of secondary structure elements and binding site residues were calculated from the input structure (PDB ID: 2RCW) using backbone atoms and heavy atoms, respectively.

Table 1
Structural resolution factors and selected binding site properties of the catalytic cleft of the crystal structures of PARP-1.

PDB code	Res. (Å)	Contact	Exposure	Balance	Donors/acceptor	Volume (Å^3)	D-score
1UK0	3.0	0.975	0.498	0.904	1.037	620.490	1.088
1UK1	3.0	0.976	0.492	0.849	1.217	630.780	1.095
1WOK	3.0	0.958	0.446	0.535	1.085	673.310	1.051
2RD6	2.3	0.947	0.519	0.793	0.687	731.960	1.050
3GJW	2.3	0.973	0.466	0.654	1.055	737.790	1.071
3GN7	2.5	1.003	0.463	0.570	0.639	664.050	1.027
3L3M	2.5	0.946	0.530	0.781	0.871	730.250	1.063
3L3L	2.5	0.977	0.514	0.831	0.753	700.410	1.064
2RCW	2.8	0.982	0.485	0.699	0.931	867.450	1.091

2.2. Statistical analysis of binding site properties

The *SiteMap* program (Schrödinger, LLC, New York, NY) was instrumental to define the binding site and calculate binding site descriptors.

A total number of eleven descriptors was calculated on the binding site identified by *SiteMap*, including the number of site points (size), exposure (exp), enclosure (encl), number of van der Waals contacts with ligand (contact), hydrophobic character (phobic), hydrophilic character (philic), hydrophobic/hydrophilic balance (balance), ratio between hydrogen bond donors and hydrogen bond acceptors (donors/acceptors), site score (S-score), drugability score (D-score), and volume. Descriptions of each of these descriptors can be found in the original publication [39].

The above binding site descriptors were used as variables to perform principal component analyses (PCA), using the XL-STAT software. In particular, two PCA were carried out in this study. PCA-1 was performed on binding site descriptors as calculated from the nine crystal structures of the catalytic domain of human PARP-1 in complex with small-molecule inhibitors; PCA-2 was run using binding site descriptors calculated on the structures (2084) collected along REMD simulations of the unbound and ligand-bound PARP-1. Accordingly, PCA-1 was performed on 9 objects and 11 independent variables, and PCA-2 was carried out on 4168 objects and 11 independent variables.

2.3. Ligand preparation and docking calculations

A library of 153 small molecules with binding data to PARP-1 was taken from reference [21] (see supplementary information, Table S1). These ligands were designed and checked with *Marvin* 5.0, 2012, ChemAxon (http://www.chemaxon.com) to adjust connectivity problems wherever present and submitted to the standard *LigPrep* protocol (Schrödinger Suite 2011), using *Epik* to define the ionization states.

Active and inactive binders of PARP-1 were defined using delta thermal shifts (ΔT_M) according to the following scheme of classification: PARP-1 binders presenting $\Delta T_M \geq 2\,^\circ$C were defined as actives (78 active ligands); ligands having $\Delta T_M < 2\,^\circ$C included the weak and inactive binders (75 inactive ligands).

Additional 1171 decoys were also included representing the PARP subset of the DUD dataset [40]. As a result, the final dataset counted 1324 ligands.

Conformations of ligand-bound and unbound PARP-1 were selected from each plot of PCA-2 (Figs. 5 and 6) to cover the first larger group (site s1, conformations a–c) and the second smaller group (site s2, conformation d) of binding sites.

Docking experiments were carried out using *Glide* SP algorithm as implemented in Schrödinger suite, and different scoring schemes including SP and SP-XP.

Although SP and XP scoring functions use similar terms, XP scoring function features a more complete treatment of lipophilic interactions, hydrogen bonds and π–cation/π–π stacking interactions [41]. Docking experiments were performed on a grid box centered on the coordinates of the centroid of compound **1** in the ligand-bound and unbound PARP-1 conformations, and of co-crystallized ligands in the other crystal structures. An inner box length of 10 Å and an outer box of 30 Å were used to define the size of the grid box.

Among the possible binding mode solutions for each ligand, only the best scored one was taken into account to calculate the active enrichment factor (EF^a) and the inactive enrichment factor (EF^i). The EF^a and EF^i in the top scoring 1% ($EF^a_{1\%}$, $EF^i_{1\%}$) of the docked ligands (14 entries) were calculated using the following equations:

$$EF^a = \frac{F_{active}}{F_d} \quad (1)$$

$$EF^i = \frac{F_{inactive}}{F_d} \quad (2)$$

where F_{active} is the fraction of active binders found, $F_{inactive}$ is the fraction of poor and inactive binders found, and F_d is the fraction of the database sampled.

3. Results

3.1. Selecting the input structure and temperature for REMD simulations

PARP-1 belongs to the transferase family of enzymes (EC 2.4.2.30) and catalyzes the poly(ADP-ribosyl)ation of nuclear proteins using NAD^+. The enzyme is a multi-domain protein, with the catalytic domain being located at the C-terminal region of the protein.

Nine co-crystal structures of the catalytic domain of human PARP-1 in complex with small-molecule inhibitors were taken from the PDB. They are endowed with different binding site properties of the relative catalytic cleft (Table 1). Fig. 1, in particular, reports the results of a PCA study of these PARP-1 structures according to *SiteMap* descriptors (PCA-1, see method for details).

The first three components of PCA-1, shown in Fig. 1, explain 85.5% of the variance of the original set of descriptors. From the inspection of the loadings of the descriptors onto the three components (Table 2), solvent exposed binding sites (PC1, loading "exposure" = 0.902) with hydrophobic properties (PC1, loading "phobic" = 0.829; PC1, loading "balance" = 0.903) fall at positive values of the first component (PC1, variance 38.2%), whereas enclosed PARP-1 sites with hydrophilic properties lay at negative values of PC1 (PC1, loading "philic" = −0.752). Likewise, while PARP-1 binding sites rich in hydrogen bond donor properties (PC2, loading "don/acc" = 0.815) and high drugability score (PC2, loading "D-score" = 0.839) spread over positive values of the second component (PC2, variance 27.1%), large PARP-1 binding sites mostly occupy high positive values of the third component (PC3, variance 20.2%; PC3, loading "size" = 0.961; PC3, loading "volume" = 0.871). Inspection of Fig. 1 reveals that eight of the nine crystal structures can be qualitatively organized in two main groups, with one

Table 2
Loadings of the molecular descriptors onto the three principal components of PCA-1 (eigenvalues are reported in brackets). The most relevant molecular descriptors with loading values higher than 0.80 are typed in bold.

