Modeling Loop Backbone Flexibility in Receptor-Ligand Docking Simulations

Johannes Flick, Frank Tristram, and Wolfgang Wenzel*

The relevance of receptor conformational change during ligand binding is well documented for many pharmaceutically relevant receptors, but is still not fully accounted for in in silico docking methods. While there has been significant progress in treatment of receptor side chain flexibility sampling of backbone flexibility remains challenging because the conformational space expands dramatically and the scoring function must balance protein–protein and protein–ligand contributions. Here, we investigate an efficient multistage backbone reconstruction algorithm for large loop regions in the receptor and demonstrate that treatment of backbone receptor flexibility significantly improves binding mode prediction starting from apo structures and in cross docking simulations. For three different kinase receptors in which large flexible loops reconstruct upon ligand binding, we demonstrate that treatment of backbone flexibility results in accurate models of the complexes in simulations starting from the apo structure. At the example of the DFG-motif in the p38 kinase, we also show how loop reconstruction can be used to model allosteric binding. Our approach thus paves the way to treat the complex process of receptor reconstruction upon ligand binding in docking simulations and may help to design new ligands with high specificity by exploitation of allosteric mechanisms.

Introduction

Fueled by strong multidisciplinary efforts in the life sciences the amount of available genetic and structural data related to many biological processes and their regulation has tremendously increased. This information is used in a variety of approaches to understand the molecular mechanisms of diseases and in the design of new drugs. It is therefore puzzling that despite these efforts the number of newly approved drugs stagnates.[1–3] The development of improved discovery approaches and validation techniques for new-scaffold small-molecule compounds may help improve this situation. In silico virtual ligand screening has long been proposed to contribute to the progress in the identification of new lead compounds, because the ever-increasing computational power makes it possible to screen increasingly large ligand databases. This approach is potentially cost-effective, because it avoids unnecessary synthesis and can exploit the large variety of available structural information.[4] Despite some success stories, limitations in current in silico screening approaches nevertheless restrict their accuracy and general applicability.[5–9]

The success rate of structure-based docking methods depends crucially on the accuracy of the structural model of the complex. The role of the receptor flexibility upon ligand binding has been well documented experimentally for several therapeutically important receptors (see Refs. 10–12 and references therein). In many cases the complexity of this problem can be illustrated by simply comparing the ligand-free (apo) and ligand-bound (holo) protein crystal structures. This comparison demonstrates the adaptation of ligand-specific conformations of the receptor in the complex, and crossdocking experiments into rigid crystal structures show that a specific receptor conformation is often unable to accommodate (or score with high affinity) ligands that correspond to another structure.

Docking simulations, which do not consider any receptor flexibility, in the following called "rigid" docking simulations, restrict the conformational space which leads to errors in the identification of the correct binding mode for ligands that require receptor reconstructions. This results in underestimation of the affinity for novel ligands. Screening applications based on a single protein structure are biased toward compounds of high molecular similarity and chemotype of the cocrystallized ligand (in holo structures).[13] The influence of the receptor flexibility in ligand binding differs among targets, ranging from only small side chain movements to large reconstruction of loop regions. The degree to which these conformational changes can be modeled can significantly impact the results of docking simulations.[14,15]

A large range of models and computational tools have been developed to accommodate protein flexibility into docking-methods: Many methods are able to treat protein side chains flexibly,[16–18] but also a flexible treatment of protein backbone regions upon docking simulations has been realized.[10,12,19] Here, different approaches were developed: the sampling of normal backbone modes,[20–23] the multiple receptor ensemble docking,[24–32] a global energy optimization in internal coordinates,[33,34] the explicit modeling of holo protein structures[35] or a backbone conformation sampling via conformer libraries.[34] Furthermore, backbone flexibility was applied to

J. Flick, F. Tristram, and W. Wenzel
Institute of Nanotechnology, Karlsruhe Institute of Technology, P.O. Box 3640, 76021 Karlsruhe, Germany
E mail: wolfgang.wenzel@kit.edu

different therapeutically important targets.[36–38] For a recent review see Refs. 39 and 40.

The conformational variability of the binding partners prior to their association (conformational selection) and their conformational adaptation upon binding (induced-fit) are thought to be responsible for conformational changes of protein and ligand in the binding process. Both mechanisms are closely related and can be used as starting points for modeling the variety of protein holo structures observed in experiments. Correspondingly, there are two main groups of methods to treat receptor flexibility: (1) methods, which use multiple receptor conformations, known from experiment or simulation, to represent flexible regions of the receptor (ensemble-docking approach) and (2) methods, which sample the receptor flexibility explicitly during the docking simulation (induced-fit approach).

Recently, there has been experimental and theoretical evidence on protein dynamics that induced-fit and conformational receptor selection are limiting cases of a wide spectrum of possible scenarios.[41] The data suggests that initial binding is dominated by conformational selection and further receptor conformational changes are based on optimization of binding interactions. The two successive stages of the binding process were observed by kinetic measurements.[42–44]

The ensemble view explains receptor conformational changes observed upon ligand binding with the intrinsic protein flexibility and the roughness in the energy landscape of the native protein.[45–47] This view assumes that the native state of a protein comprises a thermodynamically accessible ensemble of multiple possible receptor conformations. It is presumed that some of the low energy conformations (meaning free-energy microstates) are isoenergetic, hence almost equally populated. Ligand binding then redistributes the statistical weight within the ensemble to the preferred ligand-bound conformation resulting in selective binding to one specific receptor conformation. Experimental results[48] support this view: observing fluctuations between open and closed states in a adenylate kinase,[49] the free-energy landscape of myoglobin and other heme proteins with photo dissociation experiments,[50] and the detection of slow backbone movement in the DFG-motif (Aspartic acid, Phenylalanine and Glycine) region in the apo structure for p38 kinases.[51]

The in silico equivalent of the ensemble view increases the conformational space by introducing additional receptor conformations into the docking protocol, but often without considering explicitly the receptor internal free energy. The conformations used are usually deducted from experimentally known holo structures or homology models, thus 'knowledge-based," which may lead to the inclusion of unphysical conformations or to the missing of relevant conformations. Due to the neglect of receptor reorganization free energy the contribution of receptor reorganization to the affinity remains unclear. For example, it has been shown that the reorganization energies of the receptor in variable conformations can be in the same order of magnitude as the differences in binding free energy.[11]

The opposite end of the spectrum is represented by the induced fit approach,[52] which explains the structural changes in receptor conformation upon binding by the proteins response to a perturbation caused by the ligand.[40] In computational models implementing this view, the receptor conformation changes during the simulation, which requires a balanced model for the receptor reorganization energy in the scoring function and an extremely efficient sampling procedure. Its principal advantages are that the relevant receptor conformational space need not be known prior to the simulation and that the total affinity and not just the ligand binding free energy are approximated by the scoring function. In principle, this approach has a higher accuracy than the ensemble docking approach, but it has to struggle with a high number of degrees of freedom in the system.

Backbone motion is presently rarely observed in the most costly molecular simulations, thus most established docking protocols model induced fit only at the side chain level.[16,53] For some applications this limitation leads to acceptable performance, because it was shown that in ligand binding for many receptors only few side chains are moving.[54] In many cases only three or even less side chains change their conformation.[14] To reduce the computational demands the usage of rotamer libraries has been realized.[55,56] Side chain selection proceeds via physics (hydrophobicity[54] or polarity[57]) criteria, chemical institution,[53] analysis of known binding modes[58] or observed flexibility in experimental structures,[59] sequence conservation[60] or analysis of the structural features in the binding site.[61] However, it was shown that the introduction of flexibility for residues with Lennard-Jones clashes in rigid docking can improve the docking performance.[16]

While side chain flexibility is today implemented in many docking methods, flexibility for continuous backbone regions remains very challenging[40] limiting the applicability of this methodology for many important applications, e.g., in kinases and the few GPCR receptors for which crystal structures are available. In this investigation, we attempt to improve this situation by implementing a novel backbone reconstruction algorithm that can modify the conformation of preselected extended backbone regions with high efficiency during the simulations. Using this approach we demonstrate for three important kinases that binding modes that cannot be found in rigid backbone receptor simulations are indeed selected in simulations using the backbone reconstruction protocol. We also show that this approach permits modeling of allosteric effects by sampling both the DFG-in and DFG-out motifs in an application to p38α kinase, an important target for novel anti-inflammatory therapies.[62]

Methods

Docking method

All docking simulations have been performed with the all-atom receptor-ligand docking program FlexScreen,[53] which employs the following biophysically motivated scoring function:

$$S = \sum_{i,j} \left(\frac{R_{i,j}}{r_{i,j}^{12}} - \frac{A_{i,j}}{r_{i,j}^{6}} + \delta_{i,j} \frac{q_i q_j}{\varepsilon r_{i,j}} \right) + \sum_i \Phi_i q_i + \sum_{h-bonds} \cos\ \theta_{i,j} \left(\frac{\tilde{R}_{i,j}}{r_{i,j}^{12}} - \frac{\tilde{A}_{i,j}}{r_{i,j}^{10}} \right)$$

with $\delta_{i,j} = \mathbf{1}$, for the pairwise interaction of two atoms, which are allowed to change position during simulation and $\delta_{i,j} = \mathbf{0}$ otherwise. Φ_i represents the electrostatic potential of the rigid protein parts at the position of atom *i*. The scoring function contains Lennard-Jones (first two terms), electrostatic Coulomb (term three and four, $\varepsilon = 4$), and angular dependent hydrogen bond (terms five and six) potentials. Protein and ligand atoms (*i,j*) are treated on the same footing. The parameters used for Lennard-Jones and the hydrogen bonding are the same as in AutoDock.[63]

