J Comput Aided Mol Des (2009) 23:73–86
DOI 10.1007/s10822-008-9239-y

What induces pocket openings on protein surface patches involved in protein–protein interactions?

Susanne Eyrisch · Volkhard Helms

Received: 31 March 2008 / Accepted: 12 August 2008 / Published online: 6 September 2008

Abstract We previously showed for the proteins BCL-X_L, IL-2, and MDM2 that transient pockets at their protein–protein binding interfaces can be identified by applying the PASS algorithm to molecular dynamics (MD) snapshots. We now investigated which aspects of the natural conformational dynamics of proteins induce the formation of such pockets. The pocket detection protocol was applied to three different conformational ensembles for the same proteins that were extracted from MD simulations of the inhibitor bound crystal conformation in water and the free crystal/NMR structure in water and in methanol. Additional MD simulations studied the impact of backbone mobility. The more efficient CONCOORD or normal mode analysis (NMA) techniques gave significantly smaller pockets than MD simulations, whereas tCONCOORD generated pockets comparable to those observed in MD simulations for two of the three systems. Our findings emphasize the influence of solvent polarity and backbone rearrangements on the formation of pockets on protein surfaces and should be helpful in future generation of transient pockets as putative ligand binding sites at protein–protein interfaces.

Keywords Binding pocket · CONCOORD · Molecular dynamics simulation · Normal mode analysis · Protein–protein interaction inhibition · Structure-based drug design · tCONCOORD

Electronic supplementary material The online version of this article (doi:10.1007/s10822-008-9239-y) contains supplementary material, which is available to authorized users.

S. Eyrisch · V. Helms (✉)
Center for Bioinformatics, Building C7 1,
P.O. Box 15 11 50, D-66041 Saarbruecken, Germany
e-mail: volkhard.helms@bioinformatik.uni-saarland.de

Introduction

Within the last decade, the number of newly identified or designed enzyme inhibitors has been steadily rising. These advances are not only due to improved screening techniques, but also due to the continuous progress in structure-based drug design and the soaring number of available high-resolution protein structures. If the three-dimensional target structure is known one may identify the binding site and compare its biochemical and geometrical properties to those of potential ligands, e.g. from compound libraries. The computational prediction of the binding site can be based on geometric or energetic criteria. The approaches of the former type identify clefts or cavities on the protein surface (e.g. PASS [1], LigSite [2], POCKET [3], SURFNET [4]) and it has been shown that ligands usually bind in the largest pocket [5–7]. The approaches of the latter type screen the protein surface for energetically favorable interaction sites, so-called 'hot spots'. Many methods like CS-Map [8] and GRID [9] compute interaction energies between the protein and different probe types (hydrophobic, aromatic, hydrogen bond donor/acceptor, …), whereas others such as QSiteFinder [10] consider only van-der-Waals interactions.

Although all these methods have been successfully used in identifying inhibitors for interactions between proteins and small-molecule ligands, it is unclear whether they are also capable of identifying inhibitors of protein–protein interactions. Within the last few years, the design of such small molecules that target protein–protein interactions has become a very active field of research. So far, small molecule protein–protein interaction inhibitors (SMPPIIs) have been identified for several protein complexes such as, for example, MDM2–p53, BCL-X_L–Bak, IL-2–IL-2Rα, LFA-1–ICAM-1, Ras–Raf, and TNF–TNFRc1 [11–13]. Of

particular interest is the interaction of p53 with MDM2 that promotes its degradation and thus prevents the transcription of genes that control e.g. apoptosis and DNA repair [14, 15]. p53 is the most frequently inactivated protein in cancer and its interaction with MDM2 has become an important drug target in anti-cancer therapy [16, 17]. Several classes of molecules have been characterized that inhibit this interaction [18–20]. Until today, most published SMPPIIs for this class of drug targets were identified by experimental screening methods. Structure-based drug design remains a great challenge because, in contrast to the well defined binding pockets of enzymes, most free structures of proteins involved in protein–protein interactions lack clearly shaped binding pockets at the interface. Consequently, it is nearly infeasible to identify potential binding sites from the crystal structure alone [21]. Although docking methods allowing for (partial) receptor flexibility have proven to be quite promising even if rigid docking to the free protein structure failed [22–25], they depend on a definition of the known (or a potential) binding region of the protein. Some advanced methods incorporate receptor flexibility by conformational ensembles that are often taken from MD simulations. One example is the relaxed complex scheme [26, 27] that combines MD simulations with docking algorithms. If the binding site is unknown, the entire protein surface will be taken into account. In another example, new SMPPIIs for the MDM2:p53 interaction were identified by docking ligands to a dynamic receptor-based pharmacophore model that was based on the crystal structure of human MDM2 bound to p53 [28].

We have previously presented a pocket detection protocol that provides a starting point for in silico drug design in cases when no potential binding pocket could be identified so that standard screening methods would fail [29]. For the three protein systems MDM2 (murine double minute 2), BCL-X_L (basal cell lymphoma-extra large), and IL-2 (Interleukin-2), we found that large pockets not detectable in the crystal structures of the unliganded proteins opened frequently on the protein surfaces during standard MD simulations of 10 ns length at room temperature. The identified transient pockets represent potential binding sites of new inhibitors. Comparisons between pockets observed in two MD simulation runs for the same system showed that most of the pockets were reproducible. This was particularly true for the most frequent ones. At the native binding site, pockets of similar size as with a known inhibitor bound could be observed for all three systems. Docking known inhibitors with AutoDock 3.0.5 [30] into these transient pockets resulted in docking results with less than 2 Å root mean square deviation (RMSD) from the crystal structures. Although these pockets were less polar than the overall protein surface, most of them were not as hydrophobic as the native binding pocket. It is currently unclear whether the opening of more nonpolar pockets is energetically "forbidden" in water and only possible in the presence of a nearby ligand. Here we speculated that this could be circumvented by simulating the protein system in a nonpolar solvent that may allow for the opening of more and larger hydrophobic pockets, even in the absence of a ligand. Methanol appeared as a good candidate solvent as it may act as a hydrogen bond donor and acceptor and is less polar than water ($\varepsilon = 33$ [31]).

Alonso and Daggett studied the unfolding and folding of ubiquitin by MD simulations in a mixture of methanol and water to mimic the cytosolic environment of biological cells [32]. Interestingly, they observed that partially unfolded conformations with increased exposure of hydrophobic residues were only stable in the presence of methanol whereas the protein collapsed in water. MD simulations in pure methanol have mostly been applied to membrane-bound peptides so far. For example, Kovacs et al. compared simulations of an integral membrane helix of the surfactant protein C in chloroform, water, and methanol [33]. The total accessible surface of the peptide increased in methanol and chloroform, whereas it decreased in water. The increase in chloroform was driven by the increasing exposure of non-polar side chains, whereas the decrease in water was driven by the burial of aliphatic side chains. In methanol, polar as well as non-polar side-chains became exposed. Further, they observed that the helical conformation was more stable in water and methanol than in chloroform.

To understand the underlying dynamics of the opening of transient pockets on protein surfaces, we investigated the following points: (1) How stable is the native binding pocket without bound ligand? (2) Are backbone movements necessary for the opening of pockets or are side chain rotations sufficient? (3) Can pockets on the protein surface fully open in water and what is the additional benefit of simulating the proteins in a less polar solvent? For answering these questions, we applied the pocket detection protocol to different conformational ensembles for all three protein systems. The previously published MD simulations of the free proteins serve as a reference point. The first question is relatively straightforward and was answered by MD simulations of the inhibitor-bound conformations after removal of the inhibitor. The second question was addressed by comparing MD simulations with harmonic restraints on all heavy backbone atoms to the unrestrained simulation at the example of MDM2. To answer the third question, we extracted conformational ensembles from MD simulations in methanol or in water.