Descriptors	PC1 (4.20)	PC2 (2.98)	PC3 (2.22)
S-score	−0.390	**0.881**	−0.073
Size	0.003	−0.026	**0.961**
D-score	0.484	**0.839**	0.243
Volume	−0.034	−0.094	**0.871**
Exposure	**0.902**	−0.265	−0.040
Enclosure	−0.777	0.456	−0.163
Contact	−0.569	0.371	−0.300
Phobic	**0.829**	0.035	−0.446
Philic	−0.752	−0.580	−0.290
Balance	**0.903**	0.270	−0.256
Don/acc	0.036	**0.815**	0.073

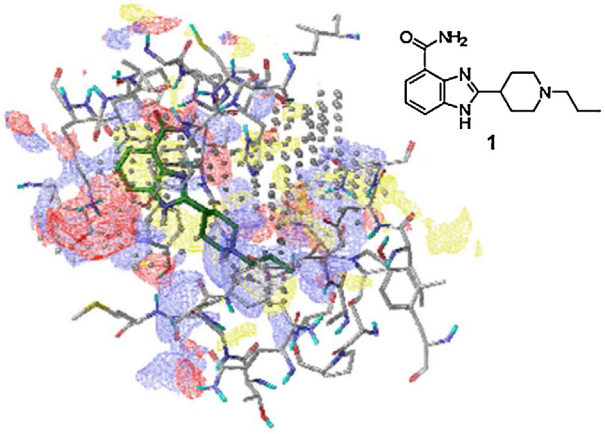

Fig. 2. Binding site properties of the selected crystal structure of PARP-1 (2RCW) in complex with the inhibitor A620223 (**1**). Yellow surfaces show hydrophobic regions; red surfaces highlight polar regions with hydrogen bond acceptor propensity; blue surfaces indicate polar regions with hydrogen bond donor propensity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

containing three structures (PDB ID: 3GN7, 1WOK, and 3GJW) and the other five structures (PDB ID: 1UK1, 1UK0, 3L3L, 3L3M, and 2RD6). Lying in the middle of the two groups, the last structure (PDB ID: 2RCW) is the co-crystallized complex of PARP-1 with the inhibitor A620223 (**1**, Fig. 2).

Since 2RCW does not belong to a specific group of structures and shows binding site properties in between those of the two groups, we deemed it an unbiased input structure for REMD simulations. In order to select the highest temperature for REMD simulations, we considered the thermal stability of the catalytic domain of PARP-1. It was previously reported that the BRCT domain of PARP-1 is endowed with a low melting temperature (43.0 °C), suggesting its poor stability at physiological condition as single domain [42]. Likewise, using differential scanning fluorimetry, a relatively low melting temperature of 48.6 ± 0.2 °C was observed for the catalytic domain of PARP-1 [21], pinpointing the use of a low highest temperature for REMD to avoid protein denaturation during simulations. Accordingly, the range of temperature for REMD simulations was set to 300–315 K.

3.2. The conformational space of ligand-bound and unbound PARP-1

A total of two REMD simulations were carried out using the catalytic domain of PARP-1 with ligand (ligand-bound form) and without ligand (unbound form). The resulting trajectories were analyzed to assess the amount of conformational diversity explored by REMD. Firstly, this was carried out by calculating the root mean square deviation (RMSD) of the backbone atoms of the enzymes from the relative input structures over the 10 ns simulations.

As a result, a large deviation is observed between the unbound and ligand-bound forms of PARP-1 in REMD simulations, with average RMSD values being 2.09 ± 0.32 Å and 1.98 ± 0.29 Å, respectively. Accordingly, a slight stabilizing effect on the conformational state of PARP-1 can be ascribed to inhibitor binding.

The analysis of the RMSD values of secondary structural elements shows that some motifs undergo to higher drifts during the simulations, including the α-helices H2, H3, H4, H8 and the D-loop (Fig. 3). This result may be explained by the high solvent exposure of these secondary elements, which all lay on the surface of the enzyme (Fig. 4). The loop-D, in particular, differs in the aminoacidic composition and length among members of the PARP family. Crystallization studies have shown that this loop is more rigid in PARP-1 than in other members of the family, such as tankyrases (TNKS-1 and TNKS-2), due to the presence of proline residues [21]. In agreement with this observation, our results show that the movements of the D-loop are similar to those of more ordered structural elements such as helices H2, H3, H4 and H8.

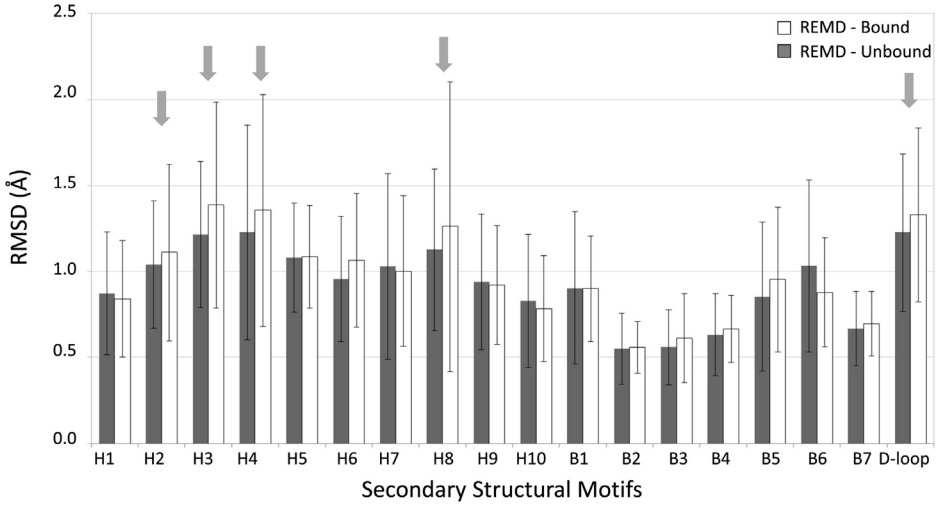

Fig. 3. RMSD fluctuations of secondary structure elements of PARP-1.

