JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Accounting for Conformational Variability in Protein−Ligand Docking with NMR-Guided Rescoring

Lars Skjærven,[†,⊥] Luca Codutti,[†,⊥] Andrea Angelini,[†] Manuela Grimaldi,[†,‡] Dorota Latek,[†] Peter Monecke,[§] Matthias K. Dreyer,[§] and Teresa Carlomagno*[,†]

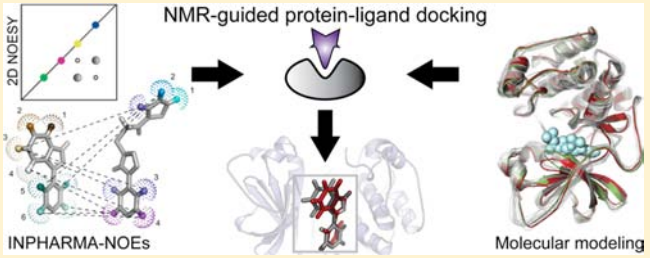
[†]EMBL, Structural and Computational Biology Unit, Meyerhofstraße 1, D-69117 Heidelberg, Germany

[‡]Department of Biomedical and Pharmaceutical Sciences, University of Salerno, Via Ponte Don Melillo 8, 84024 Fisciano (SA), Italy

[§]Sanofi-Aventis Deutschland GmbH R&D LGCR/Structure, Design & Informatics, Industriepark Höchst, Bldg. G838, D-65926 Frankfurt am Main, Germany

Supporting Information

ABSTRACT: A key component to success in structure-based drug design is reliable information on protein−ligand interactions. Recent development in NMR techniques has accelerated this process by overcoming some of the limitations of X-ray crystallography and computational protein−ligand docking. In this work we present a new scoring protocol based on NMR-derived interligand INPHARMA NOEs to guide the selection of computationally generated docking modes. We demonstrate the performance in a range of scenarios, encompassing traditionally difficult cases such as docking to homology models and ligand dependent domain rearrangements. Ambiguities associated with sparse experimental information are lifted by searching a consensus solution based on simultaneously fitting multiple ligand pairs. This study provides a previously unexplored integration between molecular modeling and experimental data, in which interligand NOEs represent the key element in the rescoring algorithm. The presented protocol should be widely applicable for protein−ligand docking also in a different context from drug design and highlights the important role of NMR-based approaches to describe intermolecular ligand−receptor interactions.



INTRODUCTION

Protein−ligand docking has emerged as an attractive and integrative computational approach in structure-based drug design (SBDD).[1,2] It allows for high throughput virtual screening of compound databases and subsequent prediction of protein−ligand binding modes. A reliable protein−ligand complex structure facilitates optimization of lead compounds to high affinity drug candidates.[3] High quality structural information is commonly obtained by X-ray crystallography, which is often limited by high costs and compatibility with crystallization—both issues potentially forcing the dismissal of the SBDD route for the target of interest. Conversely, computational docking protocols are rapid, inexpensive, and dependent only on the availability of an atomic resolution structure of the target protein. Despite being a driving force in SBDD, docking protocols suffer from a limited reliability of the prediction of the interaction mode for arbitrary protein−ligand pairs,[4,5] partly due to the inability to account for protein flexibility and entropic effects in an accurate manner.[6]

Docking routines commonly account for moderate protein flexibility either by allowing partial steric clashes, or by considering independent side-chain rotations.[7,8] Inclusion of backbone motions would in most cases greatly increase the degrees of freedom and negate on-the-fly sampling of protein conformational space. In the case of proteins which undergo structural changes beyond side-chain rearrangements upon ligand association, docking can be performed on a predefined ensemble of experimentally determined receptor structures (ensemble docking).[6,9] This seems to be a viable alternative when a representative set of structures is available.[10] Conversely, when an ensemble of experimental structures is not available, computational tools can aid in exploring the conformational space. Of these, molecular dynamics (MD) simulations have been extensively used to generate structural ensembles for docking experiments.[11−13] The limitation of ensemble docking relates to the *a priori* knowledge required to judge the available ensemble, and success might be strongly related to the presence of a member structure with high similarity to the "correct" receptor state with the query ligand bound.[7] Moreover, it inflicts additional pressure on the docking scoring functions, which should be competent to discriminate between receptor conformers, and not only between different orientations of the ligand in a defined receptor binding pocket.

Recent advances in nuclear magnetic resonance (NMR) spectroscopy have accelerated SBDD,[14,15] in particular with

Received: January 28, 2013
Published: April 8, 2013

   dx.doi.org/10.1021/ja4007468 | *J. Am. Chem. Soc.* 2013, 135, 5819−5827

respect to screening of low affinity compounds for lead identification.[16] Detailed information on the protein−ligand binding pose can be obtained by "protein-detecting" techniques, including chemical shift perturbation analysis from HSQC experiments,[17−19] and acquisition of protein−ligand NOEs[20]—both approaches facilitate filtering query docking modes. A general challenge in the "protein-detecting" techniques is the assignment of the protein resonances, which requires $^{13}C/^{15}N$—and possibly ^{2}H-labeled protein, and limits the size of the receptor to ∼50 kDa. As an alternative, "ligand-detecting" techniques can be used for low affinity ligands, such as (a) transferred nuclear Overhauser effect (NOE) experiments, with the capability to reveal the bioactive conformation of the ligand,[21,22] (b) interligand NOEs (ILOES) for identifying binding of two ligands to adjacent sites,[23] and (c) saturation transfer difference (STD) experiments, which can reveal which part of the ligand is involved in binding.[24]

The methodology entitled *INPHARMA* (Interligand NOEs for PHARmacophore MApping) exploits NOEs occurring between two ligands binding competitively to the same target protein measured in a "ligand-detecting" NOESY experiment.[25] These interligand NOE cross-peaks result from magnetization transfer mediated by receptor protons and are not limited by the protein size or availability of isotope-labeled protein. The technique requires low affinity ligands, making it attractive for lead generation. Interligand INPHARMA NOEs can be theoretically estimated for pairs of protein−ligand complexes (generated, e.g., by protein−ligand docking), and subsequently compared to the experimental ones to discriminate the correct binding modes (see ref 26 for a more extensive theoretical introduction). However, the solutions attained from fitting experimental and theoretical interligand NOEs are often not unique, and might lead to false positive hits or ambiguous answers. In previous work, the ambiguity was lifted either by manual inspection of the correlation graphs,[27] or by employing biochemical structure activity relationship (SAR) data.[28] Both ways are not viable when handling a larger number of ligands or binding poses.