To speed up the docking simulations FlexScreen splits the Coulomb and Lennard-Jones receptor energy contributions in the scoring function into different contributions for rigid and flexible parts of the system, respectively. The terms stemming from the rigid parts are evaluated via an efficient grid-based procedure. Grid files for the Lennard-Jones interactions (using FlexScreen's stand-alone program doGrid) and for the electrostatic interactions (using APBS[64]) are preprocessed. APBS provides a adaptive Poisson–Boltzmann equation solver, which yields a high accuracy electrostatic model for regions with different dielectric constants. While APBS includes electrostatic effects of the solvent on the ligand, differential ligand intramolecular electrostatic interactions and interactions between flexible side chains and the ligand are treated with a fixed-epsilon approximation (leading to a presently unavoidable imbalance in the treatment of these contributions). For rigid receptor models these contributions are obviously zero, respectively. Furthermore, FlexScreen uses a "docking center", around which the sampling is enhanced (biased moves), but this has no effect on the scoring function.

Docking simulations should ideally compute the free energy of binding, ΔG which comprises both enthalpic and entropic contributions stemming from conformational changes of the receptor and the ligand. Presently, methods for computation of absolute binding free energies are still very costly,[65] while methods computing relative binding free energies ΔG, rely on perturbation of one ligand. Neither of these approaches are presently useful to screen even medium sized libraries of structurally unrelated compounds. For this reason we continue to develop the docking methodology here, where 'energy" refers to the value of the scoring function, largely ignoring entropic contributions to the binding process. Improving upon this approximation, which presently all docking methods make, is not subject of this investigation, which focuses on conquering the sampling issues related to backbone flexibility.

The docking protocol uses a cascadic version of the stochastic tunneling algorithm (STUN),[66,67] which samples ligand and receptor degrees of freedom. The docking cascade consists of three different stages, starting with many short simulations in the first stage, from which the best nonoverlapping conformations are selected for further refinement in subsequent stages. In each stage an adjustable number of steps are performed for a population of ligand conformations. In this investigation, we used 5000/30,000/75,000 steps per ligand conformation in the first/second/third stage of the simulation. For rigid backbone simulations, we simulate 200 different ligand conformations in the first stage. The lowest four conformations with a distinct value of the scoring function are forwarded to the second stage. In the last stage, the best value of the scoring function conformation from stage two is finally relaxed (in the following referred to as a 200/4/1 sampling). The flexible backbone docking simulations used a 20/20/1 sampling in the different FlexScreen stages. To check for consistency all simulations are run repeatedly. Further information on the docking method and its applications of docking performance for several systems has been reported elsewhere.[53,66–68] For all parameters not explicitly stated, default values were used.

In each step FlexScreen changes a single degree of freedom for ligand and side chain and evaluates its influence on the scoring function. The move is accepted using the Metropolis criterion on the effective STUN energy. Ligand and side chain moves comprise translations of the ligand center of mass, random rotations of the ligand or intramolecular conformational changes for the ligand and dihedral angle changes of flexible side chains for the receptor. All of these relative moves are drawn from a Gaussian distribution. In this investigation we also use backbone reconstruction steps described below.

Treatment of backbone flexibility

In small-molecule docking backbone flexibility is, for the most part, confined to limited regions of the backbone, comprising of one or more sets of consecutive amino acids (loops). Because of enthalpic reasons, secondary structure elements are largely conserved between apo and holo structures, thus most changes occur in unstructured regions in the vicinity of the docking site. It is presently unrealistic to model protein flexibility for the whole protein in docking simulations. Ideally, the free energy changes of the protein have to be combined with the free energy changes of the ligand, neither of which are directly accessible in a docking simulation. To assess the quality of a particular binding pose, an accurate or at least compatible model free-energy of the free-energy microstate of the protein energy is required. Given the present state of force field accuracy (presently, all commonly used scoring functions in docking simulations incorporate only single protein structures and use a model to obtain a free energy estimate), it is necessary to confine protein conformational change to the most relevant degrees of freedom of the protein.

In the following, we pursue an approach, where backbone flexibility for predefined protein loop regions is sampled during the simulation using the backbone dihedral angles as the flexible degrees of freedom. In the following, we assume that the relevant flexible sections of the backbone are known before the simulation, either by comparison of crystal structures or by computational methods.[16]

Moves comprising single torsions of backbone dihedrals are not effective because the entire rest of the chain is moved that results in large scale conformational changes (and often clashing conformations). We have therefore implemented a loop reconstruction algorithm, where all dihedral angles in one of the predefined segments of the backbone are changed at once. The protein conformation outside the selected chain remains unchanged and the integrity of the backbone (angles, distances) is maintained. Such loop moves have been widely discussed previously and intensely studied: The first mathematical models of ring closure, when bond length and bond

distances are maintained fixed, were developed by Go and Scheraga[69] and applied to loops containing up to five residues.[70] More recently, different approaches for loop reconstruction have been developed, including the hierarchical loop prediction, in which the loops with the best scoring function generated from one stage are passed to the next where more focused (constrained) sampling is performed[71] or the robotics-inspired conformational sampling (kinematic closure).[72] In docking simulations, loop modeling was realized using a low-resolution protein representation to sample loop regions followed by a repacking of the protein side chains.[73] Loop prediction methods using the analytical generalized born plus nonpolar scoring function with torsion angle sampling have shown good accuracy in comparison with selected crystal structures.[74] Despite some progress most approaches still vary strongly in accuracy.[75,76]

We have implemented a scheme where backbone moves are attempted with an adjustable frequency during the simulations. Construction of a backbone move in this scheme consists of two steps after the backbone dihedral angles are randomized. First, using a global loop closure search the loop is closed (see section Docking Method). Second, the resulting conformation is minimized using a modified steepest descent algorithm (local optimization, section Treatment of Backbone Flexibility). If the loop closing routine was successful (gap is less than 0.001 nm), the new backbone conformation will conserved in the acceptance criterion of the Monte-Carlo search of the docking simulation. In the work reported here, the probabilities for a backbone reconstruction move are chosen to sample 1–2 new backbone conformations in average for one subset of the simulation (sample) in the first two stages. In the last stage, we modify the backbone conformation more often, but with smaller step size to achieve a good relaxation. Specifically, the probabilities were set to 25/10,000, 25/100,000, and 1/1000 in stages 1, 2 and 3, respectively. The backbone closure step is discussed in detail in the Supporting Information.

Receptor and ligand preparation

All receptor files were obtained from the RCSB Protein Data Bank in the PDB file format. As our software FlexScreen requires MOL2 data files as an input, we converted the receptor files with the software pdb2pqr[77] and Open Babel[78] into the MOL2 format. The software pdb2pqr provides the protonation state and the partial charges for the receptor file using the AMBER99[79] parameters. All water molecules, cofactors or ions present in the PDB files were removed. The ligand files were converted from the PDB file format into the MOL2 file format with MOE, using am1bcc[80] to determine protonation state and partial charges. Both the receptor structure and the ligand are relaxed prior to the simulation to define a reference for the energy.

Results

In the following, we present simulations for three different kinases, one of the most important families of drug targets, to demonstrate the functionality of our novel backbone algorithm. Kinases constitute about 30% of the targets in drug discovery projects today, making it the second most exploited protein family after G-protein-coupled receptors.[81] Drug development for kinases presents an enormous challenge, because the ATP binding site is highly conserved. Targeting this binding site tends to lead to unspecific binding, resulting in undesired side effects. In order to develop compounds that are active in only a single kinase or only a subfamily of kinases, specific features of that kinase/subfamily must be targeted.