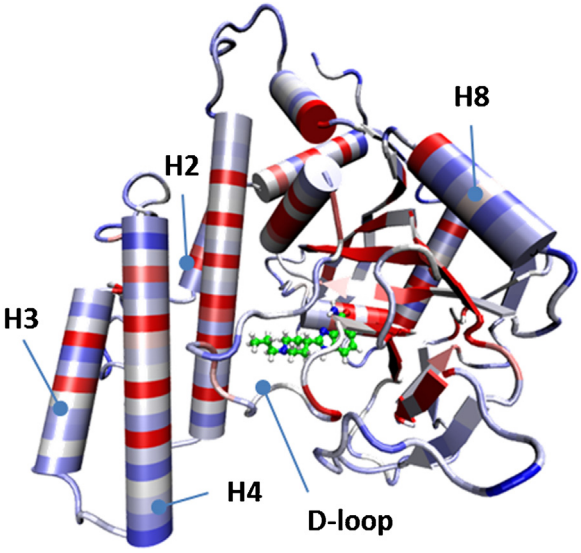

Fig. 4. Solvent exposure of secondary elements of PARP-1 (from blue – high exposure, to red – low exposure) as projected on the selected crystal structure for REMD simulations (2RCW). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Next, we wondered whether the conformational sampling of PARP-1 achieved with REMD simulations was able to gather different or redundant information about binding site properties of the catalytic cleft.

This analysis was carried out performing a further PCA (PCA-2) on the set of binding site descriptors as detailed in the method section. Collectively, the first three components of PCA-2 account for 81% of the variance of the original set of descriptors.

Inspection of the loadings (Table 3) of the original variables into the three components reveals that the first component (PC1, variance 41%) mostly encodes the enclosure (PC1, loading "enclosure" = 0.819) or the drugability (PC1, loading "D-score" = 0.841) of the binding site. On the other hand, the second component (PC2, variance 25%) correlates with the hydrophilic (PC2, loading "philic" = 0.942) or hydrophobic (PC2, loading "phobic" = −0.630; loading "balance" = −0.752) character of the cleft, while the third component (PC3, variance 15%) charts the size of the binding cleft (PC3, loading "size" = 0.620; loading "volume" = 0.748). Figs. 5 and 6 report the plots of the three principal components (PC1–PC3) for the conformations of binding site as generated in REMD trajectories of ligand-bound and unbound PARP-1.

As a first result, REMD simulations explore similar areas of the property space, generating conformational frames that recover binding site properties of PARP-1 catalytic cleft as observed in the crystal complexes.

Table 3
Loadings of the molecular descriptors onto the three principal components of PCA-2 (eigenvalues are reported in brackets). The most relevant molecular descriptors with loading values higher than 0.80 are typed in bold.

Descriptors	PC1 (4.54)	PC2 (2.72)	PC3 (1.61)
S-score	**0.861**	0.259	0.174
Size	0.551	0.355	0.620
D-score	**0.841**	−0.303	0.253
Volume	0.554	0.084	0.748
Exposure	−0.673	−0.508	0.337
Enclosure	**0.819**	0.296	−0.367
Contact	0.684	0.442	−0.476
Phobic	0.662	−0.630	−0.222
Philic	−0.205	**0.943**	−0.065
Balance	0.573	−0.752	−0.195
Don/acc	0.285	−0.195	−0.070

Secondly, two groups of binding sites are clearly observed in distinct regions of the charts. The first and larger group of binding sites (group I, sites s1) is located around the origin of PC1, PC2 and PC3 axes, and includes the experimental conformations of PARP-1 binding site. The second smaller group of binding sites (group II, sites s2) is mainly located at very negative values of PC1, thereby containing binding sites endowed with high solvent exposure. The visual inspection of PARP-1 conformations endowed with sites s2 reveals that sites s1 and s2 are not mutually exclusive; rather s2 corresponds to a narrow pocket that is contiguous to s1 and highly exposed to the solvent. This latter site is formed as a consequence of the conformational movements of Leu324, Tyr325 and Lys242 (Leu985, Tyr986, Lys903 according to the full length sequence of human PARP1). In particular, the positively charged side chain of Lys242 occupies the middle of the putative pocket s2 in 2RCW, making it inaccessible to the solvent. Folding over site s1, the conformational movements of Leu324 and Tyr325 promote the formation of a hydrogen bond between the side chain of Lys242 and the backbone of Tyr325 (Fig. 7). As a net result, Lys242 is displaced from pocket s2 that becomes accessible to the solvent, while Leu324 and Tyr325 contribute to structure the border of this site.

3.3. Effects of binding site flexibility on docking experiments

In this part of the work we wondered whether ensembles of conformations taken from different regions of the property space mapped with REMD simulations could improve the results of docking studies of a library of known active and inactive PARP-1 ligands with respect to the ensemble of crystal structures. Accordingly, four conformations of ligand-bound and unbound PARP-1 were selected from PCA-2 (Figs. 5 and 6) as representative of the first larger group (conformations a–c) and the second smaller group (conformation d) of structures.

A library of 1324 ligands, including 153 small molecules with binding data on PARP-1 [21], was collected and docked into the active site of PARP-1 conformations. In particular, single structures and ensembles of PARP-1 conformations were used as selected from REMD simulations and crystal structures. The enrichment factor $EF^{a}_{1\%}$ of active binders ($\Delta T_M \geq 2\,^{\circ}C$; 78 ligands) and $EF^{i}_{1\%}$ of poor/inactive binders ($\Delta T_M < 2\,^{\circ}C$; 75 ligands) were adopted as criteria to comparatively appraise the performance of conformations selected from REMD and crystal structures of PARP-1 on docking experiments, using two scoring schemes (SP and SP-XP, Table 4).

4. Discussion

The aim of this work was to investigate the conformational space of the catalytic domain of PARP-1 and assess the effect of binding site plasticity on docking experiments of a library of PARP-1 inhibitors. Accordingly, in the first part of this study, we have applied REMD simulations to sample conformations of the catalytic domain of the enzyme.