A limiting factor of our previously introduced pocket detection protocol is the high computational demand of MD simulations on biomolecular systems and it would be desirable to replace MD simulations by a more efficient

protocol. Therefore, we have also tested three established methods that generate conformational ensembles in a more efficient way either using random displacements along eigenvectors extracted from normal mode analysis (NMA) or by the CONCOORD [34], or the tCONCOORD [35] methods. For completeness, we will shortly introduce these methods here. Normal mode analysis is a very well-established technique to study conformational changes in biomolecules [36–38]. The shape of the potential energy function around a local minimum is assumed as harmonic, so that normal modes in these energy valleys represent harmonic vibrations. Based on the minimized structure of the biomolecule, a mass-weighted hessian matrix is calculated whose entries are the derivatives of the forces with respect to every coordinate. The eigenvectors describing the direction of the normal modes are obtained by diagonalization of this matrix [39]. In several proteins like lysozyme, myosin, and kinesin, conformational changes upon ligand binding could be described by a single, or multiple low-frequency normal modes [37, 40, 41]. Furthermore, NMA can be used for the generation of a conformational ensemble. For example, Fu et al. recently designed novel α-helical peptides as BCL-X_L ligands where they used NMA for sampling different backbone conformations of the ligand [42].

CONCOORD, on the other hand, generates protein conformations that fulfill previously calculated distance bounds. These distance bounds are determined by measuring all pairwise interatomic distances in the input structure and identifying the type of each interacting pair (e.g. covalent bonds, hydrogen bonds, tight or loose hydrophobic contact, part of the same secondary structure). It was shown that CONCOORD ensembles may closely resemble ensembles of conformations extracted from MD simulations [34]. Further, Barrett et al. used the peptide-bound MDM2 protein to show that principal component analysis of a conformational ensemble generated using CONCOORD predicted quite similar concerted motions as an MD simulation [43]. More interestingly, they observed that the first eigenvector was coupled to an opening and closing motion of the native binding pocket that was even more pronounced when the peptide was removed. These findings are in good agreement with a later study by Espinoza-Fonseca and Trujillo-Ferrara who conducted MD simulations of the same crystal structure of MDM2 with and without the bound peptide [44]. They found that most motions of this binding site were accounted for by the first eigenvector for the bound protein and by the first two eigenvectors for the unbound protein. Further, they also observed that the binding cleft was wider and more stable with the peptide bound.

tCONCOORD is an extension of CONCOORD. After analyzing the neighborhood of all hydrogen bonds, those that are estimated to be stable are translated into distance constraints. Hydrogen bonds that may possibly be broken are ignored. This strategy allows for the prediction of large-scale conformational transitions that are not observable using the CONCCORD method. When using the apo protein structures as input, tCONCOORD was able to generate conformations close to experimentally determined ligand-bound structures [35].

The main goal of this study was to investigate which aspects of the natural conformational dynamics of proteins (e.g. backbone movements, side chain movements, deformations along low-frequency normal modes) induce the formation of surface pockets. To this end, we generated conformational ensembles using different methods and compared the properties of all detected transient pockets. Furthermore, we tried to characterize for each method its appropriateness for detecting potential binding pockets. This was realized by focusing on the binding pockets of known SMPPIIs because these are the only cavities with experimentally validated small-molecule ligand binding capabilities. By docking the known SMPPIIs with the docking program AutoDock 3 into transient pockets that opened at the binding interface and comparing the computed ligand binding pose to that observed in the complex crystal structure, we could identify which methods are best suited for sampling the native binding pocket.

Methods and materials

The structure preparation, equilibration, MD simulations, and docking runs were done as previously described [29]. We summarize the main steps for completeness.

Structure preparation

The free and inhibitor-bound protein structures of IL-2 (free: 1M47, bound: 1PY2_A), MDM2 (free: 1Z1 M, bound: 1T4E_A), and BCL-X_L (free: 1R2D, bound: 1YSI) were taken from the Protein Data Bank [45]. All hetero atoms (including the ligands) were manually removed. As residues 28–81 are missing in 1R2D (residues 45–84 in 1YSI, respectively; residue numbering according to 1R2D), the two parts of the protein were modelled as two distinct chains. The missing residues in 1M47 were modelled as loops of the lowest AMBER/GBSA potential energy generated by the program RAPPER [46]. We note that these missing residues are far away from the native binding pocket. The structure of free MDM2 is represented by 24 NMR models that differ mainly in the loop regions. Since no model is defined as most representative, the first model was chosen.

All energy minimizations, MD simulations, and normal mode analysis were performed with the GROMACS 3.3.1 package [47] using the OPLS-AA force field [48].

Molecular dynamics simulations

After placing the unbound structure of each protein in cubic water boxes of 6.2–8.3 nm box dimensions, the system was relaxed by 500 steps of steepest-descent energy minimization keeping the heavy protein atoms harmonically restrained (force constant of 1,000 kJ mol^{-1} nm^{-2}). Solvent molecules were modelled explicitly using the TIP4P water model [49]. If needed, counter ions (Na^+ or Cl^-, respectively) were added to ensure a net neutral charge of the systems and the pre-equilibrating run was repeated. Equilibration continued during a 100 ps simulation in the NPT ensemble at a temperature of 300 K with periodic boundary conditions. Simulation snapshots were collected during a subsequent 10 ns simulation. Electrostatic interactions beyond the short-range cut-off of 0.9 nm were treated by the particle-mesh-Ewald method [50]. Van der Waals interactions were computed within a 0.9 nm cut-off. Temperature and pressure were kept constant at standard conditions (10^5 Pa, 300 K) by weak coupling to a temperature and pressure bath [51] with coupling constants of 0.1 ps for the temperature coupling and 1 ps for the pressure coupling. Protein, solvent, and counter ions were coupled to separate baths. The LINCS procedure [52] was used to constrain all covalent bonds. Conformational snapshots were saved every 2.5 ps yielding a total of 4001 MD snapshots. Before further processing, they were superimposed on the starting structure based on the C_α-coordinates using the VMD program [53]. For comparison, the simulations were repeated using the protein conformations in their ligand-bound structures (ligands were manually removed). The same procedure was applied for the simulations in methanol where the structures of the free proteins were placed in cubic boxes filled with methanol molecules (using parameters from the OPLS-AA force field) and the equilibration was extended to 500 ps. For MDM2, two additional MD runs were conducted with a modified protocol. The 10 ns simulations were conducted in water (run 1) or in methanol (run 2) with harmonic restraints on all heavy backbone atoms using a force constant of 1,000 kJ mol^{-1} nm^{-2} and the NMR structure coordinates as reference positions.

Normal mode analysis

The conformations of the free proteins were minimized in vacuo without constraints using the L-BFGS algorithm until the maximum force on any atom was smaller than 0.001 kJ mol^{-1} nm^{-1}. For the calculation of electrostatic interactions, the relative dielectric constant ε was set to 4r. Van der Waals interactions were calculated without cut-off. The hessian matrix of the minimized structure was calculated in vacuo using the same parameters. 50 eigenvectors representing the normal modes with lowest frequencies were derived from the diagonalized mass weighted hessian matrix. Eigenvectors 1–6 correspond to the translational and rotational degrees of freedom of the system and were therefore set to 0. Using the remaining 44 normal modes, 4001 protein structures were generated by random displacements along the eigenvectors at 300 K, where the position along each eigenvector was randomly taken from a Gaussian distribution with variance kT/eigenvalue. Finally, the superposition procedure described for the MD snapshots was applied to these normal mode ensemble structures.

Generation of structures using CONCOORD 2.0 and tCONCOORD

CONCOORD [34] and tCONCOORD [35] generate protein structures within a set of distance bounds that are calculated on the basis of interatomic interactions within the starting structure. The range of those distance bounds depends on the strength of the interaction. The energetically minimized structures generated for the normal mode analysis were used to determine these bounds using OPLS-AA van der Waals parameter and Engh-Huber bonded parameters [54]. Based on these distance bounds 4001 protein structures were generated and superimposed as described above. Note that the conformational ensembles generated by CONCOORD, tCONCOORD, and NMA used the same energy-minimized starting configuration as input but are otherwise unrelated.