Many kinases feature a flexible loop in the vicinity of the active site (Fig. 1) that partially encloses the ligand in the holo structure. This loop is either difficult to resolve in apo-conformation or may be present in a very different conformation. When a fixed apo structure is used for screening novel high affinity ligands may be blocked from docking into the fixed conformation by the loop. Although kinases are probably the protein family for which most crystal structures are available, the availability of an apo conformation is much more likely than a suitable holo conformation for some particular ligand. In high throughput applications with many different ligands it is very difficult to select an unbiased target structure. Docking into a rigid backbone conformation of a single crystal structures will exclude some ligands, which cannot be accommodated by the structure. Inspection of the large degree of backbone conformational change, in particular for the structures 2OJJ and 2OGJ (PDB-ID), suggests that it is a very difficult to generate

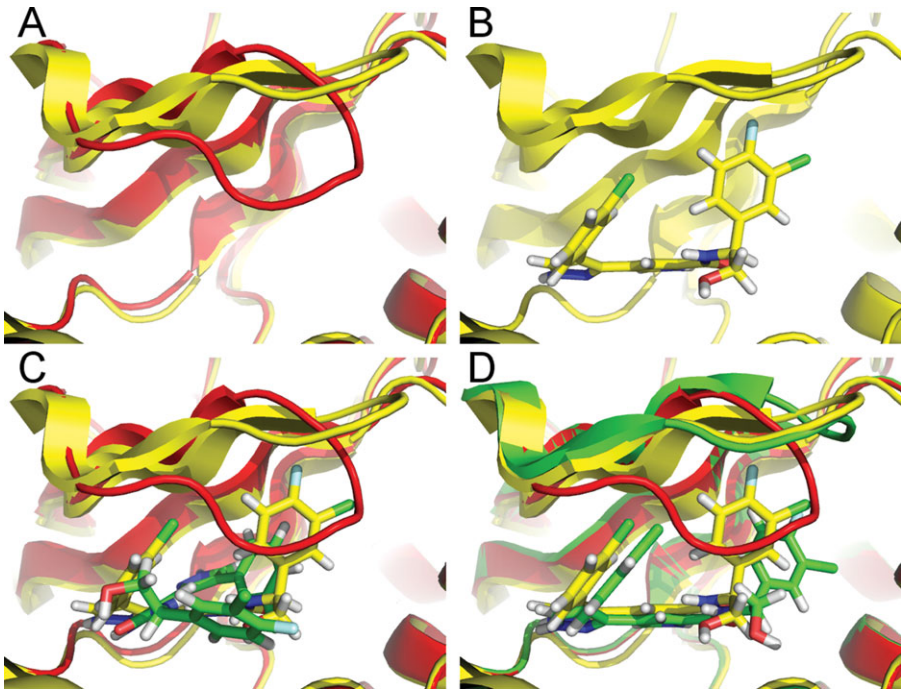

Figure 1. (A) Apo receptor (red) and holo ligand and receptor structures (yellow), (B) Native holo ligand and receptor structure, (C) Simulated ligand (green) without receptor backbone flexibility. (D) Simulated ligand and receptor (green) conformation with receptor backbone flexibility. Receptor simulation starting point was apo (red) structure. PDB IDs holo/apo structures: 2OJJ/2GPH.

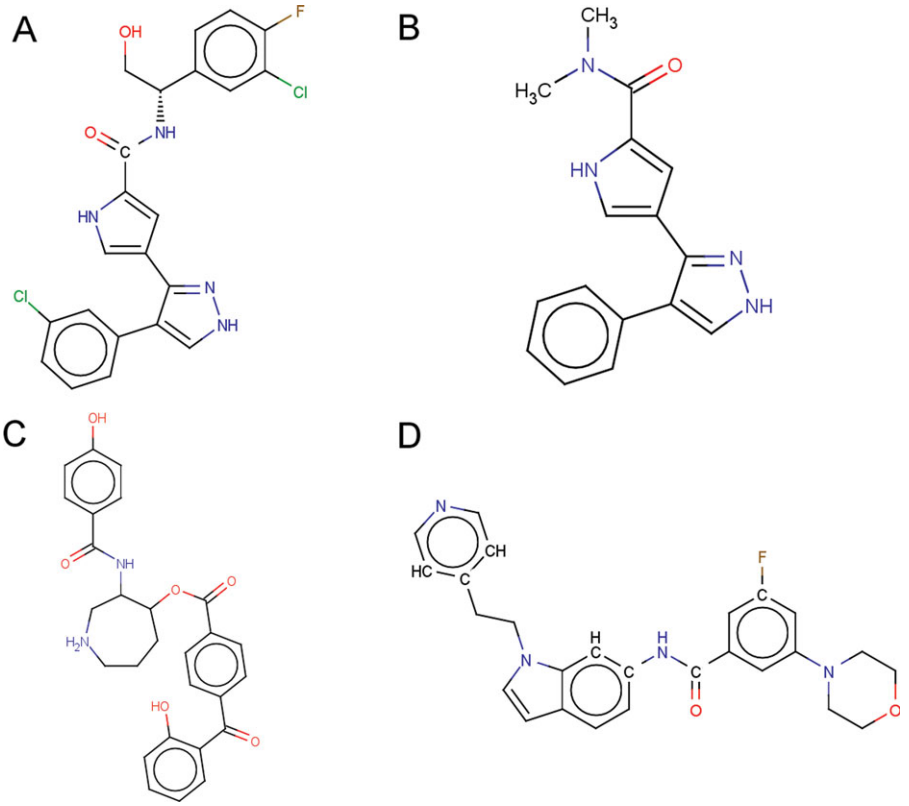

Figure 2. The four different ligands used in the docking simulation study: (A) (S) N (1 (3 chloro 4 fluoro phenyl) 2 hydroxyethyl) 4 (4 (3 chlorophenyl) 1H pyrazol 3 yl) 1H pyrrole 2 carboxamide, (B) N,N dimethyl 4 (4 phenyl 1H pyrazol 3 yl) 1H pyrrole 2 carboxamide, (C) 3 [(4 ydroxybenzoyl)amino]azepan 4 yl 4 (2 hydroxy benzoyl)benzoate, (D) N [1 (2 (4 pyridyl)ethyl) 6 indolyl] 3 fluoro 5 (4 morpholino)benzamide. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com].

the correct backbone conformation in the absence of the ligands, required for an ensemble docking approach. In Figure 1A, we show overlays of ERK2 receptor structures in apo and holo conformation. The latter is shown in Figure 1B with ligand ((S)-N-(1-(3-chloro-4-fluorophenyl)-2-hydroxyethyl)-4-(4-(3-chlorophenyl)-1H-pyrazol-3-yl)-1H-pyrrole-2-carboxamide) which illustrates this point.

In the current study, we docked different ligands into their corresponding apo receptor structures with and without our model for backbone flexibility and compared the results to the experimentally known holo structures. We will refer to these simulations as "apo-docking.' To assess the performance of our backbone reconstruction algorithm we always run our simulations twice ("rigid backbone docking simulation" and "flexible backbone docking simulation"). Rigid backbone docking simulations are performed with only side chain receptor flexibility,[16] the latter with full backbone flexibility in a selected protein region. In all of the following examples, we demonstrate that receptor backbone flexibility is required to predict the correct ligand or receptor position and binding scoring function.

The first and second examples focus on pyrazolylpyrrole analogs[82] (see Fig. 2 for an overview of all ligands) as inhibitors for ERK–kinases. ERK kinase is a widely expressed intracellular signaling molecule and is part of the Ras/Raf/MEK/ERK signal transduction pathway, which is responsible for many fundamental cellular processes, such as cell survival, proliferation, motility, and differentiation.[83–87] In many different cancer forms, i.e., lung, colon, pancreas, kidney, and ovary cancers the ERK pathway is disrupted, which leads to an enhancement of the deregulated molecular lesions in cancer. In the pathway ERK is a central point downstream of Ras, Raf and MEK, where multiple signaling pathways merge for transcription. There are two types of ERK kinases, which show a sequence identity of 88%, but for both the ATP binding pocket site is conserved to an even higher degree. Due to their crucial functions ERK kinases are intensely studied as a drug target and Aronov et al.[82] discovered new ERK inhibitors based on a structure-guided optimization of the pyrazolylpyrrol molecule. Holo crystal structures of the newly found inhibitors were deposited in the PDB database under accession numbers 2OJJ and 2OJG, which we used for our docking studies.

In the third example, we investigated binding of a balanol analogue ligand into a cyclic adenosine monophosphate-dependent protein kinase (PKA).[88] The Serine–Tyrosine kinase PKA was one of the first characterized protein kinases and is also one of the simplest and biochemically best understood.[89] Until the triggering through cyclic AMP, a messenger for hormone signaling, PKA is maintained in an inactive state. In the absence of cAMP, the enzyme complex contains two regulatory (R) and two catalytic (C) subunits. The activating signal (binding of cAMP to the R subunit) leads to dissociation of the complex into an R and two C subunits. This C subunit contains the same conserved catalytic core found in all protein kinases and is therefore often used as a model system for protein kinases. PKA has a functional role in the regulation of glycogen, sugar, and lipid metabolism.

Receptor plasticity in the ERK2 kinase

For the ERK kinases, we use the apo structure of an ERK2 kinase (PDB-ID: 2GPH), which has a resolution of 1.9 Å,[90] as the receptor starting point of our simulation. Figure 1A shows the differences between the apo (red) and the holo (yellow) structure after an alignment (structural superposition) indicating a significant flexibility in the glycine rich loop (Residue ILE29-VAL37), which severely constrains the ATP binding pocket. This receptor is a good example where comparison between apo and holo structures demonstrates a significant plasticity in the loop region. Obviously, the correct binding position will only be accessible with a flexible treatment of this loop region.