Overall, it is possible to ascribe a high conformational flexibility to PARP-1 catalytic domain that is mostly localized on some secondary structural elements exposed to the solvent (α-helices H2, H3, H4, H8, Figs. 3 and 4). This first result may be explained with the fact that PARP-1 is a multiple domain protein with diverse functions [43,44]. Accordingly, the high conformational flexibility on the surface of PARP-1 catalytic domain may have a functional ground, promoting inter-domain interactions to couple the catalytic activity of the enzyme to the DNA binding activity of the N-terminal domain and the auto-regulatory function of the central domain (WGR domain). While recent crystallization studies of a DNA double-strand break in complex with human PARP-1 domains, show that helices H2, H3 and H4 are indeed in contact with the

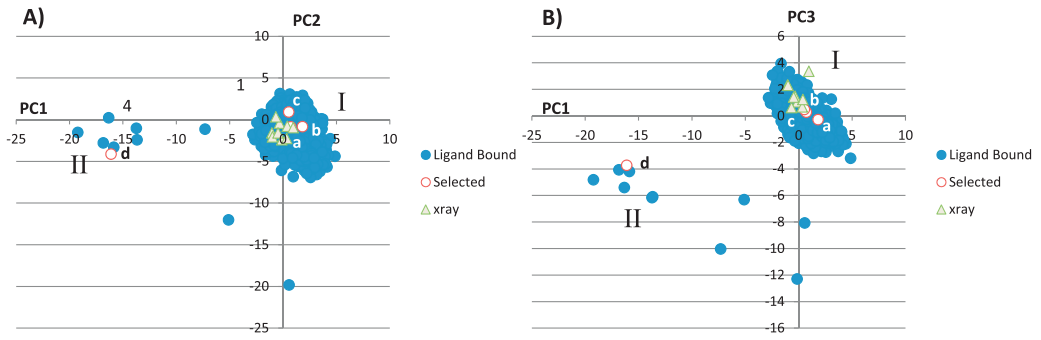

Fig. 5. Plots of PC1 and PC2 components (A), and PC1 and PC3 components (B) of PCA-2 according to the frames generated in the REMD simulation of ligand-bound PARP-1.

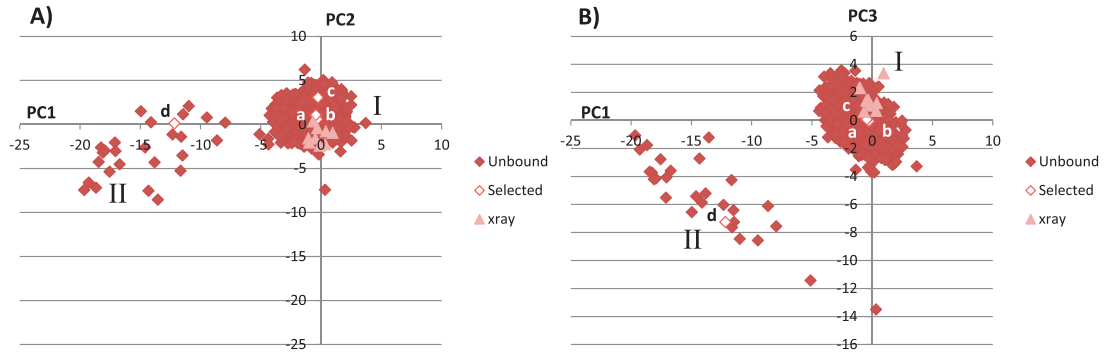

Fig. 6. Plots of PC1 and PC2 components (A), and PC1 and PC3 components (B) of PCA-2 according to the frames generated in the REMD simulation of unbound PARP-1.

central WGR domain, favoring a collapsed conformation of PARP-1 domains [45], a more recent work pinpoints the importance of inter-domain interactions, reporting the enhancement of PARP-1 binding to DNA by different inhibitors that trap the enzyme at damaged DNA regions with different potencies [44].

As a second result, REMD simulations of the ligand-bound and unbound enzyme have allowed to disclose the formation of an accessory pocket (site s2, Figs. 5 and 6) into the catalytic cleft of PARP-1, being contiguous to the main pocket (site s1). The accessory site s2 is locked in the input structure 2RCW, and it is absent in the other crystal structures that are endowed with a lower binding site volume (Table 1). Although we cannot rule out that this is due to a poor parameterization of *SiteMap*, this pocket seems the outcome of conformational movements involving Leu324, Tyr325 and Lys242 that unlock site s2 during REMD simulations (Fig. 7). Of note, Leu324, Tyr325 and Lys242 are not fully conserved among all members of ARTDs family, suggesting that the aforementioned movements may be a specific feature of PARP-1.

In the second part of this work we have performed docking experiments of a library of PARP-1 ligands, analyzing the effect of enzyme flexibility in improving the enrichment factors of active ligands. This was accomplished using structures generated from REMD simulations, and selected on the basis of binding site descriptors, in comparison with crystal structures. It should be mentioned that a recent study has reported that the use of ensemble of conformations from REMD, while providing a significant enrichment and diversity of active compounds in docking studies with respect to single conformations, was not able to perform better than an

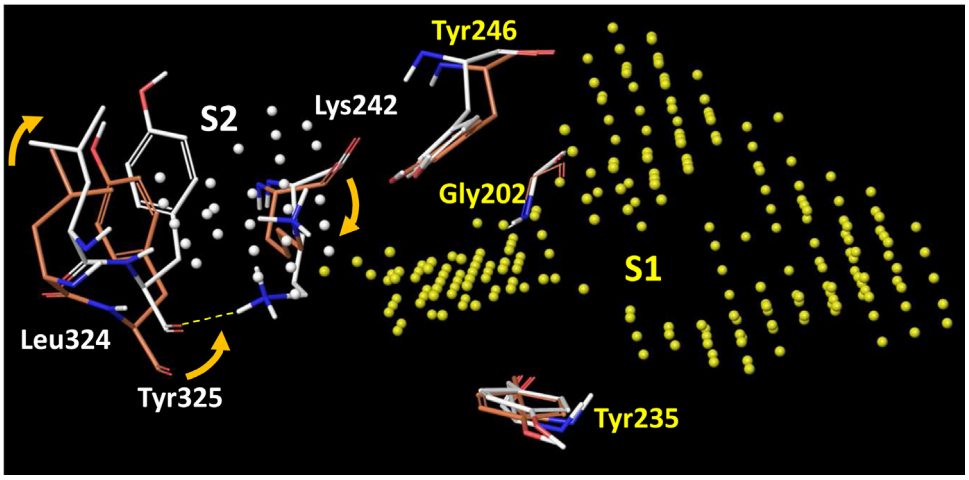

Fig. 7. Superposition of binding site residues between REMD bound conf. d (white sticks) and 2RCW (orange sticks). Sites s1 and s2 are shown as yellow dots and white dots, respectively. Arrows indicate the conformational movements of Leu324, Tyr325 and Lys242 that promote the formation of pocket s2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