Calculation of the RMS deviation and fluctuation from the crystal/NMR structures

All generated conformations were compared to the crystal/NMR structure of the unbound protein. The root mean square deviation of a conformation $Conf_i$ from the crystal/NMR structure was calculated by

$$RMSD(Conf_i) = \sqrt{\frac{1}{n}\sum_{j=0}^{n}\left(Coord_i(j) - Coord^{ref}(j)\right)^2} \qquad (1)$$

where $Coord_i(j)$ are the coordinates of the jth heavy atom in conformation i and $Coord^{ref}(j)$ are the coordinates of the corresponding atom in the reference structure. The RMS fluctuations were calculated for each C_α-atom $C_\alpha(i)$ according to

$$RMSF(C_\alpha(i)) = \sqrt{\frac{1}{n}\sum_{t=0}^{n}\left(C_\alpha^t(i) - C_\alpha^{ref}(i)\right)^2} \qquad (2)$$

where $C_\alpha^t(i)$ are the coordinates of the ith C_α-atom at time t and $C_\alpha^{ref}(i)$ are the coordinates of the corresponding C_α-atom in the reference structure.

Pocket detection using the PASS algorithm

In this work, a reimplemented version of the PASS algorithm [1] termed BALLPass (that is based on the BALL library [55]) was applied to each MD snapshot or generated structure, respectively. As previously described [29], a pocket was represented by one active site point (ASP) and the surrounding (coherent) probes. These voluminous imprints (also called probe patch) of the cavities were then used to calculate the pockets' solvent excluded surface volume (Note that defining each ASP to represent one pocket leads to the subdivision of very large cavities into two or even more pockets, when they consist of more than one ASP). For defining pocket polarities, all protein atoms found within a distance of 5 Å from these probe patches were identified as "pocket lining atoms" (PLAs). The polarity is then defined as the ratio of the sum of N, O, and S atoms over the sum of N, O, S, and C atoms within the set of PLAs. Furthermore, the similarity of two pockets i and j detected in two different MD snapshots was determined as

$$similarity(PLAs_i, PLAs_j) = \frac{|PLAs_i \cap PLAs_j|}{\min(|PLAs_i|, |PLAs_j|)} \quad (3)$$

These similarities were then used to cluster all cavities occurring in the 4001 structures of the set (ca. 11,000–20,000) using an agglomerative complete linkage approach. Care was taken during the clustering step to ensure that the similarity was at least 85% and that no cluster contained more than one cavity taken from the same structure. All cavities assigned to the same cluster were then defined to be states of the same transient pocket and labeled by the same pocket identifier (PID).

Two sets of transient pockets resulting from different structure generation methods or MD simulation trajectories can be efficiently compared by calculating the similarities of the so-called "subpockets". A subpocket of a transient pocket is defined as a set of PLAs that appeared in at least a third of its pocket states. The similarity of two subpockets can then be calculated by Eq. 3. If the similarity is greater than 50%, the two subpockets are defined to be equal and this transient pocket i is considered "reproduced" by the transient pocket j and vice versa.

Overlap volume as measurement for the size of the native binding pocket

After superposing all structures and the ligand bound counterpart as described above, we determined for each structure all BALLPass probes that overlapped with ligand atoms and used them to calculate a reduced form of a pocket that only consists of the native binding region of the ligand. The volume of this reduced pocket taken from the ligand bound structure was then used for calculating relative overlap volumes. These values were used to approximate the degree to which the native binding pocket opened in the structure sets. Note that for the calculation of the overlap volumes the pockets detected in this structure are considered all at once, so the probes that overlap with ligand atoms may belong to different pockets. The volumes and the polarities of the native binding pockets were also acquired from the BALLPass probes overlapping with ligand atoms in the bound structures.

Docking method

Docking experiments were performed with AutoDock 3.0.5 [30] as previously described [29]. The ligands N3B (4-(4-fluorophenyl)-N-[3-nitro-4-(2-phenylsulfanylethylamino) phenyl]sulfonylbenzamide) for BCL-X_L, FRH (5-[[2, 3-dichloro-4-[5-[1-[2-[[(2R)-2-(diaminomethylideneamino)-4-methylpentanoyl]amino]acetyl]piperidin-4-yl]-1-methylpyrazol-3-yl]phenoxy]methyl]furan-2-carboxylic acid) for IL-2, and DIZ ((2S)-2-(4-chlorophenyl)-2-[(3S)-3-(4-chlorophenyl)-7-iodo-2,5-dioxo-1,3-dihydro-1,4-benzodiazepin-4-yl]acetic acid) for MDM2 were extracted from the complex crystal structures and hydrogens were added with the AutoDockTools (ADT 1.4.3) modules of the Python Molecular Viewer software [56]. The same software was used to compute the Gasteiger atomic charges [57]. Finally the rotatable bonds were assigned with AutoTors. The number of flexible torsions was 10 for N3B, 17 for FRH, and 5 for DIZ. The protein structures were taken either from MD simulations, from the structural ensemble derived using the normal modes or generated using CONCOORD or tCONCOORD. The nonpolar hydrogens of the MD snapshots were removed and polar hydrogens were added for the CONCOORD and tCONCOORD structures. Kollman united-atom partial charges and solvation parameter were assigned using the AutoDockTools utility. The grid maps were calculated with AutoGrid. The grid centre was chosen to coincide with the centre of mass of the individual transient pocket. The default spacing of 0.375 Å between the grid points were used. Taking into account that only a terminal moiety of the ligands may be placed into the transient pocket, the grid dimensions were expanded to 30 Å × 30 Å × 30 Å for BCL-X_L and IL-2. For MDM2, the grid dimensions were reduced to 16.125 Å × 16.125 Å × 16.125 Å to confine the position of the smaller ligand to the transient pocket. For performance reasons, we restricted the docking to those transient pockets that were located at the interface (Docking into a transient pocket took 1–3 min on one 2.8 GHz Xeon CPU depending on the flexibility of the ligand and the size of the grid box).

For the docking procedure the standard Lamarckian Genetic Algorithm protocol was used with an initial population of 50 randomly placed individuals, a maximum number of 250,000 energy evaluations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value of 1. The probability of performing a local search on an individual was set to 0.06, and the maximum number of consecutive successes or failures before doubling or halving the local search step size was 4. The pseudo-Solis and Wets algorithm was applied for these local searches with a maximum of 300 iterations. 10 independent docking runs were carried out for each MD snapshot.

Results

For each system, we generated six conformational ensembles consisting of 4001 structures each: three sets of snapshots extracted from MD simulations of the bound structure in water or from the free structure in water or methanol, and three sets of structures generated using CONCOORD, tCONCOORD, or by deformations along normal modes. In the following, we will shortly refer to these conformational ensembles as bound MD snapshots, free MD snapshots (methanol) or (water), CONCOORD, tCONCOORD, or NMA conformations, respectively. For MDM2, two additional sets of snapshots resulting from the modified MD simulation protocols were generated. Whereas the MD simulations used the crystal/NMR structures of the free or bound protein as starting structure, the calculations of the CONCOORD, tCONCOORD, and NMA conformations were based on energy-minimized structures of the unbound proteins. We will first present the findings for the ensembles from the various MD techniques and then the findings for the CONCOORD, tCONCOORD, and NMA ensembles.