In preparing the docking simulations, we noted four additional side chains, which have a different conformation in the apo and holo structure, respectively. These residues (Residue LYS52, TYR62, ARG65, and GLN103) may also play an important role in the binding process of the ligand, because their side chains have either polar or charged side chains, meaning they are able to form hydrogen bonds with the ligand. In all simulations reported, these side chains were treated flexible. Ligand 1 (Fig. 2), (S)-N-(1-(3-chloro-4-fluorophenyl)-2-hydroxyethyl)-4-(4-(3-chlorophenyl)-1H-pyrazol-3-yl)-1H-pyrrole-2-carboxamide, was extracted from the holo crystal structure (PDB-ID: 2OJJ) and structurally randomized.

We then ran 200 independent simulations with a rigid backbone. The ligand conformation with the best score is shown in Figure 1C (green ligand), demonstrating that the ligand will bind in a position very far from the native pose (RMSD (root mean square deviation) >5 Å). Indeed, none of the 200 simulations managed to generate a final pose (independent of the value of the scoring function) with a significantly better RMSD.

In the second set of simulations, we enabled backbone flexibility for the nine consecutive amino acids, starting with the residue ILE29. Considering the rotational, spatial and internal flexibility of the ligand, and the receptor side chains, we have a system with a total of 61 degrees of freedom to minimize with respect to the scoring function for receptor and ligand.

We performed 2000 independent simulations with the flexible backbone protocol using the apo structure as the receptor starting point. Of the five ligand structures with the lowest values of the scoring function we obtained four conformations representing the same cluster of conformations, (RMSD variations under 0.1 Å). Because the scoring function of these conformations varies more than in the rigid docking simulations (Fig. 3), presumably because of the larger number of degrees

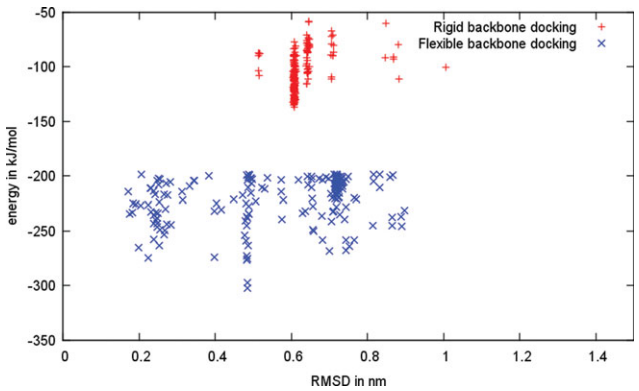

Figure 3. Scoring function vs. RMSD plot for rigid backbone docking simulations (red) and flexible backbone docking simulations without further relaxation (blue) of ligand 1.

of freedom, we performed two additional relaxation simulations starting with the conformation of the top scoring clusters. We used an iterative procedure to converge the scoring function: We reseeded the simulation with the previously obtained conformation and ran 200 independent simulations until the final scoring function of the simulation converged in comparison with the previous cycle. After three cycles both conformations had converged in scoring function. The best scoring structure is shown in Figure 1D (green ligand and receptor) and agrees well (after alignment of the receptor) with the native holo structure (final RMSD 2.32 Å). In comparison with the apo-structure, we found that the binding of the ligand forces the receptor to adopt a more open conformation, resulting in a significant movement of the glycine rich loop.

In the next example, we investigated a different ligand, ligand 2 (Fig. 2): N,N-dimethyl-4-(4-phenyl-1H-pyrazol-3-yl)-1H-pyrrole-2-carboxamide, lacking the second benzene ring, into the same apo receptor as used in example 1. The corresponding holo crystal structure has the PDB-ID 2OGJ (Fig. 4A). In comparison with the previous example, we find that the resulting loop movement upon ligand binding is larger for the ligand 1, which may be caused by the additional second benzene ring.

The rigid backbone simulation runs consisted of 200 runs with a stage composition of 200/4/1. The resulting RMSD value here was 2.44 Å, so the resulted ligand is not as away from the native position as in the previous example. For the flexible backbone simulation runs, we used the same receptor flexibility explained in example 1. The simulations consisted of 2000 independent runs using a 20/20/1 sampling scheme. The resulting ligand structure is shown in Figure 4B. In these simulations, the RMSD was reduced significantly to 1.74 Å. In comparison to the previous example there were also no competing structures with scoring function, so further refinement was not needed. It is interesting to analyze the difference between the two holo structures. Figure 4C illustrates the differences between the two holo structures of the two ligands for comparison. The glycine rich loop is more open for ligand 1 (blue structure), than for ligand 2 (yellow structure). To test the induced-fit theory, a docking simulation with ligand 2 into the best scoring receptor structure of the docking simulation from ligand 1 was performed with rigid backbone, but flexible side chain docking. The result is shown in Figure 4D, with an resulting RMSD of 2.80 Å, thus very similar to the rigid backbone docking simulation, which started from the apo structure, demonstrating that also crossdocking can only be accomplished with consideration of backbone flexibility.

Receptor plasticity in the PKA kinase

As the third example, we investigated the binding of 3-[(4-hydroxybenzoyl)amino]azepan-4-yl 4-(2-hydroxybenzoyl)benzoate (ligand 3, Fig. 2C), a fungal metabolite, into a cyclic adenosine monophosphate-dependent protein kinase (PKA). For the receptor, we used the available PKA apo structures resolved in 1992 to a resolution of 2.5Å (PDB-ID 2CPK[89]). The ligand was later modified to yield potent inhibitors with sufficient selectivity as is desirable for pharmaceutical use. Ligand modifications cause plasticity of the glycine rich loop (Residue LEU49-VAL57) in the ATP binding pocket of the PKA kinase.[88] This loop stabilizes the binding position of the ligands, but is also very flexible. At first sight it appears that holo and apo structures differ only very little. The differences between the

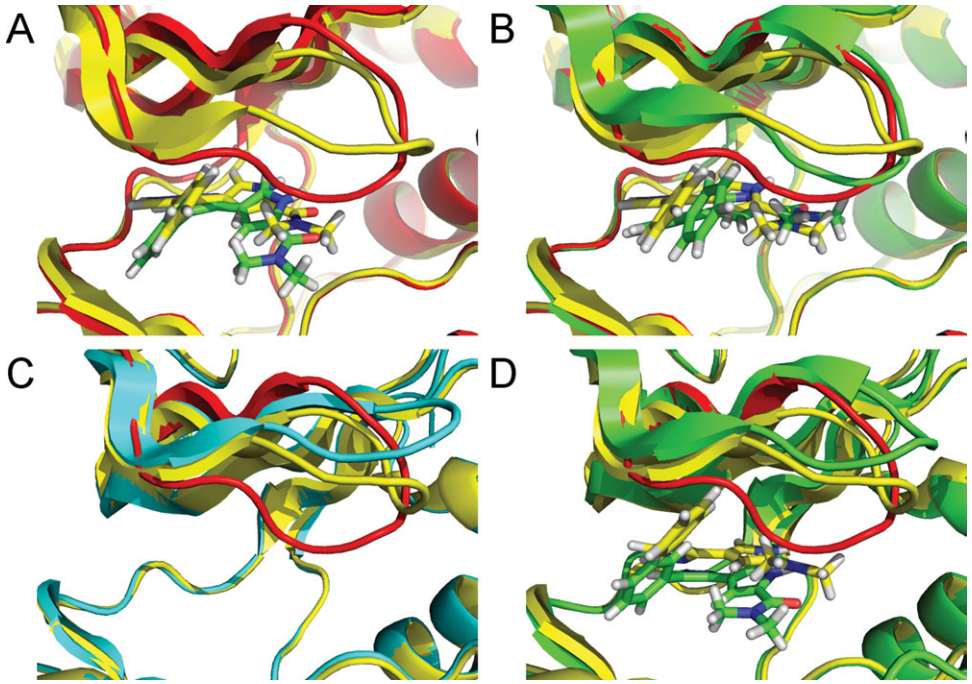

Figure 4. Apo receptor (red) and holo ligand and receptor structures (yellow) as compared to simulated ligand and receptor conformations (green). Receptor simulation starting point was apo (red) structure. (A) Simulated ligand (green) without receptor backbone flexibility. (B) Simulated ligand and receptor (green) conformation with receptor backbone flexibility. PDB IDs holo/apo structures: 2OGJ/2GPH (C) Holo structures of ligand 1 (blue) and ligand 2 (yellow), (D) Simulated ligand 2 (green) docked into best scoring receptor structure of ligand 1. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com].

apo receptor and the holo receptor with ligand are shown in Figure 5A. Further, we noted several side chains LYS72, LYS78, LEU82, GLU91, MET120, GLU127, LYS168, ASN171, THR183, and ASP184 in the two structures which may play a role in the binding process and were therefore made flexible in the docking simulations.

To compare the performance of our new backbone algorithm in the docking simulations we made two sets of simulations: In the rigid backbone simulations we considered 21 flexible side chains but without backbone flexibility. We used the same protocols in both simulations, i.e., 2000 parallel runs for the flexible backbone simulations. The result of the rigid backbone simulations is shown in Figure 5A. Without the flexibility of the glycine rich loop it is not possible to reproduce the correct position of the ligand. The RMSD value is very large (>9Å) and the overall orientation of the ligand is inverted along the vertical plane perpendicular to the figure. The ligand assumes a compact configuration (native: extended), resulting in a lack of contacts between the ligand with the hinge region (Residue GLU121-VAL123) of the binding pocket.