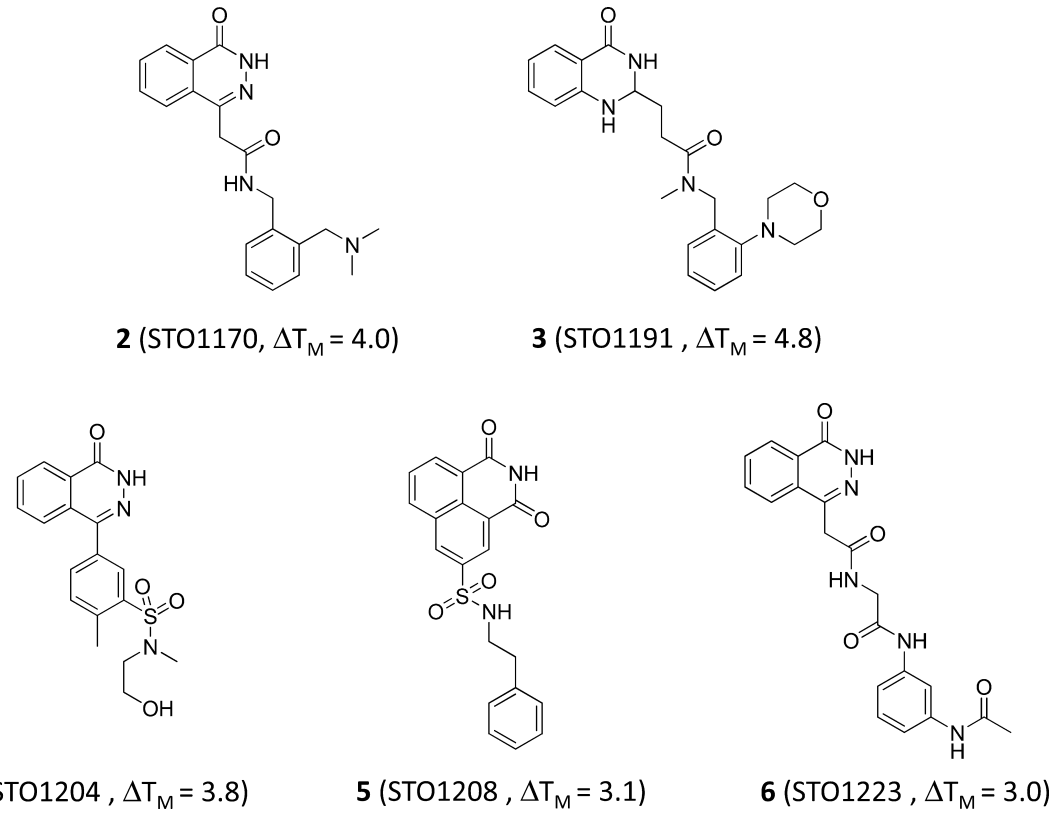

Fig. 8. Chemical structures of selected active ligands that are specifically present in the in the top scoring 1% list (SP-XP scores) of REMD bound conf. d.

Table 4
Enrichment factors ($EF_{1\%}$) of docking experiments using single structures and ensembles of conformations of PARP-1, according to different scoring schemes. The number of active ligands found in the top 1% of hits is reported in brackets.

Conformations	SP – $EF^{a}_{1\%}$	SP – $EF^{i}_{1\%}$	SP-XP – $EF^{a}_{1\%}$	SP-XP – $EF^{i}_{1\%}$
2RCW	9.7 (8)	7.6 (6)	9.7 (8)	7.6 (6)
1UK0	6.1 (5)	11.3 (9)	8.5 (7)	6.3 (5)
1UK1	7.3 (6)	10.1 (8)	8.5 (7)	8.8 (7)
1WOK	7.3 (6)	10.1 (8)	7.3 (6)	10.1 (8)
2RD6	7.3 (6)	10.1 (8)	10.9 (9)	6.3 (5)
3GJW	7.3 (6)	10.1 (8)	10.9 (9)	6.3 (5)
3GN7	9.7 (8)	7.6 (6)	9.7 (8)	7.6 (6)
3L3L	8.5 (7)	8.8 (7)	9.7 (8)	7.6 (6)
3L3M	8.5 (7)	8.8 (7)	9.7 (8)	7.6 (6)
Ensemble (9)	3.6 (3)	11.3 (9)	8.5 (7)	7.6 (6)
REMD Bound conf. a (s1)	7.3 (6)	8.8 (7)	7.3 (6)	10.1 (8)
REMD Bound conf. b (s1)	1.2 (1)	8.8 (7)	8.5 (7)	8.8 (7)
REMD Bound conf. c (s1)	6.1 (5)	11.3 (9)	8.5 (7)	8.8 (7)
REMD Bound conf. d (s2)	6.1 (5)	11.3 (9)	12.1 (10)	5.0 (4)
Ensemble (4)	12.1 (10)	2.5 (2)	7.3 (6)	8.8 (7)
REMD Unbound conf. a (s1)	1.2 (1)	3.8 (3)	1.2 (1)	8.8 (7)
REMD Unbound conf. b (s1)	4.8 (4)	3.8 (3)	2.4 (2)	7.6 (6)
REMD Unbound conf. c (s1)	1.2 (1)	0 (0)	2.4 (2)	2.5 (2)
REMD Unbound conf. d (s2)	3.6 (3)	7.6 (6)	1.2 (1)	7.6 (6)
Ensemble (4)	1.2 (1)	3.8 (3)	2.4 (2)	7.6 (6)

ensemble of crystal structures, with the latter yielding higher enrichment and diversity [27]. The inspection of Table 4 shows that these observations in part apply to our case study, though the use of different REMD protocols and scoring schemes may bias such a comparison.