Despite promising results for a variety of NMR-based techniques in SBDD, the majority of them have only been benchmarked using the receptor structures in which the protein−ligand was crystallized. Thus, their performance in more realistic docking experiments, where the exact conformation of the receptor structure is not known *a priori*, is still unclear. In previous work,[29] we have demonstrated that the performance of INPHARMA deteriorates when the starting structure for the docking experiments is inaccurate. In this work we aim at establishing a rigorous approach for induced-fit docking using a combination of NMR-derived NOE restraints and extensive molecular modeling. Protein Kinase A (PKA) is used as the model system, due to evident structural changes upon ligand binding. We present a novel algorithm, entitled *INPHARMA-STRING*, which robustly filters out false positive docking modes, even in the presence of only sparse NMR information. We demonstrate the performance of the new rescoring function on five diverse ligands and a series of realistic docking scenarios, ranging from the fully open apoenzyme to the inhibitor bound closed form of PKA. The highly complementary methods employed in this study reveal a previously unexplored potential of obtaining robust docking results for difficult scenarios.

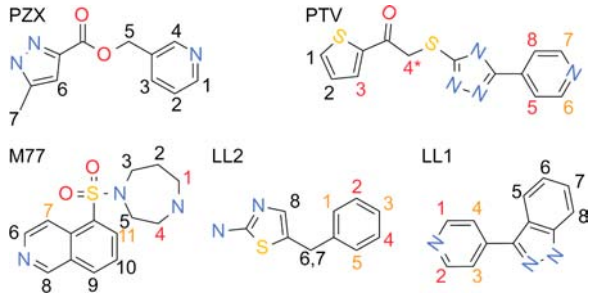


Figure 1. Structures of PKA ligands. Positions for NMR-detectable hydrogen moieties are marked numerically. Red and orange fonts denote overlapping chemical shifts between two or more protons, whereas protons whose resonance overlaps with the solvent are marked by an asterisk.

MATERIALS AND METHODS

Materials. A set of five PKA inhibitors was collected from our in-house compound library based on their chemical diversity and low affinity for PKA (Figure 1). The ligands, referred to by their PDB ligand ids (LL1, LL2, M77, PTV, and PZX) obtain dissociation constants (K_D) ranging from 6 to 30 μM. High resolution structures of PKA bound to ligands LL1, LL2, and M77 were collected from PDB with ids: 3DNE, 3DND, and 1Q8W, respectively. Structures of PKA bound to ligands with ids PTV and PZX were obtained by X-ray crystallography (see below).

Hamster PKA for NMR experiments was prepared starting from constructs described elsewhere,[27] and stored in perdeuterated PBS buffer (pH 7.4, [NaCl] = 150 mM). The buffers for both the ligand characterization and the INPHARMA experiments contained an addition of 5% in volume of DMSO. NMR samples contained ligands:protein in a 10:1 stoichiometric ratio for a final protein concentration of 25 μM.

INPHARMA Experiments and Calculations. We collected a total of 10 INPHARMA experiment sets, corresponding to all pairwise combinations of the five ligands. Each INPHARMA set was recorded as a fully interleaved scan by scan series of NOESY spectra with mixing times 300, 500, and 700 ms. Experimental acquisition details are listed in the Supporting Information. Spectral assignment and peak integration was made using NMRViewJ[30] and Felix 2007. The relative K_Ds of the ligands were estimated by means of STD competition experiments.[31] Ligands τ_c was set to 0.1 ns, a value in agreement with estimates obtained from T_1 inversion recovery and T_2 CPMG experiments performed with the *J*-coupling free PROJECT sequence.[32] The R_1 and R_2 values obtained were compared with theoretical decays of isotropic spheres with increasing τ_c. The concentration of the ligands was 250 μM, with the exception of PTV, which degraded over time. Its concentration was estimated for every experiment set and varied between 150 μM and 80 μM. The concentration of the complex was tuned by fitting experimental intraligand NOEs (Table S1). Theoretical interligand INPHARMA NOEs were calculated for all protein−ligand pairs in each of the docking scenarios using an in-house written program in accordance to the previous theoretical development, using a direct exchange model to describe the protein−ligand kinetics.[26]

Crystallization. Bovine PKA in complex with ligands PTV and PZX was crystallized following previously described procedures[33] (Supporting Information and Table S2). Coordinates are available from the RCSB protein data bank with accession codes 4IJ9 (ligand PTV) and 4IE9 (ligand PZX).

Structural Analysis and Cross-Docking. All PKA structures with a sequence identity higher than 98% to the human sequence were collected from the RCSB protein databank,[34] and analyzed using the Bio3D package.[35] Five representative high resolution structures with PDB id 3DND, 3AGM, 1Q8W, 1CDK, and 1CMK were selected as targets for cross-docking. In addition, a homology model of human PKA was generated with Modeller[36] based on the high resolution

dx.doi.org/10.1021/ja4007468 | J. Am. Chem. Soc. 2013, 135, 5819−5827

structure of protein kinase B (PKB; PDB id 2JDO). The evaluation of cross-docking of ligands LL1, LL2, M77, PTV, and PZX was carried out with Surflex[37] and Glide (Schrödinger), in which 20 docking modes were generated for each protein−ligand combination. Ensemble docking was carried out with Surflex on the PKA X-ray structures with PDB ids 3DND, 3AGM, 1Q8W, 1CDK, and 1CMK (see also Supporting Information for docking using a larger ensemble). All docking experiments were conducted using flexible ligands. Default values of parameters for both programs were used apart from an RMSD similarity filter set to 1 Å for the Surflex docking.