Next, we performed flexible backbone simulations with the very same side chain flexibility, but including the backbone flexibility. In total, we found four different clusters of ligand conformation (Fig. 6). The cluster with the best scoring function corresponds to the native pose of the ligand in the holo crystal structure and includes the conformation with the best scoring function. The improvement under the consideration of the flexibility of the glycine rich loop is shown in Figure 5B. The RMSD value of the docked pose is very small (0.6Å), meaning the native configuration was reproduced very accurately. The cluster with the highest RMSD (>9Å) has a lower scoring function, but corresponds to the apo form of the glycine rich loop and is also the result of the rigid docking. Because the scoring function of this conformation is only 18.8 kJ/mol lower than the top scoring state (in the uncalibrated scale of the scoring function this reflects a small difference among populations), this conformation may be an intermediate state in the binding process. We also performed an additional 7000 independent flexible backbone simulations to check for convergence but found no notable difference in the results.

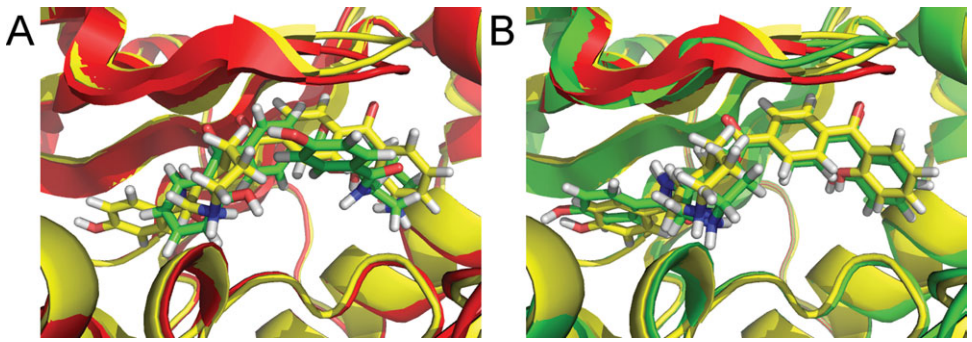

Figure 5. Apo receptor (red) and holo ligand and receptor structures (yellow) as compared to simulated ligand and receptor conformations (green). Receptor simulation starting point was apo (red) structure. Left, simulated ligand (green) without receptor backbone flexibility. Right, simulated ligand and receptor (green) conformation with receptor backbone flexibility. PDB IDs holo/apo structures: 1RE8/2CPK. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com].

Allosteric binding in the MAP kinase p38α

One very challenging approach to develop drugs specific for one kinase or a subfamily of kinases is to exploit allosteric binding effects, where the ligands does not bind in the active site, but a different location. Binding to the allosteric

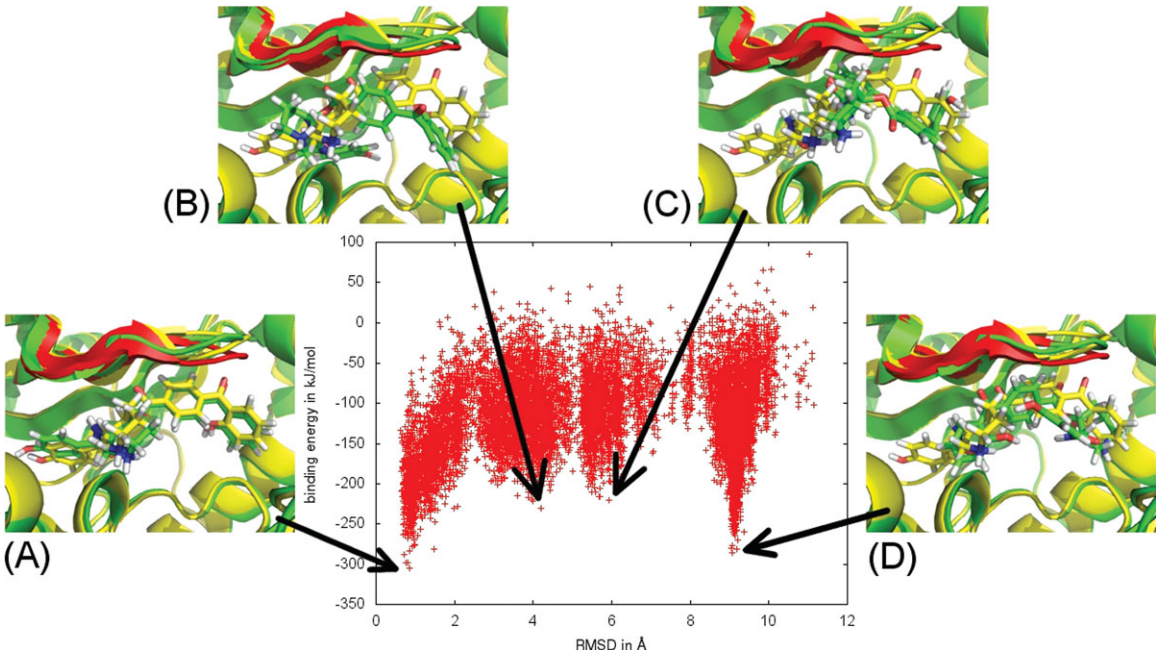

Figure 6. The lowest scoring function conformation with simulated receptor and ligand conformation corresponding to the experimental holo structure (A), an intermediate state (D) with simulated receptor conformation corresponding to experimental apo structure and two local minima conformations (B),(C) drawn from 14280 independent docking simulations, starting from apo structure (red). For comparison the crystal holo structure is colored yellow and simulated ligand and receptor conformation are colored in green (PDB IDs of holo/apo structure: 1RE8/2CPK). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com].

site, for example illustrated in Figure 7 induces a conformational change that affects the active site. As allosteric mechanisms often involve backbone conformational changes, treatment of protein conformational flexibility is important to model allosteric binding.

One interesting examples is the mitogen-activated protein kinase (MAP) p38α kinase is an intracellular serine/threonine kinase and can be activated by many environmental stimuli such as TNF-α, IL-1β or stress caused by osmotic shock or UV light.[91] The activation of p38α is caused by phosphorylation on the Thr180-Gly181-Tyr182 activation loop.[92] In its active state p38α phosphorylates many intracellular protein substrates which regulate the biosynthesis of TNF-α and IL-1β. Furthermore, the MAP kinase plays an important role in the regulation of the COX-2 gene expression.[93] It is understood that the excessive production of TNF-α and IL-1β may lead to many inflammatory diseases such as rheumatoid arthritis, Crohn's disease, inflammatory bowel disease and psoriasis.[94]

It has been shown that inhibition of p38α using SB203580 significantly reduces TNF-α and IL-1β production, which makes p38α a very promising target for novel anti-inflammatory therapy.[62] Until now, there are two anti-inflammatory agents on the market, which specifically inhibit the TNF-α production: Enbrel[95] (a TNF-α receptor) and Remicade[96] (a human TNF monoclonal antibody). The development of orally bioavailable small molecule inhibitors promises new ways in therapies. Indeed a number of promising p38α small molecule inhibitors have entered already human clinical trials over the last years.[96]

In the following, we discuss one specific allosteric binding mechanism for p38α kinases, which is directly connected to the DFG-motif[97] (see Fig. 7). The DFG-motif, named after the one-letter code of the constituent amino acids, is part of the activation loop, where a large conformational change of the ASP-PHE-GLY residues leads to the formation of a largely lipophilic pocket, into which an inhibitor may insert. Different conformations for this loop have been found in experiments, which are classified into two categories, the "DFG-In" or "DFG-Out" mode, respectively. This categorization refers to the position of the ASP residue in the motif: In a "DFG-In' conformation the ASP residue points into the ATP binding pocket, where it coordinates the magnesium ion, necessary for the dephosphorylation of ATP to ADP.[98] In the "DFG-Out" conformation the ASP residue points away from the ATP binding site, which leads to an inactive conformation of the kinase, because the residues in the active site are not oriented to allow the catalyzation of the phosphotransfer. The conformation of the activation loop also blocks the ATP binding in the binding pocket.[99] This is easily seen in the volume representation of ATP in both conformations, which leads to significant clashes for the DFG-out conformation (Figs. 7C and 7D). The DFG-out conformations have been observed crystallographically for Abl, p38α, b-Raf, EGFR, Kdr, c-Kit, and Aurora A kinase and it still remains unclear, why not all kinases can adopt a DFG-out flip.[99]

For our study, we used the crystal structure with PDB-ID 1R3C as the apo p38α kinase receptor reference structure with a "DFG-In" and the structure with accession number 1WBT[94] as a reference DFG-out structure. The ligand in this structure is