RMS deviations between the conformational MD ensembles and the crystal/NMR structures

Figure 1 shows the crystal/NMR structures of the free and the bound proteins and the minimized structures. The backbone RMSDs of these minimized structures from the unbound crystal/NMR structures were 1.0 Å for BCL-X_L, 1.4 Å for IL-2, and 6.6 Å for MDM2. As the bound structure of BCL-X_L includes more loops than the free structure (see Fig. 1a), MD simulation of the bound structure displayed a slightly different dynamics and thus revealed larger RMSDs (up to 5.2 Å) than the other conformational ensembles that are based on the free structure. For example, the RMSD of the free MD snapshots taken from the simulation in water stayed constantly under 2.0 Å, whereas it increased to about 2.5 Å during the simulation in methanol and appeared to stabilize at around 7.7 ns simulation time (Fig. 7a, Supplementary Materials). For IL-2, both simulations in water and methanol gave similar RMSD profiles (Fig. 7b, Supplementary Materials). As shown in Fig. 1c, the large backbone RMSD of the minimized structure of MDM2 was caused by the floppy terminal loops of the NMR structure that folded back on the protein surface during the minimization in vacuo. Since this conformation appeared too compact as a starting structure, we repeated the minimization and the subsequent generation of the conformational ensembles with the stable part of the MDM2 protein (residues 17–111), to which we will refer as "truncated MDM2". This resulted in an RMSD of 1.6 Å from the crystal structure. These unstable terminal loops are missing in the bound structure of MDM2. In the free MD snapshots, they caused RMSDs of up to 8 Å for the simulation in water and up to 9.5 Å for the simulation in methanol. In contrast, the RMSD of residues 17–111 was 4.6 Å for the simulation in methanol and 3.6 Å for the simulation in water (Fig. 7c, Supplementary Materials).

For BCL-X_L and MDM2, simulating in methanol solution allowed for transitions to regions of the conformational space that were not sampled in the simulations in aqueous solution at room temperature. On the other hand, using harmonic restraints on backbone atoms restricted the sampled conformation to a small range around the conformation of the starting structure (see Fig. 7a–c, Supplementary Materials).

The differences between the conformational ensembles and the free crystal/NMR structures were further investigated by plotting the RMS fluctuations of the C_α-atoms (see Fig. 8, Supplementary Materials). The most mobile regions of the proteins were captured by all five methods, but the magnitude of the fluctuations differed significantly. Interestingly, many residues forming the native binding pocket (residues 93, 96, 97, 101, 104, 105, 108, 126, 130, 136–139, 141, 146, 190, 195 for BCL-X_L, residues 34, 35, 38, 39, 41–45, 62, 72, 76 for IL-2, and residues 16, 17, 54, 57, 58, 61, 62, 67, 72, 91, 93, 96, 99, 100 for MDM2) belong to these mobile regions.

As expected, the simulation in methanol leads to a more pronounced exposure of hydrophobic side chains and thus a lower overall surface polarity than during the simulation in water (see Fig. 9, Supplementary Materials).

Surface pockets detected in the starting structures

Before applying BALLPass to the conformational ensemble for the identification of transient pockets, it was applied to the free and bound crystal/NMR structures to obtain reference values. The volumes and the polarities of the native binding pockets in the protein-inhibitor complexes

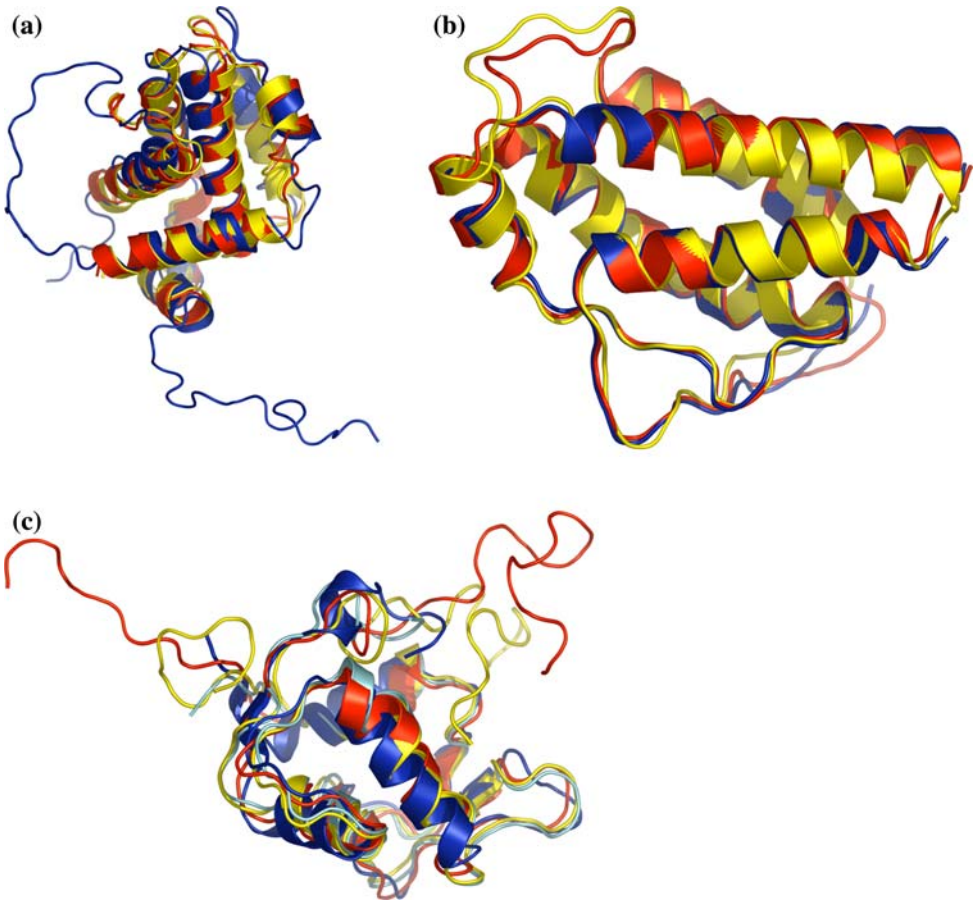

Fig. 1 Crystal/NMR structures of free (red), bound (blue), and minimized (yellow) BCL-X_L (**a**), IL-2 (**b**), and MDM2 (**c**) in cartoon representation. For BCL-X_L (panel (**a**)), the terminal loops as well as the loop formed by residues 28–44 are missing in the free structure. In MDM2 (panel (**c**)), the two terminal loops are missing in the bound structure. The minimized structure of full MDM2 is shown in yellow, the minimized structure of truncated MDM2 is shown in green

are 493.1 $Å^3$ and 0.26 for BCL-X_L, 400.3 $Å^3$ and 0.27 for IL-2, and 445.9 $Å^3$ and 0.25 for MDM2. The native binding pocket is only partly open in the free structures of BCL-X_L and IL-2 (23.1 and 25.9% volume relative to the volume of the ligand-binding pocket in the X-ray structure of the complex, respectively). For free MDM2, the native binding pocket was too compact to be detected in any NMR model. These values differ from those we published in our first study [29], because we used a different pocket detection program here (BALLPass instead of PASS).

Transient pockets detected in the structure sets

The pocket detection protocol was then applied to all conformational ensembles for identifying transient pockets. Table 1 summarizes the main properties of the identified pockets. When focusing on the differences between the pockets detected in the free and bound MD snapshots of the simulation in water, one should keep in mind that for MDM2 and BCL-X_L, the simulated proteins were not of equal size. So the larger total number of pockets (the number of pockets before clustering) and the larger number of distinct transient pockets (the number of pockets after clustering) detected in the bound form of BCL-X_L (the smaller number of pockets detected in the bound form of MDM2, respectively) is not important. However, the volume of the native binding pocket covers a broader range in the simulations started from the bound conformation as indicated by the mean and the maximum overlap volume. The pockets opening during the simulation of the free protein structure in methanol have interesting properties. For all three systems, the opening of the native binding pocket seems to be eased in the simulation of the free structure in methanol compared to the simulation of the same structure in water as the maximum overlap volume shows. Besides, more (in terms of total number) and on average larger pockets opened in methanol, suggesting that the less hydrophilic solvent facilitates the formation of cavities.