MD Simulations and Ensemble Docking. All-atomic models of the high resolution X-ray structures with PDB id 1CDK, 1CMK, and the PKA homology model (see above) were generated with AmberTools[38] using the corresponding Amber99SB force field.[39] MD simulations were carried out in explicit solvent employing a 1 fs time step under periodic boundary conditions (constant volume). Bonds involving hydrogen atoms were kept rigid using SHAKE, and particle mesh Ewald was applied with an 11 Å cutoff (see also Supporting Information for more details). Ligands, with which the structure was crystallized, were included in the simulations for PDB id 1CDK and the homology model. Ligand parameters were generated using the generalized Amber force field.[40]

Ensemble docking to cluster representatives derived from MD simulations was performed with Glide, and up to 20 docking modes per ligand were collected for each receptor structure. All protein−ligand models were issued to a 5 ps long MD simulation (generalized Born solvation model), and a total of 6 frames from each MD simulation were collected for further analysis (see Supporting Information for more details). The binding free energy was estimated by MM/GBSA calculations and the 50% highest energy binding modes (energetically unfavorable) were discarded to reduce the computational load. Simulations and trajectory analysis were performed with the Amber 11 software suite.

RESULTS

Structural Characterization and Docking Evaluation. PKA is an extensively characterized pharmacological target with more than 130 hits in the RCSB protein databank. To investigate this structural diversity, we collected all 88 structures with sequence identity >98% to human PKA. This includes ATP bound conformers as well as structures bound to ATP analogues, ligands, and peptides. Principal component analysis on the set of PKA structures, combined with structure based clustering, reveals that the PKA structures can be divided into three distinctive conformational states: (1) fully open, (2) intermediate, and (3) fully closed, deviating up to 1.9 and 2.1 Å in terms of Cα and heavy-atom binding site RMSD, respectively (Figure 2A; Figure S1).

Due to the relatively large structural variability, PKA appears as an ideal target to evaluate the ability of current docking programs to reproduce the binding poses of ligands while using receptor structures with an increasing structural deviation from that of the receptor−ligand complex. We therefore carried out cross-docking experiments, using two state-of-the-art docking programs and a representative set of receptor structures ranging from the fully open apo enzyme (PDB id 1CMK) to the fully closed conformation (PDB id 1CDK), as well as three structures showing an intermediate conformation (PDB ids 3DND, 3AGM, 1Q8W). In addition, the analysis was extended by performing ensemble docking to a pool of X-ray structures, and moreover, by including a low quality homology model of human PKA built from PKB (PDB id: 2JDO). We limit the docking evaluation to a set of five diverse ligands (LL1, LL2, M77, PTV, and PZX; Figure 1) also relevant for the subsequent NMR experiments. For this evaluation PKA in combination with the two new fragments with ids PTV and PZX were

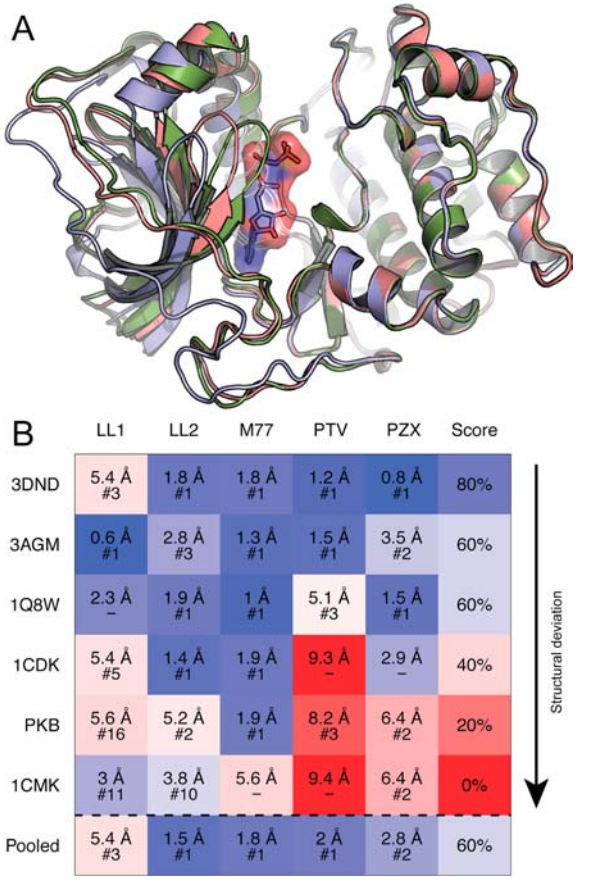


Figure 2. Structural variation of protein kinase A and cross-docking. (A) Visualization of three representative PKA structures superimposed on the large lobe: fully open apo (blue; PDB id 1CMK), intermediate (red; PDB id 3DND), and fully closed (green; PDB id 1CDK). Ligand present in PDB id 1CDK is shown in stick and surface representation. (B) Cross-docking (with Surflex) of five selected ligands (column-wise) to a set of six representative PKA structures (row-wise). Last row ("Pooled") shows the results of X-ray ensemble docking (3DND, 3AGM, 1Q8W, 1CDK, and 1CMK). Structural deviation from the first docking mode to the X-ray structure (RMSD value in each cell) is indicated as a heatmap in color range blue (low RMSD) to red (high RMSD). The ranking of the first docking mode with an RMSD < 2 Å to the X-ray binding pose is noted inside each cell. The docking score, defined as the proportion of ligands in which the correct binding pose is ranked as #1 by the docking program, is shown in the last column for each receptor structure. PKB: Homology model of PKA built from PKB.

crystallized and the protein−ligand structures were obtained at 2.6 and 1.9 Å resolution, respectively.