In particular, we have observed that docking into the ensemble of experimental structures does not result in an improvement of the enrichment factor of active ligands ($EF^{a}_{1\%}$) with respect to the use of single experimental conformations of the catalytic domain of PARP-1. Conversely, docking into conformations selected from REMD simulations of ligand bound PARP-1 and using SP scores provides the highest enrichment factor of active ligands ($^{SP}EF^{a}_{1\%} = 12.1$) over single conformations, combined with a poor enrichment factor of inactive ligands ($^{SP}EF^{i}_{1\%} = 2.5$). This level of enrichment of active ligands is also observed in the case of SP-XP scores using a single PARP-1 conformation endowed with site s2 (REMD bound conf. d; $^{SP\text{-}XP}EF^{a}_{1\%} = 12.1$), suggesting that the formation of this pocket may in part account for the improvement of the enrichment factor of active ligands in docking experiments.

In order to understand how some active ligands exploit pocket s2, we have analyzed the binding poses of compounds **2–6** (Fig. 8). These compounds rank in the top scoring 1% list (SP-XP scores) of REMD bound conf. d, but they are out from the top 1% lists of REMD bound conf. a–c devoid of site s2. The inspection of the relative binding poses in REMD bound conf. d suggests that the aromatic moiety of these ligands would form steric clashes with the side chain of Lys242 as observed in REMD bound conf. a–c, whereas it nicely fits pocket s2 in REMD bound conf. d (Fig. 9).

Another interesting observation concerns the different performance observed when using conformations retrieved from the ligand-bound PARP-1 simulation with respect to the unbound PARP-1 simulation. As a premise, it should be mentioned that the

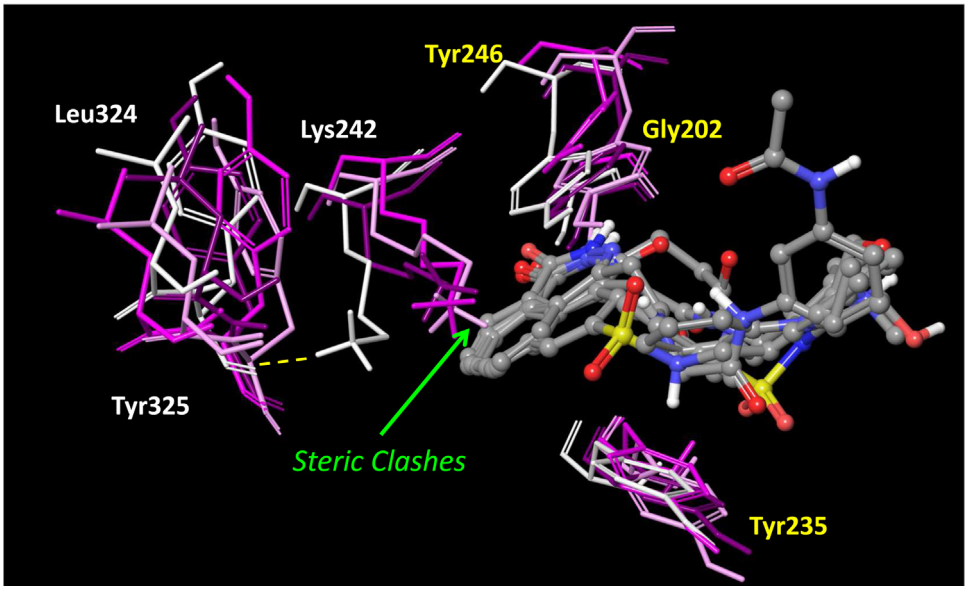

Fig. 9. Binding poses of compounds **2–6** in REMD bound conf. d (white sticks). Binding site residues of REMD bound conf. a (magenta sticks), REMD bound conf. b (purple sticks) and REMD bound conf. c (pink sticks) are superimposed to show the steric clashes (green arrow) that would form between the aromatic moiety of ligands **2–6** and the side chain of Lys242.

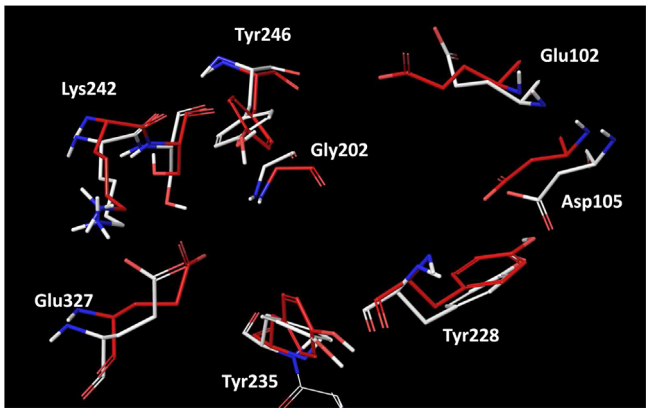

Fig. 10. Superposition of binding site residues between REMD bound conf. d (white sticks) and REMD unbound conf. d (red sticks). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

presence of ligand **1** in the catalytic cleft of PARP-1 does not bias the results of ligand-bound simulation, since the resulting conformations have similar binding site properties to those generated in the unbound simulation (Figs. 5 and 6).