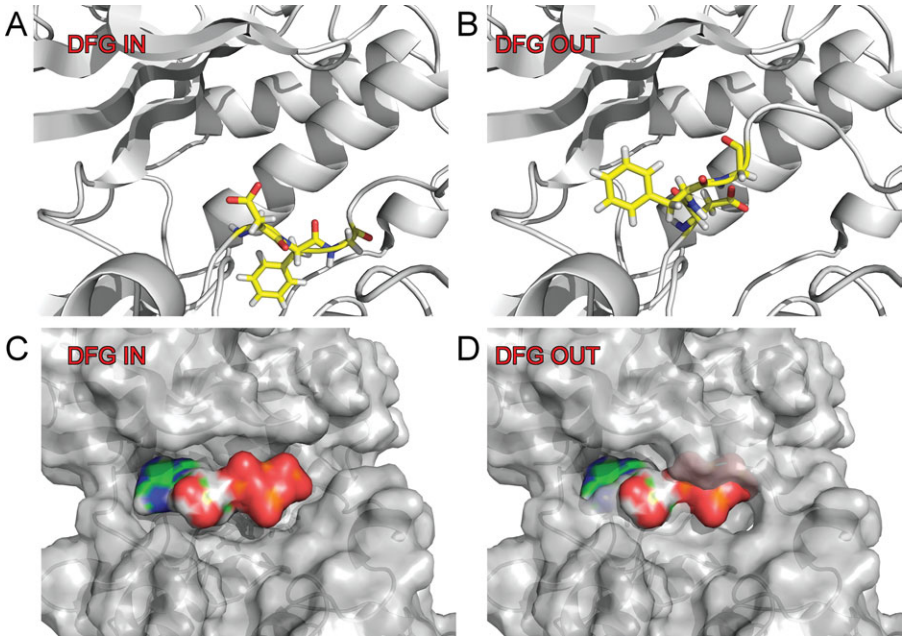

Figure 7. DFG residues in (A) 'In' and (B) 'Out' conformation (yellow). Surface representation of aligned ATP in the receptor binding pocket in a (C) 'DFG In' and (D) 'DFG Out' conformation. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com].

a modified benzamide molecule, which was found in a high throughput X-ray crystallographic screening. We ran 2000 independent simulations with backbone sampling using a 20/20/1 sampling scheme. The backbone and side chains in the region ILE166-ASP176 were treated flexibly. In addition, side chains which may influence the ligand binding process (TYR35, ARG67, GLU71, LEU74, LEU75, MET78, ILE84, THR106, ILE141, GLU178) were treated flexible, leaning to a total of 21 flexible side chains.

Starting from a DFG-in conformation both modes of the DFG-motif were extensively sampled, but the lowest scoring function was found for conformations with the DFG-in conformation, which corresponds to the correct apo conformation of the complex. Here, the ligand has a high RMSD with 10.8Å (Fig. 8B). One conformation with a very low RMSD value (Fig. 8A) did also occur in the simulation run. The ligand RMSD here is 1.9Å, but the total binding scoring function was more than 51.30 kJ/mol lower than the scoring function in the conformation with the highest scoring function. Interestingly, we found in low ligand RMSD value conformations stronger direct interactions between ligand and receptor than in high RMSD conformations (difference of 34.09 kJ/mol). The main differences in scoring function are caused by the internal receptor contribution to the scoring function. In conformations close to the apo structure the internal receptor scoring function is 84.79 kJ/mol higher than in conformations similar to the holo structure, which can be traced to changes in the electrostatic interaction of side chains that are not resolved in the crystal structure and hence difficult to analyze. Nevertheless, our simulation protocol was able to reproduce, without prior knowledge of the apo conformation, the complex conformational changes associated with the reorientation of the DFG-motif that is responsible for the allosteric control of the p38α kinase.

Discussion

Backbone receptor flexibility is known to have a significant impact on the ability to correctly describe and rank the binding poses of different ligands in many pharmaceutically relevant receptors. In this study, we have succeeded to implement an efficient backbone reconstruction algorithm that permits to sample backbone reconstruction events in the framework of receptor-ligand docking simulations. By this increases the numerical effort, as discussed below, but we could also demonstrate that state-of-the-art computational hardware permits

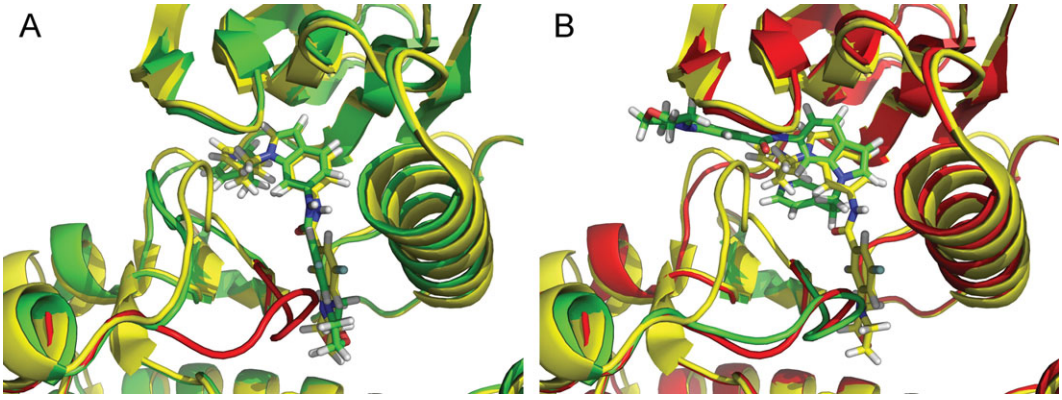

Figure 8. Apo receptor (red) and holo ligand and receptor structures (yellow) as compared to simulated ligand and receptor conformations (green). Low RMSD simulated ligand with a holo like conformation (Left). Best scoring simulation result with a apo like receptor conformation and a high RMSD ligand (Right). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com].

thousands of such calculations in a reasonable timeframe. In the following, we discuss a few specific points that are relevant when considering backbone-flexibility:

Balancing backbone-backbone and backbone-ligand energies

Balancing backbone-backbone and backbone-ligand energies emerges as an obvious novel challenge in flexible-backbone simulations and scoring functions must be adapted, possibly by learning from molecular-dynamics force field simulations, to better address these contributions. In our simulations, we often found large energetic differences for relatively small backbone movements, which may be caused by an overestimation of the internal receptor contribution in the scoring function. This issue was most pertinent at the p38α kinase, where the receptor internal scoring function for receptor conformations similar to the apo conformation was much better than for receptor conformations similar to the holo structure. However, for this particular case the change in the receptor scoring function could be traced to the conformation of a side chain (ARG173) that is not resolved in the apo crystal structure. An interaction of this highly charged group with the solvent, which are not accounted for in the scoring function, may partially compensate this effect. A more accurate solvent treatment, for example using the Generalized Born model,[100] may help ameliorate this problem. In our simulations, we always generate a reference conformation for receptor and ligand apart from each other. This avoids problems with high-energies when starting simulations with a specific experimental structure and is always required in cases where the experimental structure is not fully resolved (missing loops/residues). This also permits comparison between binding energies of a single ligand to different kinases in principle, although given the resolution of presently available scoring functions, such comparisons are meaningful only for closely related receptors. It is possible to postprocess the results to get more accurate absolute values of binding energy with other methods. Common methods for this are the linear interaction energy method[101] or the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM-PBSA) and the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA).[102,103]

Selection of the receptor flexibility

Even using our efficient sampling technique, only a fraction of the protein side chain and backbone can be treated flexibly. Automatization of flexible side chain selection has been realized by using Lennard-Jones clashes in the scoring function of rigid docking as well as other methods.[16,40] This approach showed reasonable results in balancing docking performance and computational time. The increase of the computational cost by making certain backbone regions flexible is much steeper than for enhanced side chain sampling. We therefore presently advocate manual selection of flexible backbone regions by comparison of apo and holo structures or by the identification of unstructured regions of the receptor in the vicinity of the binding pocket. Some receptor classes, such as kinases, show a very high similarity in the catalytic region[104] and the detail models for the selection of flexible backbone regions are available. Recent studies[51,98] have shown that the DFG-motif in all p38 and also in ABL kinases is very flexible and this flexibility therefore should be included in docking simulations, if ligand exploration in the lipophilic pocket due to DFG-motif movement, is desired.

Computational costs

In the current study, the required computational time depended strongly on the total number of degrees of freedom. Docking a single ligand with only receptor side chain flexibility required from 3–5 min in a single simulation run, while a single simulation including receptor backbone flexibility for 8–11 residues required 6–8 h of computational time (IntelPC-86-64, 1.8 GHz processor). As a larger degree of backbone flexibility leads to a larger conformational space, more sampling is necessary to find local minima in the scoring function. One important advantage is enhancement of reliability by running multiple of simulations parallel. For some receptors up to 12,000 independent parallel simulations were performed (Fig. 6) to sample accurately the conformational space. This investigation demonstrates that an increasing number of simulations has little influence on the selected conformations of both backbone and ligand (compared to only 2000 independent parallel runs), but a much strong influence on the final scoring function.