Further analysis addressed the differences in the properties between those pockets opening at the native binding site and those opening somewhere else on the protein surface (see Fig. 10, Supplementary Materials). Note that several distinct pockets are opening at the native binding site. As already mentioned in the method section, they represent different subpockets of the native binding pocket and thus may possess different polarities. In general, the cavities identified in the free MD snapshots of the simulation in methanol belong to the least polar pockets, especially when focusing on the largest pockets. In most

cases, transient pockets that are located at the native binding site are on average larger than the other pockets and have a slightly reduced polarity. Very small pockets (mean volume $\leq$ 200 Å^3) tend to be either very polar or very nonpolar. This analysis further reveals that for BCL-X_L and MDM2, all transient pocket opening at the native binding site are on average more polar than the native binding pocket, except for those opening during the MD simulation in methanol. In contrast, for IL-2, transient pockets with an average volume and polarity comparable to the native binding pocket can be identified in both MD simulations of the free structure.

Transient pockets of MDM2 detected in snapshots from restrained MD simulations

In all MD simulations presented so far, the whole protein was flexible so that it is hard to say whether side chain movements alone are sufficient for the observed pocket openings or not. Therefore, we now harmonically restrained the backbone atoms of the MDM2 protein in the modified MD simulation protocol. The impact on the pocket properties is shown in Table 1. As expected, the number of pockets (their total number and the number of distinct transient pockets) and their volume were reduced when their formation depends exclusively on side chain movements. Using methanol as solvent not only led to larger pockets, but also their number increased considerably compared to the simulation in water. Interestingly, this effect is observable in the unrestrained and the restrained MD simulations. The native binding pocket opened in all MD simulations, but its native volume was only reached when simulating without restraints in water or methanol.

Figure 10c (Supplementary Materials) compares the properties of the pockets opening during the different MD simulations. Interestingly, in all MD simulations the

Table 1 Comparison of the pockets detected in the conformational ensembles for each system

	# Pockets before clustering	# Pockets after clustering	Mean pocket volume (Å^3)	Maximum pocket volume (Å^3)	Mean overlap volume (%)	Maximum overlap volume (%)
BCL-X_L						
Free MD snapshots (water)	17,079	24	375.1	1363.9	43.5	91.3
Free MD snapshots (methanol)	22,818	23	380.7	1357.9	48.3	95.0
Bound MD snapshots (water)	39,596	46	395.2	1606.9	38.7	94.6
CONCOORD	8,226	11	348.1	1013.1	34.3	67.1
tCONCOORD	17,135	23	378.7	1761.3	42.5	99.9
NMA	7,774	6	342.7	889.8	32.0	50.8
IL-2						
Free MD snapshots (water)	14,721	29	335.2	1013.0	35.3	89.2
Free MD snapshots (methanol)	24,513	24	352.7	1633.9	35.0	94.3
Bound MD snapshots (water)	18,412	33	394.4	1340.7	32.6	128.1[a]
CONCOORD	14,096	18	320.1	965.2	34.8	88.5
tCONCOORD	17,356	28	365.4	1422.9	33.3	92.3
NMA	15,345	11	261.1	722.4	50.1	87.9
MDM2						
Free MD snapshots (water)	26,419	35	372.7	1641.3	60.1	106.8[a]
Free MD snapshots (methanol)	34,022	42	407.6	2204.5	47.6	108.0[a]
Bound MD snapshots (water)	11,737	14	384.6	1213.8	57.5	114.3[a]
Free MD snapshots (water/restrained BB)	21,568	17	330.5	839.9	52.9	81.2
Free MD snapshots (methanol/restrained BB)	25,204	22	351.8	1024.3	52.2	82.0
CONCOORD	13,090	15	267.1	641.8	13.2	55.3
CONCOORD[b]	13,519	10	348.6	828.2	48.0	70.1
tCONCOORD[b]	16,090	14	388.6	1237.7	2.4	8.4
NMA	16,546	7	260.4	433.9	0	0
NMA[b]	14,776	7	322.2	670.2	45.4	61.4

[a] Overlap volume is larger than in complex crystal structure

[b] Minimized structure of truncated MDM2 (residues 17–111) used as starting structure

transient pockets with the largest mean volume opened at the native binding site. However, the pockets are more polar than in the complex structure except for those in the unrestrained simulation in methanol. As already shown in Table 1, the transient pockets observed during the restrained MD simulations are relatively small. Pockets with mean volumes that exceed the volume of the native binding pocket were only found in the MD snapshots of the restrained simulation in methanol.

Focusing on the similarity of the sets of transient pockets opening under different simulation conditions as shown in Fig. 2c reveals that some pockets can only open when backbone movements are allowed. This emphasizes the intrinsic influence of backbone movements on pocket openings.

How "open" must a pocket be to accommodate an inhibitor?

In many structures, the native binding pocket is only partly open (see Table 1) and it is unclear whether an overlap value of 50–90% is sufficient to accommodate a ligand. However, this measure is only a rough estimate. Therefore, ligand docking was performed using as starting points all transient pockets that opened at the interface in all snapshots extracted from MD simulation of the unbound structure in water or methanol (see Table 2). Interestingly, for BCL-X_L and MDM2 much better docking results were obtained when using the MD snapshots in methanol compared to those in water reflected by the better docking scores and the reduced score rank. Only for IL-2, docking into water snapshots gave better results (although the maximal overlap in Table 1 gives another impression). A possible explanation for this may be that the pockets opening in methanol at the native binding site are too small for the native ligand (see Fig. 10b, Supplementary Materials), even though the pockets are on average larger than those opening in water.

The docking results for the snapshots of MDM2 extracted from the MD simulations with restrained backbone shown in Table 2 reflect the importance of backbone movements on the opening of the native binding pocket. Although side chain movements are sufficient to open new cavities, suitable backbone movements are also needed to achieve enough depth and plasticity. Consistent with Table 1, the results are slightly better in the docking using snapshots from a restrained simulation in methanol solvent. The effect of methanol as solvent is most striking for the snapshots taken from the unrestrained MD simulations. The relative score rank of a correct docking result improved to 0.2%. This means that even if the native binding pose is unknown, it can be found among the 0.2% of the docking results with best score. Among 1,000 docking poses, this would correspond to the second best scored docking solution. Furthermore, the score was even 1.7 kcal/mol more favorable than when docking into a snapshot extracted from the unrestrained MD simulation in water. This indicates that, in addition to backbone movements, the effect of a less polar solvent promotes the opening of pockets even further. Examples for different shapes of the binding pockets resulting form different simulation conditions and the docked ligand poses reported in Table 2 are shown in Fig. 3 and in Figs. 12–14 (Supplementary Materials).

Results for CONCOORD, tCONCOORD, and NMA

For all three systems, the CONCOORD and the NMA methods generated ensembles with much smaller RMS deviations from the experimental structure than the MD methods (see Fig. 7, Supplementary Materials). The RMSD values for the NMA conformations were nearly constant (RMSD variation 0.1–0.2 Å) and only slightly higher than the RMSD of the minimized structure, whereas the RMSD values of the CONCOORD conformations varied up to 2.0 Å. In contrast, due to the enhanced conformational sampling of the tCONCOORD method, it generated conformations that differ up to 4.2 Å from the NMR or crystal structures and thus lie in about the same order of magnitude as most MD simulations.

As these conformational ensembles were generated from the minimized structures of the unbound proteins, it is of interest to know whether and to which extent the native binding pocket is open in these structures. Interestingly, minimizing free IL-2 in vacuo resulted in a further opening to 74.8%, whereas in BCL-X_L the native binding pocket closed. In the minimized structure of full MDM2, the N-terminal loop buries the native binding site. In the minimized structure of truncated MDM2, the native binding pocket opened to 50.8%.