Nevertheless, conformations of the unbound simulation yield worse enrichment factors of active ligands ($EF^{a}_{1\%} < 5$) than conformations of ligand-bound simulation. In some of these cases, the enrichment factor of inactive ligands is even significantly higher than that of active ligands (e.g. REMD unbound conf. a; $^{SP\text{-}XP}EF^{a}_{1\%} = 1.2$, $^{SP\text{-}XP}EF^{i}_{1\%} = 8.8$). Although such an observation may be counterintuitive on the basis of the similar binding site properties reported for these two set of conformations (Figs. 5 and 6), a more careful analysis of binding site residues unveils subtle differences in side chain rotamers that may affect the binding of ligands (see supplementary information, Fig. s1). For instance, comparing the binding site of the best performing conformation REMD bound conf. d to REMD unbound conf. d, which is still endowed with site s2 though being a poor performing conformation, it is worth noting the flip of the side chains of Tyr235 and Tyr246 from the face to the edge rotamer (Fig. 10). Since these residues are involved in conserved face to face π-stacking interactions with PARP-1 inhibitors [25], the above conformational movement results detrimental for a productive binding of ligands to the enzyme, thereby accounting for the low enrichment factors of REMD unbound conf. d ($^{SP}EF^{a}_{1\%} = 3.6$; $^{SP\text{-}XP}EF^{a}_{1\%} = 1.2$).

5. Conclusions

In this study we have first used REMD simulations to sample different conformations of the catalytic domain of PARP-1 in the ligand-bound and unbound forms. Then, we have assessed the impact of the emerging enzyme flexibility in docking experiments.

Both ligand-bound and unbound simulations of PARP-1 unveil the formation of an accessory pocket (s2) in the catalytic cleft of the enzyme as a consequence of the conformational movements of Leu324, Tyr325 and Lys242. The use of a PARP-1 conformation endowed with s2 pocket from REMD simulation of the ligand bound enzyme results in the improvement of the enrichment factor of active ligands ($EF^{a}_{1\%}$) in docking studies. Furthermore, despite having similar binding site properties, conformations generated from ligand-bound simulation and unbound simulation show different performance in docking experiments, with the latter yielding lower enrichment factors. The inspection of binding site residues suggests that this may be ascribed to the presence of different side chain rotamers that badly affect the geometries of ligand–protein interactions.

Collectively, these results pinpoint the beneficial effect of including enzyme flexibility in docking studies of PARP-1 ligands. However, it is also worth noting that this effect is appreciated when conformations selected from the ligand-bound REMD simulation of the enzyme are used for calculations.

6. Supporting information available

Supporting information for this article is available. Table s1 contains the list of PARP-1 active ligands used in this study. Fig. s1 shows the superposition of the side chains of binding site residues as observed in conformations selected from ligand-bound and unbound REMD simulations of PARP-1. Figs. s2 and s3 show the plot of the replicas as a function of time for REMD simulations of ligand-bound and unbound PARP-1.

Acknowledgements

We thank Elisabet Wahlberg and Herwig Schüler (Karolinska Institutet, Stockholm, Sweden) for providing us the data on the melting temperatures of the catalytic domain of recombinant human PARP-1. This research was funded by the Catalan Government grant 2012FIB1-00175 (to A.A.A.), the Spanish Ministerio de Ciencia e Innovación grant BIO2011-26669 (to J.M.), and the Instituto de Salud Carlos III (to J.M.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2013.08.006.

References

[1] P.O. Hassa, M.O. Hottiger, The diverse biological roles of mammalian PARPS, a small but powerful family of poly-ADP-ribose polymerases, Front. Biosci. 13 (2008) 3046–3082.
[2] V. Schreiber, F. Dantzer, J.C. Ame, G. de Murcia, Poly(ADP-ribose): novel functions for an old molecule, Nat. Rev. Mol. Cell Biol. 7 (2006) 517–528.
[3] M. Rouleau, A. Patel, M.J. Hendzel, S.H. Kaufmann, G.G. Poirier, PARP inhibition: PARP1 and beyond, Nat. Rev. Cancer 10 (2010) 293–301.
[4] T. Kalisch, J.C. Ame, F. Dantzer, V. Schreiber, New readers and interpretations of poly(ADP-ribosyl)ation, Trends Biochem. Sci. 37 (2012) 381–390.
[5] X. Luo, W.L. Kraus, On PAR with PARP: cellular stress signaling through poly(ADP-ribose) and PARP-1, Genes Dev. 26 (2012) 417–432.
[6] Y. Ji, A.V. Tulin, The roles of PARP1 in gene control and cell differentiation, Curr. Opin. Genet. Dev. 20 (2010) 512–528.
[7] W.L. Kraus, Transcriptional control by PARP-1: chromatin modulation, enhancer-binding, coregulation, and insulation, Curr. Opin. Cell Biol. 20 (2008) 294–302.
[8] R. Alvarez-Gonzalez, Genomic maintenance: the p53 poly(ADP-ribosyl)ation connection, Sci. STKE 415 (2007) pe68.
[9] P. Bai, C. Canto, The role of PARP-1 and PARP-2 enzymes in metabolic regulation and disease, Cell Metab. 16 (2012) 290–295.
[10] B.A. Gibson, W.L. Kraus, New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs, Nat. Rev. Mol. Cell Biol. 13 (2012) 411–424.
[11] E.M. Horvath, C. Szabo, Poly(ADP-ribose) polymerase as a drug target for cardiovascular disease and cancer: an update, Drug News Perspect. 20 (2007) 171–181.
[12] C.A. de la Lastra, I. Villegas, S. Sanchez-Fidalgo, Poly(ADP-ribose) polymerase inhibitors: new pharmacological functions and potential clinical implications, Curr. Pharm. Des. 13 (2007) 933–962.
[13] D.V. Ferraris, Evolution of poly(ADP-ribose) polymerase-1 (PARP-1) inhibitors. From concept to clinic, J. Med. Chem. 53 (2010) 4561–4584.
[14] T. Ekblad, E. Camaioni, H. Schüler, A. Macchiarulo, PARP inhibitors: polypharmacology versus selective inhibition, FEBS J. (2013), http://dx.doi.org/10.1111/febs.12298.
[15] E.S. Ratner, A.C. Sartorelli, Z.P. Lin, Poly(ADP-ribose) polymerase inhibitors: on the horizon of tailored and personalized therapies for epithelial ovarian cancer, Curr. Opin. Oncol. 24 (2012) 564–571.
[16] T.D. Penning, Small-molecule PARP modulators – current status and future therapeutic potential, Curr. Opin. Drug Discov. Dev. 13 (2010) 577–586.
[17] F.N. Novikov, V.S. Stroylov, O.V. Stroganov, V. Kulkov, G.G. Chilov, Developing novel approaches to improve binding energy estimation and virtual screening: a PARP case study, J. Mol. Model. 15 (2009), 1337–1147.