Interaction analysis

It is known that all of presently available scoring functions correlate only weakly with the measured affinities.[7,105] Many enthalpic contributions to receptor ligand interactions are extremely short range, e.g., hydrogen bonding and pi-pi stacking interactions have a steep distance and angle dependency. In our simulations, we observed that for several cases the additional interaction points, such as hydrogen bonds, were found after backbone relaxation. In particular, for incomplete crystal structures consideration of backbone flexibility may thus be important to fully resolve all interactions between receptor and ligand and to rank the ligands in a database correctly.

Conclusions

It is well-known that the receptor flexibility plays an important role in ligand recognition by pharmaceutically relevant protein receptors. While there has been significant progress in sampling side chain degrees of freedom during receptor-ligand docking simulations, there has been little progress in treatment of the backbone receptor flexibility due to the large associated computational costs and difficulties to balance inter molecular and intra molecular contributions to the scoring function. Ensemble docking methods have been proposed to ameliorate this problem, but these methods have a difficulty to account for the contribution of receptor-reorganization energy to the affinity and in the selection of adequate target structures, when only apo structures with unstructured or structurally unresolved loops are available.

Here, we pioneered an approach that permits sampling of large preselected loop regions in the receptor during the docking simulation with an adequate accounting of receptor reorganization energy in the scoring function. We demonstrate in docking simulations starting from the apo structure for three different kinase examples good agreement between the observed binding poses and interactions in comparison with experimentally known holo structures. Generally, we find the calculated scoring function in the relaxed backbone conformation to be significantly better than in the control simulations with rigid backbone. Using the flip of the DFG-motif of p38α kinase as an example, we also demonstrate that this approach can be used to model allosteric binding effects that can be exploited to increase ligand specificity in families of proteins with structurally conserved binding pockets, such as kinases. This approach increases the computational cost by approximate an order of magnitude in comparison to rigid-backbone simulations, but still remains feasible using of-the-shelf hardware for many interesting applications.

In our view, the most significant remaining challenge is the improvement of scoring functions, particularly with respect to the balance of intermolecular and intramolecular interactions, which remains a sore point for almost all in silico screening approaches. However, we have found improvement from an unexpected source by reevaluating the interactions in relaxed holo structures and found that the scoring functions detect short-range interactions in the relaxed receptor conformations that are not detected in the experimental holo structures. This obviously does not mean that these interactions are not present in the experimental structures, but simply reflects the fact that presently available scoring functions are not able to adequately account for these interactions using just the experimental backbone conformation. As a result some of the deficiencies in affinity prediction may be reduced by using backbone reconstruction in scoring the final poses.

Our results suggest that performance of docking simulations can be significantly improved by including receptor flexibility in continuous backbone parts into the docking simulations. A main advance of this new flexible receptor docking approach is its capability to use apo structures as input for the docking simulations. With these demonstrated applications we have shown that docking simulations incorporating backbone flexibility can improve the results of in silico screens and may help to discover new-scaffold ligands that remain presently undetected in rigid-backbone screening protocols.

Keywords: in silico screening · induced fit · receptor flexibility · allosteric effect · kinase inhibitor

[1] B. Hughes, *Nat. Rev. Drug Discov.* **2010**, *9,* 89.
[2] A. L. Hopkins, C. R. Groom, *Nat. Rev.* **2002**, *1,* 727.
[3] B. M.Bolten, T. DeGregorio, *Nat. Rev.* **2002**, *1,* 335.
[4] S. Ghosh, A. Nie, J. An, Z. Huang, *Curr. Opin. Chem. Biol.* **2006**, *10,* 194.
[5] D. Plewczynski, M. Łaźniewski, R. Augustyniak, K. Ginalski, *J. Comput. Chem.* **2011**, *32,* 742.
[6] G. Klebe, *Drug Discov. Today* **2006**, *11,* 580.
[7] G. L. Warren, C. W. Andrews, A. M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, S. Senger, G. Tedesco, I. D. Wall, J. M. Woolven, C. E. Peishoff, M. S. Head, *J. Med. Chem.* **2006**, *49,* 5912.
[8] C. Sotriffer, H. Matter, In Virtual Screening: Principles, Challenges, and Practical Guidelines; C. Sotriffer, Ed.; Wiley VCH: Weinheim, **2011**; pp 177 221.
[9] B. Lin, K. Y. Wong, C. Hu, H. Kokubo, B. M. Pettitt, *J. Phys. Chem. Lett.* **2011**, *2,* 1626.
[10] S. J. Teague, *Nat. Rev. Drug Discov.* **2003**, *2,* 527.
[11] T. Lazaridis, A. Masunov, F. Gandolfo, *Protein. Struc. Func. Genet.* **2002**, *47,* 194.
[12] V. Mohan, A. C. Gibbs, M. D. Cummings, E. P. Jaeger, R. L. DesJarlais, *Curr. Pharm. Des.* **2005**, *11,* 323.
[13] O. Rabal, G. Schneider, J. I. Borrell, J. Teixido, *BioDrugs* **2007**, *21,* 31.
[14] C. W. Murray, C. A. Baxter, A. D. Frenkel, *J. Comput. Aided Mol. Des.* **1999**, *13,* 547.
[15] D. L. Mobley, A. P. Graves, J. D. Chodera, A. C. McReynolds, B. K. Shoichet, K. A. Dill, *J. Mol. Biol.* **2007**, *371,* 1118.
[16] D. B. Kokh, W. Wenzel, *J. Med. Chem.* **2008**, *51,* 5919
[17] V. Kairys, M. K. Gilson, *J. Comput. Chem.* **2002**, *23,* 1656.
[18] G. Bottegoni, I. Kufareva, M. Totrov, R. Abagyan, *J. Comput. Aided Mol. Des.* **2008**, *22,* 311.
[19] H. A. Carlson, J. A. McCammon, *Mol. Pharmacol.* **2000**, *57,* 213.
[20] A. May, M. Zacharias, *J. Med. Chem.* **2008**, *51,* 3499.
[21] S. Leis, M. Zacharias, *J. Comput. Chem.* **2011**, *32,* 3433.
[22] M. Rueda, G. Bottegoni, R. Abagyan, *J. Chem. Inf. Model.* **2009**, *49,* 716.
[23] Y. Zhao, M. F. Sanner, *Proteins* **2007**, *68,* 726.
[24] C. N. Cavasotto, R. A. Abagyan, *J. Mol. Biol.* **2004**, *337,* 209.
[25] S. Chaudhury, J. J. Gray, *J. Mol. Biol.* **2008**, *381,* 1068.
[26] R. S. Armen, J. Chen, C. L. Brooks, *J. Chem. Theory Comput.* **2009**, *5,* 2909.
[27] J. Subramanian, S. Sharma, C. B Rao, *ChemMedChem* **2008**, *3,* 336.
[28] J. Chatterjee, O. Ovadia, G. Zahn, L. Marinelli, A. Hoffman, C. Gilon, H. Kessler, *J. Med. Chem.* **2007**, *50,* 5878.
[29] T. Polgár, G. M. Keserü, *J. Chem. Inf. Model.* **2006**, *46,* 1795.
[30] M. L. Teodoro, L. E. Kavraki, *Curr. Pharm. Des.* **2003**, *9,* 1635.
[31] I. Antes, *Protein.: Struct. Funct. Bioinformatics* **2010**, *78,* 1084.
[32] J. Meiler, D. Baker, *Proteins* **2006**, *65,* 538.
[33] M. Totrov, R. Abagyan, *Proteins* **1997**, (Suppl 1):215.
[34] I. W. Davis, D. Baker, *J. Mol. Biol.* **2009**, *385,* 381.
[35] D. Seeliger, B. L. de Groot, *PLoS Comput. Biol.* **2010**, *6,* e1000634.
[36] G. M. Arantes, *Proteins* **2010**, *78,* 3017.
[37] M. S. Zoda, M. Zacharias, S. Reissmann, *J. Pept. Sci.* **2010**, *16,* 403.
[38] D. F. Cummings, S. S. Ericksen, J. A. Schetz, *J. Neurochem.* **2009**, *110,* 45.
[39] M. A. Lill, *Biochemistry* **2011**, *50,* 6157.
[40] D. B. Kokh, R. C. Wade, W. Wenzel, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1,* 298.
[41] P. Csermely, R. Palotai, R. Nussinov, *Trends Biochem. Sci.* **2010**, *35,* 539.
[42] L. C. James, D. S. Tawfik, *Proc. Natl. Acad. Sci. USA* **2005**, *102,* 12730.
[43] D. D. Boehr, D. McElheny, H. J. Dyson, P. E. Wright, *Proc. Natl. Acad. Sci. USA* **2010**, *107,* 1373.
[44] T. R. Weikl, C. von Deuster, *Proteins* **2009**, *75,* 104.
[45] G. M. Verkhivker, D. Bouzida, D. K. Gehlhaar, P. A. Rejto, S. T. Freer, P. W. Rose, *Curr. Opin. Struct. Biol.* **2002**, *12,* 197.
[46] G. G. Hammes, Y. C Chang, T. G. Oas, *Proc. Natl. Acad. Sci. USA* **2009**, *106,* 13737.
[47] R. B. Best, *Proc. Natl. Acad. Sci.* **2006**, *103,* 10901.
[48] D. D. Boehr, R. Nussinov, P. E. Wright, *Nat. Chem. Biol.* **2009**, *5,* 789.
[49] K. A. Henzler Wildman, V. Thai, M. Lei, M. Ott, M. Wolf Watz, T. Fenn, E. Pozharski, M. A. Wilson, G. A. Petsko, M. Karplus, C. G. Hübner, D. Kern, *Nature* **2007**, *450,* 838.
[50] F. Schotte, M. Lim, T. A. Jackson, A. V. Smirnov, J. Soman, J. S. Olson, G. N. Phillips, Jr., M. Wulff, P. A. Anfinrud, *Science* **2003**, *300,* 1944.