Again, pockets on the protein surfaces were identified in the CONCOORD, tCONCOORD, and NMA conformational ensembles using the standard pocket detection protocol (see Table 1). The truncated MDM2 resulted in significantly larger cavities and the opening of the native binding pocket was not hindered. Although the total number of cavities found in the CONCOORD or the NMA conformational ensembles is quite high, they are not as diverse (indicated by the number of pockets after clustering) as those opening during MD simulations. Especially when using NMA for structure generation one ends up with only a few transient pockets that all appear unusually frequent in the structure set (see Supplementary Materials, pie charts in Fig. 11a–c). Moreover, the pockets found in the CONCOORD and NMA conformational ensembles are smaller than those in the MD snapshots. More importantly,

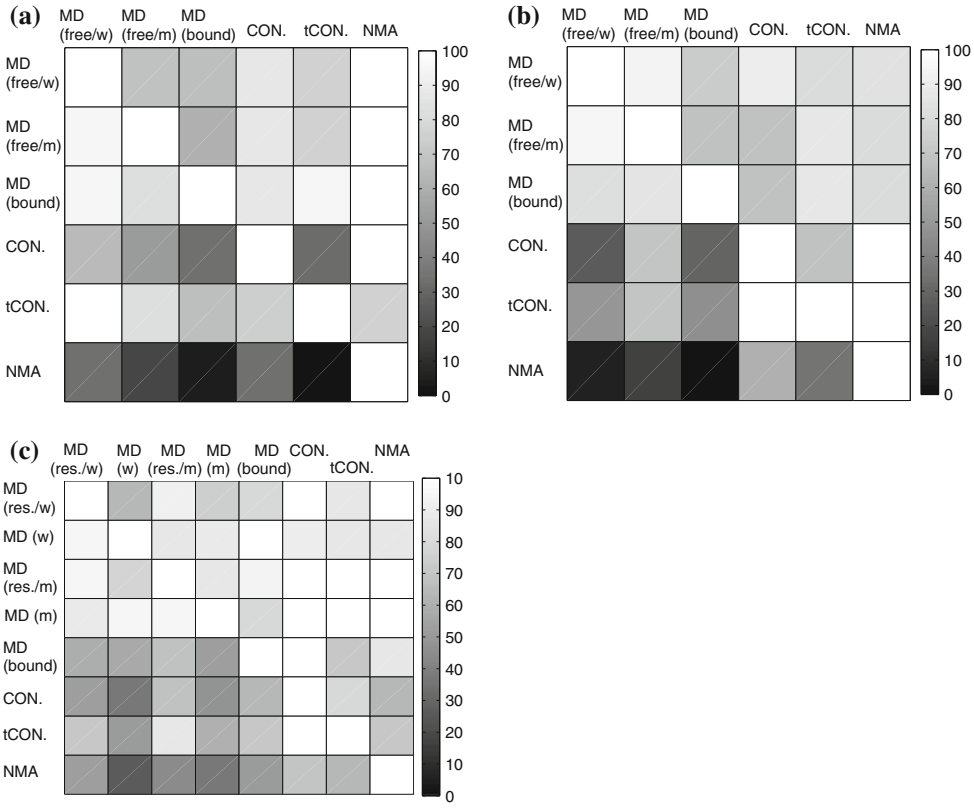

Fig. 2 Reproducibility according to Eq. 3 of the detected transient pockets in one conformational ensemble by the transient pockets detected in another conformational ensemble for BCL-X_L (**a**), IL-2 (**b**), and MDM2 (**c**). In each column, the percentage of reproduced transient pockets of one method by the other methods (rows) is shown

they are also smaller than the native binding pockets and too polar, except for IL-2. For all three systems, varying interatomic distances in the CONCOORD approach or random deformations along normal modes approach sometimes resulted in an opening or enlargement of the native binding pocket, but not to the same extent as observed in MD simulations. Only for IL-2, the overlap volumes of the NMA and CONCOORD conformations are of the same magnitude than those of the free MD snapshots.

Using the tCONCOORD instead of the CONCOORD method significantly improves the sampling of transient pockets. Not only the total number of cavities increased, but also their diversity as indicated by the number of pockets after clustering. Moreover, these pockets are considerably larger than those detected in the CONCOORD and NMA conformational ensembles. Their volumes are even in the same order of magnitude as the pockets opening during the MD simulations, but except for IL-2 they are too polar (see Fig. 10, Supplementary Materials).

Finally, we tested whether transient pockets found in one conformational ensemble are also contained in other ones. Figure 2 shows that most pockets found in NMA structures were also found in all other conformational ensembles, because, as stated above, the NMA structures tend to be quite similar and thus possess only a small number of distinct pockets. Apparently, the slow normal modes are not involved in the dynamics of all pocket opening regions. Although NMA, CONCOORD, and tCONCOORD structures are based on the same starting structure, the pockets detected in the conformational ensemble generated with CONCOORD and tCONCOORD are more diverse, but they were also not able to reproduce most pockets found in MD snapshots. Here again, the advantage of tCONCOORD over CONCOORD is obvious as the cavities detected in the tCONCOORD conformations include more of those detected during the MD simulations. Taking into account the different size of the bound and free structure of BCL-X_L and MDM2, all MD simulations yielded a relatively similar set of pockets. These results indicate that although CONCOORD and NMA could produce conformations with pockets not present in the crystal structure, these pockets are not as diverse as those observable in MD snapshots. At least for BCL-X_L and MDM2, however, tCONCOORD seems to be an efficient alternative to MD simulations.

As before, all transient pockets that opened at the interface in conformational ensembles generated using normal modes, CONCOORD, and tCONCOORD were used as starting points for docking experiments. The results shown in Table 2 emphasize that none of the NMA or CONCOORD conformations contains a ligand binding

Table 2 Best ranked correct (rmsd ≤ 2 Å) docking results or docking results with lowest RMSD per conformational ensemble and system

	RMSD (Å)	AutoDock score (kcal/mol)	AutoDock score rank[a] (%)
BCL-X_L			
Free MD snapshots (water)	1.9	−10.2	1.5
Free MD snapshots (methanol)	1.7	−11.8	0.5
CONCOORD	1.8	−7.6	29.0
tCONCOORD	2.0	−11.1	1.5
NMA	2.3	−8.1	16.3
IL-2			
Free MD snapshots (water)	1.9	−8.5	0.7
Free MD snapshots (methanol)	2.0	−6.9	8.4
CONCOORD	2.0	−6.6	7.9
tCONCOORD	2.4	−5.4	25.8
NMA	1.8	−7.0	7.1
MDM2			
Free MD snapshots (water)	1.5	−11.8	8.0
Free MD snapshots (methanol)	1.7	−13.5	0.2
Free MD snapshots (water/restrained backbone)	2.0	−8.8	82.6
Free MD snapshots (methanol/restrained backbone)	2.0	−9.5	68.5
CONCOORD	4.85	−7.5	87.1
CONCOORD[b]	2.1	−8.7	61.6
tCONCOORD[b]	1.6	−11.2	0.1
NMA	–	–	–
NMA[b]	3.3	−9.7	13.0

[a] Relative rank defined as the rank of this solution after sorting all results by increasing docking score in relation to the total number of docking results

[b] Minimized structure of truncated MDM2 (residues 17–111) used as starting structure

–, No docking results because no pockets opened at the interface

pocket as close to its native conformation as was found in the MD simulations started from the unbound structure. Only for BCL-X_L and IL-2, RMSD values below 2 Å could be achieved when docking into CONCOORD conformations (for IL-2 also when docking into NMA conformations), but the scores of these results demonstrate that the structural and/or the biochemical environment is not as appropriate as it may get using MD simulations without restraints. Docking into conformations generated by tCONCOORD led to surprising results. For IL-2, the docking results were significantly worse than those for the CONCOORD conformations. In contrast, the results for BCL-X_L and MDM2 are comparable to those obtained when docking into the snapshots extracted from MD simulation of the free structures in water. This emphasizes the ability of tCONCOORD of sampling ligand-bound conformations even if the unbound structure was used as input.

Discussion

During all MD simulations, transient pockets of similar size as with the known inhibitor bound could be observed at the native binding site. In fact, a quite diverse set of transient pockets was observable during MD simulations. We have previously shown that these pockets are reproducible [29]. Furthermore, the results of this study indicate that most pockets are also reproducible in MD simulations under different conditions. Interestingly, the properties of the cavities depended crucially on the complexation status of the starting structure (apo vs. bound) and the solvent. The general impressions seen for all three systems were quite similar. When the ligand bound structure was used as starting structure, the volume of the native binding pocket showed the largest fluctuations. When using methanol as solvent, more and larger pockets opened on the protein surface compared to the simulation of the same starting structure in water. The restrained simulation of MDM2 showed that side chain movements alone indeed lead to the formation of pockets, but the number and the volume of the pockets are reduced. These findings suggest that pocket openings are induced by movements of the protein backbone and side chains that are coupled to the solvent.