[18] T.V. Le, J.H. Suh, N. Kim, H.J. Park, In silico identification of poly(ADP-ribose)polymerase-1 inhibitors and their chemosensitizing effects against cisplatin-resistant human gastric cancer cells, Bioorg. Med. Chem. Lett. 23 (2013) 2642–2646.
[19] D. Branca, M. Cerretani, P. Jones, U. Koch, F. Orvieto, M.C. Palumbi, M. Rowley, C. Toniatti, E. Muraglia, Identification of aminoethyl pyrrolo dihydroisoquinolinones as novel poly(ADP-ribose) polymerase-1 inhibitors, Bioorg. Med. Chem. Lett. 19 (2009) 4042–4045.
[20] C.D. Andersson, T. Karlberg, T. Ekblad, A.E. Lindgren, A.G. Thorsell, S. Spjut, U. Uciechowska, M.S. Niemiec, P. Wittung-Stafshede, J. Weigelt, M. Elofsson, H. Schüler, A. Linusson, Discovery of ligands for ADP-ribosyltransferases via docking-based virtual screening, J. Med. Chem. 55 (2012) 7706–7718.
[21] E. Wahlberg, T. Karlberg, E. Kouznetsova, N. Markova, A. Macchiarulo, A.G. Thorsell, E. Pol, A. Frostell, T. Ekblad, D. Oncu, B. Kull, G.M. Robertson, R. Pellicciari, H. Schuler, J. Weigelt, Family-wide chemical profiling and structural analysis of PARP and tankyrase inhibitors, Nat. Biotechnol. 3 (2012) 283–288.
[22] A.A. Antolin, X. Jalencas, J. Yelamos, J. Mestres, Identification of Pim kinases as novel targets for PJ34 with confounding effects in PARP biology, ACS Chem. Biol. 7 (2012) 1962–1967.
[23] R. Pellicciari, E. Camaioni, A.M. Gilbert, A. Macchiarulo, J.A. Bikker, F. Shah, J. Bard, G. Costantino, A. Gioiello, G.M. Robertson, P. Sabbatini, F. Venturoni, P. Liscio, A. Carotti, D. Bellocchi, A. Cozzi, A. Wood, C. Gonzales, M.M. Zaleska, J.W. Ellingboe, F. Moroni, Discovery and characterization of novel potent PARP-1 inhibitors endowed with neuroprotective properties: from TIQ-A to HYDAM-TIQ, MedChemComm 6 (2011) 559–565.
[24] D. Bellocchi, A. Macchiarulo, G. Costantino, R. Pellicciari, Docking studies on PARP-1 inhibitors: insights into the role of a binding pocket water molecule, Bioorg. Med. Chem. 4 (2005) 1151–1157.
[25] G. Costantino, A. Macchiarulo, E. Camaioni, R. Pellicciari, Modeling of poly(ADP-ribose)polymerase (PARP) inhibitors. Docking of ligands and quantitative structure–activity relationship analysis, J. Med. Chem. 23 (2001) 3786–3794.
[26] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chem. Phys. Lett. 314 (1999) 141–151.
[27] D.J. Earl, M.W. Deem, Parallel tempering: theory, applications, and new perspectives, Phys. Chem. Chem. Phys. 23 (2005) 3910–3916.
[28] K. Ostermeir, M. Zacharias, Advanced replica-exchange sampling to study the flexibility and plasticity of peptides and proteins, Biochim. Biophys. Acta 1834 (2013) 847–853.
[29] N. Metropolis, S. Ulam, The Monte Carlo method, J. Am. Stat. Assoc. 247 (1949) 335–341.
[30] I.H. Park, C. Li, Dynamic ligand-induced-fit simulation via enhanced conformational samplings and ensemble dockings: a survivin example, J. Phys. Chem. B 15 (2010) 5144–5153.
[31] D.J. Osguthorpe, W. Sherman, A.T. Hagler, Exploring protein flexibility: incorporating structural ensembles from crystal structures and simulation into virtual screening protocols, J. Phys. Chem. B 23 (2012) 6952–6959.
[32] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 1 (2000) 235–242.
[33] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. Di Nola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[34] G.J. Martyna, M.E. Tuckerman, D.J. Tobias, M.L. Klein, Explicit reversible integrators for extended system dynamics, Mol. Phys. 87 (1996) 1117–1157.
[35] S.J. Nose, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. 81 (1984) 511–519.
[36] W.G. Hoover, Canonical dynamics: equilibrium phase-space distributions, Phys. Rev. A 31 (1985) 1695–1697.
[37] E. Rosta, N.V. Buchete, G. Hummer, Thermostat artifacts in replica exchange molecular dynamics simulations, J. Chem. Theory Comput. 5 (2009) 1393–1399.
[38] A. Patriksson, D. van der Spoel, A temperature predictor for parallel tempering simulations, Phys. Chem. Chem. Phys. 15 (2008) 2073–2077.
[39] T.A. Halgren, Identifying and characterizing binding sites and assessing druggability, J. Chem. Inf. Model. 2 (2009) 377–389.
[40] N. Huang, B.K. Shoichet, J.J. Irwin, Benchmarking sets for molecular docking, J. Med. Chem. 49 (2006) 6789–6801.
[41] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes, J. Med. Chem. 49 (2006) 6177–6196.
[42] P.A. Loeffler, M.J. Cuneo, G.A. Mueller, E.F. DeRose, S.A. Gabel, R.E. London, Structural studies of the PARP-1 BRCT domain, BMC Struct. Biol. 11 (2011) 37.
[43] C. Guetg, R. Santoro, Noncoding RNAs link PARP1 to heterochromatin, Cell. Cycle 12 (2012) 2217–2218.
[44] J. Murai, S.Y. Huang, B.B. Das, A. Renaud, Y. Zhang, J.H. Doroshow, J. Ji, S. Takeda, Y. Pommier, Trapping of PARP1 and PARP2 by clinical PARP inhibitors, Cancer Res. 21 (2012) 5588–5599.
[45] M.F. Langelier, J.L. Planck, S. Roy, J.M. Pascal, Structural basis for DNA damage-dependent poly(ADP-ribosyl)ation by human PARP-1, Science 6082 (2012) 728–732.