[51] M. Vogtherr, K. Saxena, S. Hoelder, S. Grimme, M. Betz, U. Schieborr, B. Pescatore, M. Robin, L. Delarbre, T. Langer, K. U. Wendt, H. Schwalbe, *Angew. Chem.* **2006**, *118,* 1008.
[52] D. E. Koshland, *Proc. Natl. Acad. Sci.* **1958**, *44,* 98.
[53] B. Fischer, H. Merlitz, W. Wenzel, *Comput. Life Sci. Proc.* **2005**, *3695,* 186.
[54] A. C. Anderson, R. H. O'Neil, T. S. Surti, R. M. Stroud, *Chem. Biol.* **2001**, *8,* 445.
[55] A. R. Leach, *J. Mol. Biol.* **1994**, *235,* 345.
[56] T. M. Frimurer, G. H. Peters, L. F. Iversen, H. S. Andersen, N. P. H. Møl ler, O. H. Olsen, *Biophys. J.* **2003**, *84,* 2273.
[57] R. Najmanovich, J. Kuttner, V. Sobolev, M. Edelman, *Proteins* **2000**, *39,* 261.
[58] M. A. Lill, A. Vedani, M. Dobler, *J. Med. Chem.* **2004**, *47,* 6174.
[59] S. B. Nabuurs, M. Wagener, J. de Vlieg, *J. Med. Chem.* **2007**, *50,* 6507.
[60] F. Glaser, Y. Rosenberg, A. Kessel, T. Pupko, N. Ben Tal, *Protein.: Struct. Funct. Bioinformatics* **2004**, *58,* 610.
[61] K. Gunasekaran, R. Nussinov, *J. Mol. Biol.* **2007**, *365,* 257.
[62] T. Gallagher, *Bioorg. Med. Chem. Lett.* **1995**, *5,* 1171.
[63] G. M. Morris, D. S. Goodsell, R. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19,* 1639.
[64] N. A. Baker, *Proc. Natl. Acad. Sci.* **2001**, *98,* 10037.
[65] T. Steinbrecher, A. Labahn, *Curr. Med. Chem.* **2010**, *17,* 767.
[66] H. Merlitz, W. Wenzel, *Chem. Phys. Lett.* **2002**, *362,* 271.
[67] W. Wenzel, K. Hamacher, *Phys. Rev. Lett.* **1999**, *82,* 3003.
[68] B. Fischer, S. Basili, H. Merlitz, W. Wenzel, *Protein.:Struct. Funct. Bioin formatics* **2007**, 195.
[69] N. Go, H. A. Scheraga, *Macromolecules* **1970**, *3,* 178.
[70] K. A. Palmer, H. A. Scheraga, *J. Comput. Chem.* **1992**, *13,* 329.
[71] B. D. Sellers, K. Zhu, S. Zhao, R. A. Friesner, M. P. Jacobson, *Protein.: Struct. Funct. Bioinformatics* **2008**, *72,* 959.
[72] D. J. Mandell, E. A. Coutsias, T. Kortemme, *Nat. Methods* **2009**, *6,* 551.
[73] C. Wang, P. Bradley, D. Baker, *J. Mol. Biol.* **2007**, *373,* 503.
[74] A. K. Felts, E. Gallicchio, D. Chekmarev, K. A. Paris, R. A. Friesner, R. M. Levy, *J. Chem. Theory Comput.* **2008**, *4,* 855.
[75] D. Baker, A. Sali, *Science* **2001**, *294,* 93 96.
[76] C. B Rao, J. Subramanian, S. D. Sharma, *Drug Discov. Today* **2009**, *14,* 394.
[77] T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe, N. A. Baker, *Nucleic Acids Res.* **2007**, *35,* W522.
[78] R. Guha, M. T. Howard, G. R. Hutchison, P. Murray Rust, H. Rzepa, C. Steinbeck, J. Wegner, E. L. Willighagen, *J. Chem. Inf. Model.* **2006**, *46,* 991.
[79] J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, *21,* 1049.
[80] A. Jakalian, D. B. Jack, C. I. Bayly, *J. Comput. Chem.* **2002**, *23,* 1623.
[81] R. Morphy, *J. Med. Chem.* **2010**, *53,* 1413.
[82] A. M. Aronov, C. Baker, G. W. Bemis, J. Cao, G. Chen, P. J. Ford, U. A. Germann, J. Green, M. R. Hale, M. Jacobs, J. W. Janetka, F. Maltais, G. Martinez Botella, M. N. Namchuk, J. Straub, Q. Tang, X. Xie, *J. Med. Chem.* **2007**, *50,* 1280.
[83] J. S. Sebolt Leopold, R. Herrera, *Nat. Rev. Cancer* **2004**, *4,* 937.
[84] K. S. M. Smalley, *Int. J. Cancer* **2003**, *104,* 527.
[85] M. Kohno, J. Pouyssegur, *Prog. Cell Cycle Res.* **2003**, *5,* 219.
[86] F. Chang, L. S. Steelman, J. G. Shelton, J. T. Lee, P. M. Navolanic, W. L. Blalock, R. Franklin, J. A. McCubrey, *Int. J. Oncol.* **2003**, *22,* 469.
[87] R. A. Hilger, M. E. Scheulen, D. Strumberg, *Onkologie* **2002**, *25,* 511.
[88] P. Akamine, Madhusudan; L. L. Brunton, H. D. Ou, J. M. Canaves, N. h. Xuong, S. S. Taylor, *Biochemistry* **2004**, *43,* 85.
[89] D. R. Knighton, J. H. Zheng, L. F. Ten Eyck, V. A. Ashford, N. H. Xuong, S. S. Taylor, J. M. Sowadski, *Science* **1991**, *253,* 407.
[90] Y. Zhou, H. Zhou, C. Zhang, S. Liu, *Cell Biochem. Biophys.* **2006**, *46,* 165.
[91] J. Han, J. D. Lee, L. Bibbs, R. J. Ulevitch, *Science* **1994**, *265,* 808.
[92] J. Raingeaud, A. J. Whitmarsh, T. Barrett, B. Dérijard, R. J. Davis, *Mol. Cell. Biol.* **1996**, *16,* 1247.
[93] J. L. Adams, A. M. Badger, S. Kumar, J. C. Lee, *Prog. Med. Chem.* **2001**, *38,* 1.
[94] A. L. Gill, M. Frederickson, A. Cleasby, S. J. Woodhead, M. G. Carr, A. J. Woodhead, M. T. Walker, M. S. Congreve, L. A. Devine, D. Tisi, M. O'Reilly, L. C. A. Seavers, D. J. Davis, J. Curry, R. Anthony, A. Padova, C. W. Murray, R. A. E. Carr, H. Jhoti, *J. Med. Chem.* **2005**, *48,* 414.
[95] B. Jarvis, D. Faulds, *Drugs* **1999**, *57,* 945.
[96] N. L. Carteron, *Mol. Med. Today* **2000**, *6,* 315.
[97] Y. Shan, M. A. Seeliger, M. P. Eastwood, F. Frank, H. Xu, M. O. Jensen, R. O. Dror, J. Kuriyan, D. E. Shaw, *Proc. Natl. Acad. Sci.* **2008**, *106,* 139.
[98] N. M. Levinson, O. Kuchment, K. Shen, M. A. Young, M. Koldobskiy, M. Karplus, P. A. Cole, J. Kuriyan, *PLoS Biol.* **2006**, *4,* e144.
[99] B. Okram, A. Nagle, F. J. Adrián, C. Lee, P. Ren, X. Wang, T. Sim, Y. Xie, X. Wang, G. Xia, G. Spraggon, M. Warmuth, Y. Liu, N. S. Gray, *Chem. Biol.* **2006**, *13,* 779.
[100] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112,* 6127.
[101] T. Hansson, J. Marelius, J. Aqvist, *J. Comput. Aided Mol. Des.* **1998**, *12,* 27.
[102] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, T. E. Cheatham, *Acc. Chem. Res.* **2000**, *33,* 889.
[103] G. Rastelli, A. D. Rio, G. Degliesposti, M. Sgobba, *J. Comput. Chem.* **2010**, *31,* 797.
[104] M. E. M. Noble, *Science* **2004**, *303,* 1800.
[105] R. Kim, J. Skolnick, *J. Comput. Chem.* **2008**, *29,* 1316.