The results of the docking evaluation are summarized in Figure 2B (Surflex) and demonstrate a general decline of performance with increased structural deviation from the receptor conformation in which the ligands were crystallized (see also Figure S2 for Glide results and a more extensive analysis of ensemble docking). Docking to the high-resolution X-ray structures with PDB ids 3DND, 3AGM, and 1Q8W yields a success rate of 60−80% (defined as the proportion of ligands in which the correct binding pose is ranked as #1 by the docking program). These structures correspond to the intermediate conformation and are representative of the majority of the crystallized structures with a bound inhibitor (Figure S1C). Docking to the fully closed structure (PDB id 1CDK) resulted in a lower success rate (40%); furthermore,

dx.doi.org/10.1021/ja4007468 | J. Am. Chem. Soc. 2013, 135, 5819−5827

among the 20 first docking modes for ligands PTV and PZX there were no poses with an RMSD < 2 Å to the X-ray binding pose. Finally, the PKA homology model and the fully open apo enzyme (PDB id 1CMK) yielded docking success rates as low as 20% and 0%, respectively. This demonstrates the general challenge of obtaining reliable results with current docking protocols, even when high-resolution structures are available.

NMR-Based Rescoring. The deficiency in cross-docking experiments prompted us to develop a new rescoring protocol based on experimentally obtained NMR data. To this end, we measured 30 INPHARMA experiments corresponding to three NOESY mixing times for each pairwise combination of the five ligands of Figure 1 in the presence of the target protein (Figures S3–S5).

Of the 10 ligands combinations, LL2-PTV represents the most challenging case, due to the ambiguousness of the experimental data set (Figure S3). The majority of the 20 INPHARMA NOEs recorded (6–7 per mixing time) involve overlapping resonances of PTV hydrogen atoms 3, 5, and 8, located in two opposite parts of the molecule (Figure 1). On the contrary, ligand pairs such as LL1-M77 display more than 40 INPHARMA NOEs with fewer overlapping resonances.

As a first test case, we evaluated the ability to rescore the docking modes to the PKA structure in the intermediate conformational state (PDB id 3AGM). We thus estimated the theoretical INPHARMA NOEs expected for all combinations of the docking modes (i, j) for each ligand pair (L_A, L_B). The resulting theoretical NOEs were evaluated with respect to their agreement with the experimental NOEs using the Pearson correlation coefficient (R), which we refer to as *INPHARMA score*. The value of the INPHARMA score for two representative ligand pairs is correlated in Figure 3 to the structural deviation of the ligands from their correct orientations. Docking modes in agreement with the X-ray structure obtained a high correlation with the experimental NOEs for all 10 combinations (Figures 3 and S6). Notably, LL2 and PZX, which docked incorrectly to the 3AGM receptor (Figure 2B), were correctly rescored using a single INPHARMA run (Figure S6). However, the observation of false positives for many of the ligand pairs limits the confidence of the selection (Figures 3A and S6). The extensive presence of false positive underlines the need for a better rescoring algorithm that would leverage the ambiguities and deliver a precise and accurate selection of binding modes. Such an algorithm is described in the next paragraph.

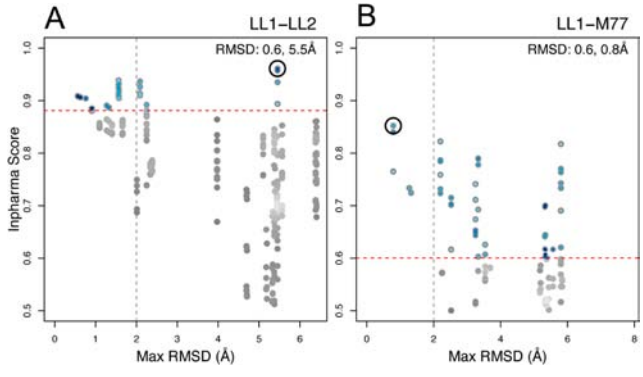


Figure 3. Evaluation of the INPHARMA score. The INPHARMA score vs the structural deviation (RMSD) to the correct binding pose is shown for 2 representative ligand pairs: (A) LL1-LL2, and (B) LL1-M77. Each dot represents one combination of two docking modes (L_{Ai}, L_{Bj}), and the maximum RMSD of the two ligands to the correct crystallographic binding pose is shown. The black circle marks the pair with highest INPHARMA score, and the related RMSD values are shown in the top right corner of each plot. Blue dots (above the red dashed line) depict the 12% highest scoring (by INPHARMA) docking modes and correspond to the set of binding modes needed to obtain a minimum consensus selection between all ligand pairs (see paragraph INPHARMA-STRING).

INPHARMA-STRING: A Consensus Criterion for Cross-Docking. In view of the presence of false positives in the INPHARMA selection, we sought a method that could simultaneously interpret the INPHARMA information from multiple ligand pairs to exclude false positive hits, in an automated approach. Since each individual INPHARMA experiment comprises a pair of ligands (i.e., L_A–L_B), each ligand also occurs in combinations with other ligands (e.g., L_A with L_C, L_D, etc.). The property of such a composition of the experiments can be exploited to determine a *consensus selection* of the docking modes. Let (L_A, L_B, L_C) be a set of 3 ligands whose docking modes are characterized by indices i, j, k, respectively. A string of docking modes (L_{Ai}, L_{Bj}, L_{Ck}) is selected if and only if all combinations of ligand docking modes (L_{Ai}–L_{Bj}, L_{Ai}–L_{Ck}, L_{Bj}–L_{Ck}) score high with respect to the experimental INPHARMA data. This approach is entitled INPHARMA-STRING, as a string of binding poses is scored and selected instead of pairs (Figure 4).