The analysis of the transient pockets detected in the six conformational ensembles revealed that only a limited number of distinct cavities are observable when CONCOORD or NMA were used to generate the structural ensemble. Almost all of these transient pockets were also found during MD simulations. Moreover, these pockets are relatively small and in most cases not appropriate for ligand binding. The quality of the docking results with the CONCOORD or NMA ensembles was not as good as for the MD ensembles with explicit solvent. Therefore, the applicability of CONCOORD or NMA for the purpose of inducing pocket openings seems to be limited by their neglect of solvent effects. tCONCOORD seems to overcome this problem by not considering hydrogen bonds that may be attacked by solvent molecules in the definition of the distance constraints. This results in an enhanced conformational sampling compared to CONCOORD and, thus, to pockets of an increased variety and volume that is even comparable to that of pockets opening during MD simulations.

The conformational changes necessary for the opening of the native binding pocket do seem to be related somehow to normal modes as suggested by the overlap volumes in NMA conformations, but deformations along these normal modes alone are not sufficient to induce full pocket

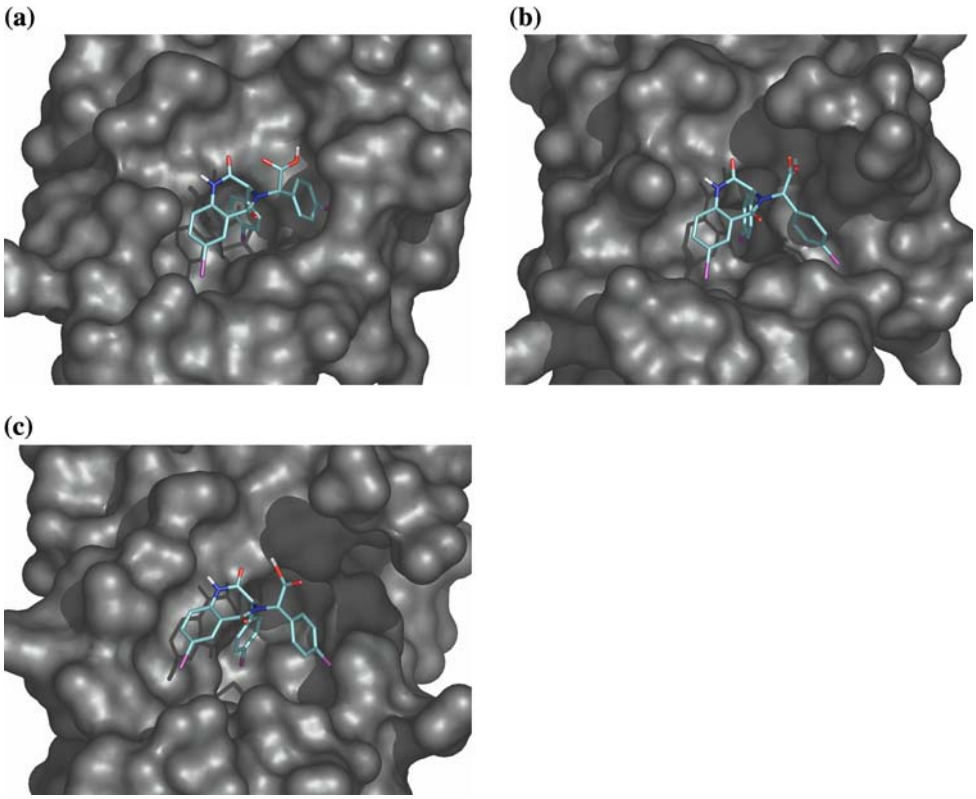

Fig. 3 The shape of the binding site in different protein conformations of MDM2 (shown in surface representation) and the corresponding ligand (shown in licorice representation) binding mode. Shown are the ligand bound complex crystal structure (**a**) and the snapshots extracted from MD simulations in water (**b**) and methanol (**c**) without restraints that achieved the best ranked docking pose with RMSD $\leq$ 2 Å (listed in Table 2). For better visibility, the N-terminal loop (residues 1–16) was deleted in these representations

openings. Even if only those normal modes were considered that have the highest impact on the binding site and the temperature was increased to 350 K, no significant improvement in terms of pocket volumes and docking scores could be obtained (Table 4, Supplementary Materials). This finding that the opening of the native binding pocket is weakly related to normal modes is in agreement with the results reported by Barrett et al. [43] and Espinoza-Fonseca and Trujillo-Ferrara [44] for MDM2. However, these authors only observed the opening and closing movement and did not measure whether the increase of the pocket volume is sufficient to bind a ligand.

In our study, only the native binding pocket of IL-2 could be approximately reproduced by application of CONCOORD and by deformations along the normal modes of the protein. One should keep in mind, however, that here the native binding pocket was already open to almost 75% in the minimized conformation that was used as starting structure. For this system the opening of the native binding pocket appears energetically quite favorable and may be observed by a variety of methods that sample low-energy conformations. We suggest that no conformation near the ligand-bound structure could be generated here using tCONCOORD because the minimized conformation we used as input was already quite similar to the bound structure. Whereas CONCOORD and NMA sample conformations quite similar to the starting structure, tCONCOORD allows for large-scale conformational transitions that drive the sampling further away from the ligand-bound structure. Thus, no native-like binding pose could be found when docking into the tCONCOORD conformational ensemble. In the case of BCL-X_L on the other hand, minimization resulted in the closure of the previously partly open native binding pocket. Here, only CONCOORD and tCONCOORD produced conformations that were able to accommodate the native ligand.

The quality of the CONCOORD and NMA results crucially depend on the starting structure used as the example of MDM2 pointed out (see Table 1). In the minimized conformation of the full protein, the N-terminal loop obstructs the p53 binding groove and thus hinders the opening of the native inhibitor binding pocket. Such a conformation was suggested earlier [58], but in their publication describing the NMR models we used for this study, Uhrinova et al. stated that no long-range NOEs were observed for residues 2–17 [59]. This rules out the possibility that this loop occupies the p53 binding groove in a stable fashion and suggests that our approach of truncating these residues is valid.

Conclusions

On the example of three protein systems we showed that the native conformational dynamics of proteins on picosecond timescales as captured by MD simulations leads to

the opening of many diverse pockets on the protein surface. Most of them were observable using either apo or bound starting structures and in water and methanol solvent. For all three model systems, transient pockets can be identified that opened at the native binding site. Docking a known inhibitor into these pockets reproduced the native binding modes with less than 2 Å RMS deviation from the crystal structure conformation. During the MD simulations in methanol, the total number of pocket openings and their volumes increased compared to the simulation in water. For two out of three systems, the simulations in methanol led to the formation of nonpolar pockets at the interface which significantly improved the docking results. This suggests that a more hydrophobic solvent facilitates the opening of the native binding pocket. Comparing MD simulations of MDM2 with full flexibility or with harmonically restrained backbone atoms revealed that although side chain movements alone lead to the formation of surface cavities, the required depth and plasticity for ligand binding can only be achieved by including the backbone movements. The efficient structure generation methods CONCOORD and NMA were capable of producing conformations with pockets not observable in the starting structure, but with limited diversity and volume. Although the formation of some pockets is coupled to low frequency normal modes, deformations along these modes are not sufficient to achieve full pocket openings. On the other hand, conformations generated by tCONCOORD possessed pockets with volumes and diversity that were comparable to those of pockets opening during MD simulations. For two out of the three test systems, this method was even able to generate conformations suitable to accommodate the native ligand.

These findings open promising avenues for structure-based drug design on protein surfaces.