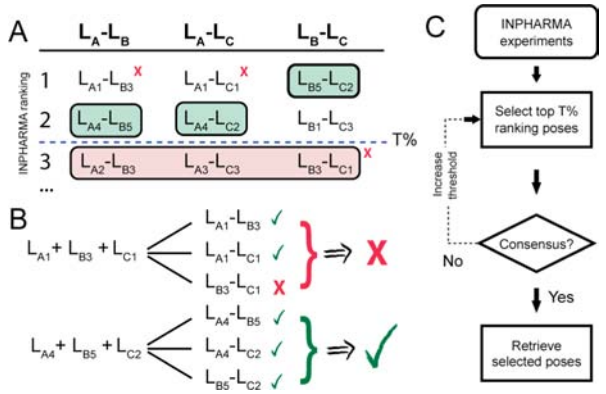


Figure 4. INPHARMA-STRING filtering protocol. Ranking (by INPHARMA score) of docking modes associated with three ligand pairs (L_A–L_B, L_A–L_C, L_B–L_C) is shown in (A). A consensus criterion is searched for each "string" of docking modes in which all pairwise combinations should be above the threshold (T%). The consensus is searched among the top T% (above blue dashed line) ranking docking modes. Here, pairs of docking modes (L_{A1}–L_{B3} and L_{A1}–L_{C1}) are discarded because (L_{B3}–L_{C1}) is below the threshold (A) and the string of docking modes (L_{A1}, L_{B3}, L_{C1}) does not fulfill the consensus criterion (B). If a minimal consensus is not found, the threshold (T) is increased and a new iteration is initiated including new docking modes for each experiment (C).

INPHARMA-STRING is applied in an iterative way. First, we select the top T (e.g., top 1%) combinations of binding modes for each ligands pair (ranked by their INPHARMA score). Among these selected poses, one or more N-tuple of binding poses is searched for which the consensus criterion is satisfied. If this N-tuple is found, the k docking modes for the N ligands are returned as the final selection. If no N-tuple in the pool satisfies the INPHARMA-STRING criterion, the percentage of poses (T) is increased until such an N-tuple is found.

dx.doi.org/10.1021/ja4007468 | J. Am. Chem. Soc. 2013, 135, 5819–5827

This iterative, data-driven approach is used to avoid selection bias by imposing a hard cutoff on the INPHARMA score.

We applied the consensus criterion for the docking modes of 3AGM by iteratively considering an increasing amount of docking modes, ranked by their INPHARMA scores, until we find a consensus quintuplet of binding modes for the five ligands. Using 12% of the top ranking docking modes yields a consensus selection of binding poses for all five ligands, showing a remarkably low RMSD to the X-ray structures of 0.6, 1.6, 0.8, 1.3, and 1.2 Å, respectively (Figure 5). Similar results

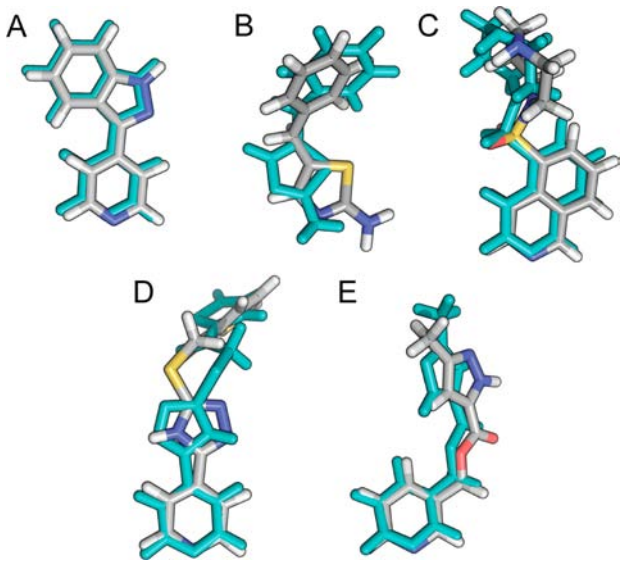


Figure 5. Consensus selection of docking modes. The final selection of docking modes to PDB id 3AGM of the ligands LL1 (A), LL2 (B), M77 (C), PTV (D), and PZX (E). The ligand binding pose in the X-ray structures are shown as gray sticks, while docking modes selected by INPHARMA-STRING are shown in cyan sticks. The related RMSD values to the X-ray structure for the selected binding modes are noted in Table 1.

are obtained for the rescoring of the ligand docking modes to the two other receptor structures in the intermediate conformation (Table 1). In addition, we benchmarked the same approach using only three ligands (the minimum input for INPHARMA-STRING). The performance is, as expected, lower than when using five ligands, and depends on the number and nature of the INPHARMA NOEs available for the three chosen ligands (Table S3). For this set of ligands, an unambiguous result was obtained starting from five ligands. Employing the NMR-based rescoring and the consensus criterion selection protocol effectively filters out false positives suggested by the docking program and by the pairwise INPHARMA selection. These results provide prospects for an improvement of binding pose prediction also in the more difficult scenarios presented in Figure 2B.

Table 1. Performance of INPHARMA-Guided Docking[a]

	LL1	LL2	M77	PTV	PZX
3AGM	0.3−0.6	0.6−1.6	0.8	1.3	0.5−1.3
3DND	0.4−1.1	1.5	0.8	2.1	0.2
1Q8W	1.9	1.6−1.9	2.2	1.2−4.3	1.6
Pooled	0.4	1.6	1.0	4.3*	0.2
1CDK$_{MD}$	1.3	1.3−1.5	0.9	2.3	0.9
PKB$_{MD}$	1.3	0.9	1.3	0.6	1.7
1CMK$_{MD}$	0.8	1.5	1.6	0.9	2.1

[a]RMSD values (in Å) for the predicted binding modes (by INPHARMA-STRING; to their respective X-ray structures) of the five ligands (column-wise) are provided for each docking scenario (row-wise). Results are presented for conventional and X-ray ensemble docking above the empty row, and MD ensemble docking (noted with MD in subscript) below the empty row. PKB: Homology model of PKA built from PKB (PDB id: 2JDO). *Correct orientation, wrong internal conformation.

Evaluation of Docking Experiments. In the case of the fully closed receptor structure (PDB id 1CDK) only LL2 and M77 were correctly predicted by the docking experiment (Figure 2B). Moreover, PTV and PZX did not reach an RMSD < 2 Å for the 20 first docking modes considered here, possibly due to the more closed binding site, which restrains the space available for the ligands (Figure 2A). This is evidently problematic, since the INPHARMA filtering protocol is dependent on the presence of the correct binding orientation in the pool of binding modes. To assess whether the proposed protocol would be able to identify the absence of correct binding modes in the pool, we repeated the INPHARMA-STRING protocol using the docking modes obtained for the fully closed receptor structure (Figure S7). With a cutoff on the INPHARMA score of 0.5 and using up to 50% of the pairs for each ligand combination, the consensus selection was not found. Thus, INPHARMA-STRING correctly identifies the absence of the correct binding mode in the pool of docked complexes.

INPHARMA-Guided Ensemble Docking. The deficiency to sample the correct binding pose of PTV and PZX in the fully closed receptor structure possibly relates to the shape of the binding pocket that cannot accommodate these ligands in the correct orientation. Consequently, a single static receptor structure might in many cases not be sufficient for an optimal docking experiment. To address this problem, we expanded the protocol described above to incorporate full protein flexibility by employing ensemble docking to MD-generated conformers.

For this purpose we conducted a 5 ns long MD simulation of PKA starting from the fully closed conformational state with bound ligand in an explicit solvent environment. The simulation provided a model of the accessible conformational space not only in the immediate vicinity of the starting structure, but also in the direction of the intermediate structure (defined by PDB id 3DND; Figure 6). To reduce the abundance of MD-generated conformers to a more affordable level for ensemble docking, the trajectory was processed by employing a binding site RMSD-based clustering. This resulted in a set of 10 representative conformers (Figure 6), which were subsequently used in a redocking experiment. For each cluster representative, we generated 20 docking modes per ligand. Each complex was subjected to a short local MD refinement to allow for mutual adaptation of the protein and ligand. Six snapshots per complex were retained along this trajectory, yielding a total of 1200 structures per protein−ligand complex. Of this ensemble, only the 50% with the lowest MM/GBSA free energy was retained (see Supporting Information for details). For the retained 600 models, the theoretical INPHARMA-NOEs were estimated and assessed with respect to the agreement with the experimental INPHARMA-NOEs. The corresponding evaluation of the INPHARMA score and RMSD is shown in Figure S8, and shows that 3 of 10 individual INPHARMA runs yielded the correct ligands orientations

dx.doi.org/10.1021/ja4007468 | J. Am. Chem. Soc. 2013, 135, 5819−5827

Figure 6. Conformational ensemble generated by molecular dynamics. The ensemble of cluster representatives obtained from an MD simulation starting from the closed (green; PDB id 1CDK) state of PKA is shown. The intermediate target conformation is colored red (PDB id 3DND), and cluster representatives are colored gray.

directly, while the remaining runs provided 2−3 alternative docking orientations per ligand. Finally, by applying INPHARMA-STRING, a selection of the five binding modes was provided, in which LL1, LL2, M77, and PZX, were below the critical threshold of 2 Å to the X-ray structure, and PTV obtained the RMSD of 2.3 Å (Table 1, Figure S8, Figure S9).

The same approach was followed for the homology model (PKA built from PKB, with a sequence identity of 42%), and an even better agreement with the X-ray structures was observed after applying INPHARMA-STRING, yielding RMSDs of 1.3, 0.9, 1.3, 0.6, and 1.7 Å for LL1, LL2, M77, PTV, and PZX, respectively (Figure S9, Figure S10). Thus, the proposed combination of MD simulations, ensemble docking, and INPHARMA-STRING rescoring provides a remarkable improvement of the ligand binding pose prediction, even when applied to receptor structures with relatively large structural deviation from the target receptor (Table 1).

Application to an Induced-Fit Scenario. The final test case in this study represents the fully open apo enzyme of PKA (PDB id 1CMK), in which the binding poses of all five ligands were incorrectly predicted in the conventional docking experiment (Figure 2B). Also in this case we attempted to rescore the docking modes (obtained on the single receptor) by INPHARMA-NOEs. Employing the consensus criterion yielded a selection only when considering the 36% top hits, again suggesting insufficient sampling of the correct poses in the docking experiments.

Consequently, we followed the approach for INPHARMA-guided ensemble docking to MD-generated conformers. In the case of the homology model and the fully closed state, the conformational variability of the protein binding pocket was modeled by MD simulations of the protein containing the ligand with which the crystallographic structure was obtained. Although these ligands differ from the five query ones, the presence of a ligand in the binding pocket is potentially important for maintaining a reasonable "holo shape" of the binding site throughout the MD simulation. Thus, for the fully open apo structure, an alternative strategy was necessary to model the conformational variability while conserving a holo shape of the binding site. Following previous work, which suggested the use of the largest ligand for generating by MD simulations conformational ensembles for docking,[41] we conducted a 40 ns long MD simulation of the top docking hit of the receptor−M77 complex. Starting from a binding site Cα RMSD of 1.4 Å, the MD simulation sampled the conformational space in the direction of the intermediate conformational state (PDB id 3DND) of PKA, which decreased the RMSD down to 0.8 Å (Figure S11). The resulting trajectory was then processed following the clustering analysis and ensemble docking protocol described above. Applying INPHARMA-STRING resulted in a selection of docking modes in a remarkable agreement with the X-ray structures (RMSD values of 0.8, 1.5, 1.6, 0.9, and 2.1 Å, for ligands LL1, LL2, M77, PTV, and PZX, respectively; Table 1, and Figure S12). Moreover, the receptor structures associated with each selected ligand docking mode display a good agreement with the intermediate conformational state of PKA (Figure 7).