Acknowledgments This work was funded by a grant of the Deutsche Forschungsgemeinschaft (Center for Bioinformatics, Saarland University). We thank Jan Fuhrmann and Dirk Neumann for making available their BALLPass implementation to us.

References

1. Brady GP, Stouten PFW (2000) J Comput Aided Mol Des 14:383. doi:10.1023/A:1008124202956
2. Hendlich M, Rippmann F, Barnickel G (1997) J Mol Graph Model 15:359. doi:10.1016/S1093-3263(98)00002-3
3. Levitt DG, Banaszak LJ (1992) J Mol Graph 10:229. doi:10.1016/0263-7855(92)80074-N
4. Laskowski RA (1995) J Mol Graph 13:323. doi:10.1016/0263-7855(95)00073-9
5. Laskowski RA, Luscombe NM, Swindells MB, Thornton JM (1996) Protein Sci 5:2438
6. Liang J, Edelsbrunner H, Woodward C (1998) Protein Sci 7:1884
7. Nayal M, Honig B (2006) Proteins 63:892. doi:10.1002/prot.20897
8. Landon MR, Lancia DR Jr, Yu J et al (2007) J Med Chem 50:1231. doi:10.1021/jm061134b
9. Goodford PJ (1985) J Med Chem 28:849. doi:10.1021/jm00145a002
10. Laurie ATR, Jackson RM (2005) Bioinformatics 21:1908. doi:10.1093/bioinformatics/bti315
11. Berg T (2003) Angew Chem Int Ed Engl 42:2462. doi:10.1002/anie.200200558
12. Yin H, Hamilton AD (2005) Angew Chem Int Ed Engl 44:4130. doi:10.1002/anie.200461786
13. Wells JA, McClendon CL (2007) Nature 450:1001. doi:10.1038/nature06526
14. Ryan KM, Phillips AC, Vousden KH (2001) Curr Opin Cell Biol 13:332. doi:10.1016/S0955-0674(00)00216-7
15. Michael D, Oren M (2003) Semin Cancer Biol 13:49. doi:10.1016/S1044-579X(02)00099-8
16. Hollstein M, Sidransky D, Vogelstein B, Harris CC (1991) Science 253:49. doi:10.1126/science.1905840
17. Zheleva DI, Lane DP, Fischer PM (2003) Mini Rev Med Chem 3:257. doi:10.2174/1389557033488178
18. Fry DC, Emerson SD, Palme S et al (2004) J Biomol NMR 30:163. doi:10.1023/B:JNMR.0000048856.84603.9b
19. Vassilev LT, Vu BT, Graves B et al (2004) Science 303:844. doi:10.1126/science.1092472
20. Grasberger BL, Lu T, Schubert C et al (2005) J Med Chem 48:909. doi:10.1021/jm049137g
21. Arkin MR, Wells JA (2004) Nat Rev Drug Discov 3:301. doi:10.1038/nrd1343
22. Apostolakis J, Plueckthun A, Caflisch A (1997) J Comput Chem 19:21. doi :10.1002/(SICI)1096-987X(19980115)19:1<21::AID-JCC2>3.0.CO;2-0
23. Zacharias M (2004) Proteins 54:759. doi:10.1002/prot.10637
24. Meiler J, Baker D (2006) Proteins 65:538. doi:10.1002/prot.21086
25. Sousa FS, Fernandes PA, Ramos MJ (2006) Proteins 65:15. doi:10.1002/prot.21082
26. Lin JH, Perryman AL, Schames JR, McCammon JA (2002) J Am Chem Soc 124:5632. doi:10.1021/ja0260162
27. Amaro RE, Baron R, McCammon JA (2008) An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. J Comput Aided Mol Des. doi:10.1007/s10822-007-9159-2
28. Bowman AL, Nikolovska-Coleska Z, Zhong H et al (2007) J Am Chem Soc 129:12809. doi:10.1021/ja073687x
29. Eyrisch S, Helms V (2007) J Med Chem 50:3457. doi:10.1021/jm070095g
30. Morris GM, Goodsell DS, Halliday RS et al (1998) J Comput Chem 19:1639. doi :10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B
31. Lide DR (2005) Handbook of chemistry and physics. CRC Press, Boca Raton, FL
32. Alonso DO, Daggett V (1995) J Mol Biol 247:501. doi:10.1006/jmbi.1994.0156
33. Kovacs H, Mark AE, Johansson J, van Gunsteren WF (1995) J Mol Biol 247:808
34. de Groot BL, van Aalten DMF, Scheek RM et al (1997) Proteins 29:240. doi :10.1002/(SICI)1097-0134(199710)29:2<240::AID-PROT11>3.0.CO;2-O
35. Seelinger D, Haas J, de Groot BL (2007) Structure 15:1482. doi:10.1016/j.str.2007.09.017
36. Brooks B, Karplus M (1983) Proc Natl Acad Sci USA 80:6571. doi:10.1073/pnas.80.21.6571
37. Levitt M, Sander C, Stern PS (1985) J Mol Biol 181:423. doi:10.1016/0022-2836(85)90230-X

38. Ma J (2005) Structure 13:373. doi:10.1016/j.str.2005.02.002
39. van Aalten DMF, de Groot BL, Findlay JBC et al (1997) J Comput Chem 18:169. doi :10.1002/(SICI)1096-987X(19970130)18:2< 169::AID-JCC3>3.0.CO;2-T
40. Zheng W, Doniach S (2003) Proc Natl Acad Sci USA 100:13253. doi:10.1073/pnas.2235686100
41. Bruccoleri RE, Karplus M, McCammon JA (1986) Biopolymers 25:1767. doi:10.1002/bip.360250916
42. Fu X, Apgar J, Keating AE (2007) J Mol Biol 371:1099. doi: 10.1016/j.jmb.2007.04.069
43. Barrett CP, Hall BA, Noble ME (2004) Acta Crystallogr D Biol Crystallogr 60:2280. doi:10.1107/S0907444904019171
44. Espinoza-Fonseca LM, Trujillo-Ferrara JG (2006) Biopolymers 83:365. doi:10.1002/bip.20566
45. Berman HM, Westbrook J, Feng Z et al (2000) Nucleic Acids Res 28:235. doi:10.1093/nar/28.1.235
46. de Bakker PI, DePristo MA, Burke DF, Blundell TL (2003) Proteins 51:21. doi:10.1002/prot.10235
47. Lindahl E, Hess B, van der Spoel D (2001) J Mol Model 7:306
48. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) J Am Chem Soc 118:11225. doi:10.1021/ja9621760
49. Jorgensen WL, Chandrasekhar J, Madura JD et al (1983) J Chem Phys 79:926. doi:10.1063/1.445869
50. Darden T, York D, Pedersen L (1993) J Chem Phys 98:10089. doi:10.1063/1.464397
51. Berendsen HJC, Postma JPM, van Gunsteren WF et al (1984) J Chem Phys 81:3684. doi:10.1063/1.448118
52. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) J Comput Chem 18:1463. doi :10.1002/(SICI)1096-987X(199709)18:12< 1463::AID-JCC4>3.0.CO;2-H
53. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14:33. doi:10.1016/0263-7855(96)00018-5
54. Engh RA, Huber R (1991) Acta Crystallogr A 47:392. doi: 10.1107/S0108767391001071
55. Kohlbacher O, Lenhof HP (2000) Bioinformatics 16:815. doi: 10.1093/bioinformatics/16.9.815
56. Sanner MF (1999) J Mol Graph Model 17:57
57. Gasteiger J, Marsili M (1980) Tetrahedron 36:3219. doi: 10.1016/0040-4020(80)80168-2
58. McCoy MA, Gesell JJ, Senior MM, Wyss DF (2003) Proc Natl Acad Sci USA 100:1645. doi:10.1073/pnas.0334477100
59. Uhrinova S, Uhrin D, Powers H et al (2005) J Mol Biol 350:587. doi:10.1016/j.jmb.2005.05.010