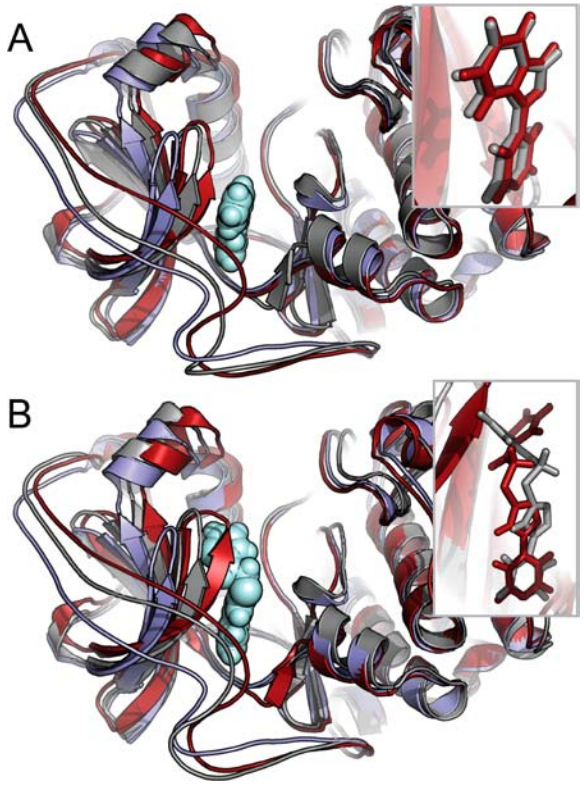


Figure 7. Selected protein−ligand conformers. Receptor conformations in complex with selected binding modes of LL1 (A; maximum deviation) and PTV (B; minimum deviation) are shown for the ensemble docking to the fully open apo structure. Modeled receptor structure is colored gray, while the X-ray structures of the open apo and intermediate conformational states are shown in blue and red, respectively. Cα binding site RMSD of the selected receptor structures are 1.2, 1.0, 1.0, 0.9, and 1.1 Å, for the hits corresponding to ligands LL1, LL2, M77, PTV, and PZX, respectively. The inset shows the predicted binding pose for the two ligands (gray) in comparison to the X-ray structure (red).

Although the deviation to the target X-ray structure is ∼1 Å, all selected receptor structures are considerably closer to the intermediate conformation than the starting structure. This demonstrates the ability of the INPHARMA-guided ensemble docking protocol to rigorously predict protein−ligand binding modes even in the presence of large structural changes.

dx.doi.org/10.1021/ja4007468 | J. Am. Chem. Soc. 2013, 135, 5819−5827

New Protocol for INPHARMA-Guided Rescoring. The docking experiments conducted here portray representative scenarios, in which the available target structure is (1) a structure with a different inhibitor/substrate, (2) a low quality homology model, or (3) an apo structure. On the basis of these experiments we aimed at designing an extended protocol for rescoring protein−ligand binding modes obtained from either conventional or ensemble based docking experiments (Figure 8). In essence, we propose a protocol, which contains two paths

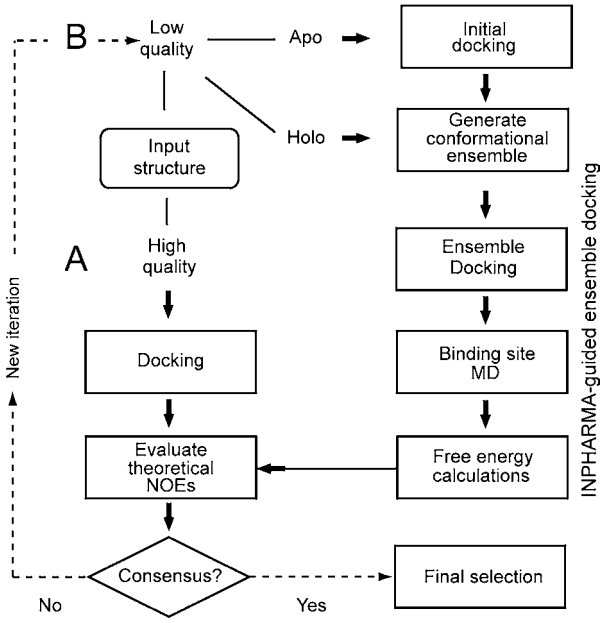


Figure 8. Overview of the INPHARMA-guided ensemble docking protocol. Depending on the input structure two paths are proposed: (A) In the case of a high quality "rigid" receptor structure, filtering by INPHARMA-STRING is performed directly on the docking modes to the single static receptor. In this scenario, flexibility is only incorporated through minimization of the docking modes. (B) Alternatively, with a low quality structure, or in the case of a flexible target, INPHARMA-guided ensemble docking can be performed on MD-generated conformers. Here, the receptor structures are collected from MD simulations, in which protein flexibility is explicitly modeled in a solvated environment.

determined either by the *a priori* knowledge of the query system, or by INPHARMA-guided evaluation of the docking modes. Under the assumption that the target protein bound with the query ligands would adopt a similar conformation as the available protein structure, a direct rescoring of the docking modes obtained on a single receptor structure can be performed (Figure 8A). In the presence of several X-ray structures of the target protein, the same approach can be utilized for ensemble docking to the pool of X-ray structures.

Conversely, if the INPHARMA based consensus selection indicates insufficient conformational sampling of the ligand binding space, or when docking to a low-quality homology model or an apo receptor, we suggest using a more extensive approach, utilizing MD-generated ensemble docking (Figure 8B). In this branch of the protocol, we propose the generation of conformers through MD simulations of a holo state of the receptor. Ideally, an available protein−ligand X-ray structure can be used as a starting structure for the MD simulation. If only an apo structure is available, a more realistic "holo"-like conformational ensemble can be generated by performing an initial docking to the apo structure, in which one or more starting structures can be issued to subsequent MD simulations. Prior to the calculation of the theoretical INPHARMA-NOEs we have included two additional computational steps: a short binding site MD of each docking mode, which allows for mutual protein−ligand adaptation, and subsequent filtering of high energy conformations.

■ DISCUSSION

This work demonstrates the potential of combining extensive molecular modeling with NMR-guided ensemble docking to decipher receptor−ligand interactions in the case of induced-fit, and highlight the importance of NMR in SBDD.[42,43] NMR derived interligand INPHARMA-NOEs represent the key element in the new rescoring function of computationally generated ligand binding modes. Based on docking scenarios containing ligand-induced domain movements of PKA, we develop an extended protocol (INPHARMA-STRING), which rigorously determines the correct protein−ligand binding modes.

Modeling induced-fit effects represents one of the major challenges in protein−ligand docking, and the ability to rapidly obtain a reliable model of the complex would potentially accelerate SBDD. Although side-chain movements are found in only 50% of the structures crystallized with multiple ligands, and backbone motions in only 7%,[44] it is commonly recognized that proteins can exists in a number of substates.[45,46] The apparent "lack" of conformational variability suggested by the PDB has been attributed to the fact that the database may be artificially enriched by "rigid" proteins[7] as compared to "flexible" proteins or even intrinsically unstable proteins, which only obtain a defined structure upon ligand-binding.[47] For this reason the ability to model protein flexibility beyond side-chain rearrangements has been a hot topic for over a decade, and impressive development has opened the way to improved docking performance in the presence of moderate induced-fit effects.[48−50] At the same time, MD simulations have taken an important role in modeling large scale ligand-induced conformational changes (see, e.g., ref 51). Modeling techniques have shown the ability to complement experimental approaches, in particular in light of the limited picture a single crystallographic structure represents.[52] For the application of binding pose prediction, the protein conformational space can be searched by MD to generate a pool of conformers for ensemble docking protocols.[11,13] The primary benefit of employing MD simulations to generate conformational ensembles is the fully flexible model guided by empiric force fields in a realistic explicit solvent environment. Although detailed classical simulations yielded reasonable results here, alternative approaches such as accelerated simulation protocols,[51] loop-modeling, or more specialized protocols for generating holo structures based on apo structures[53,54] could be suitable for other systems. In the case of PKA we illustrate by combining modeling and INPHARMA-guided docking that selected receptor conformers obtain a relatively small structural deviation from the target X-ray structures. This demonstrate the remarkable feature of the INPHARMA data, which contain information not only on the ligand binding pose, but also, indirectly, on the conformation of the binding pocket mediating the magnetization transfer.

The difficulty of cross-docking has been noted in several previous studies using different docking software and scoring functions.[5,55−57] Depending on the data set the overall docking

dx.doi.org/10.1021/ja4007468 | J. Am. Chem. Soc. 2013, 135, 5819−5827

performance ranges from 100% in retrospective self-docking to ~50%, or even lower, for cross-docking.[56] The reported success rates are in good agreement with the ones we observe. Here, we associate the evaluation of the docking performance to the structural deviation between the target complex and the query receptor structure by utilizing receptor structures with a binding site RMSD ranging from 0.5 to 1.5 Å (Figure 2B). Although the general docking performance seems to correlate with an increased structural deviation, there is also a significant deviation of the binding pose predictions between the two docking programs used in this study (Figure 2B; Figure S2). While ensemble docking might offer a solution to the structural variability of the receptor, the lack of a robust computational approach for estimating binding free energies of the query docking modes limits its value.[7] Ensemble docking protocols have shown good performance with regard to sampling binding modes that include the correct one;[10] on the other hand, the inaccurate ranking based on *in silico* binding free energies does not allow the unambiguous identification of the "correct" binding mode in many cases. Consequently, employing a scoring function that uses experimental information appears as a promising alternative to the computed free binding energies.

Among the two major techniques employed to obtain structural information on protein−ligand complexes, NMR emerges for its easy applicability to weakly binding ligands. The benefit with the INPHARMA technique employed in this study is its applicability to protein targets of any size. In addition, $^{13}C/^{15}N$ labeling of the receptor from expression systems and assignment of the receptor resonances are not needed.[25] Conversely, the requirement is on suitable ligand binding constants, as the observation of INPHARMA-NOEs requires a dissociation constant of at least 100 Hz.[26] Given a sufficient number of interligand INPHARMA-NOEs the correct binding modes can be detected with high accuracy. However, fitting theoretical INPHARMA-NOEs estimated from a pair of docking modes of two ligands to the experimental NOEs does not always provide a unique solution (Figure 3A).[27] In this work we address this ambiguity by considering multiple ligand pairs simultaneously. Each new ligand pair provides new information that leverages the ambiguities associated with any individual pair. For this purpose we have presented a sophisticated computational procedure, which exploits a criterion of consensus between individual INPHARMA experiments and their corresponding ligand pairs. We show with five ligands that the approach is robust with respect to false positives and allows the simultaneous determination of the correct binding mode of multiple ligands. The procedure also facilitates the remarkable feature of identifying docking experiments in which the correct ligand orientation has not been sampled.

Inclusion of additional experiments, which similarly do not depend on the NMR resonances of the receptor, can potentially be an important supplement for the filtering routine presented here. More specifically, obtaining intraligand restraints (i.e., proton−proton distances) from transferred NOE experiments, or ligand dihedral angles by transferred cross-correlated relaxation rates experiments,[58,59] would provide an additional source of filtering the pool of generated docking modes. Alternatively, such information could also be used to dock with a rigid ligand, whose conformation has been determined beforehand through transferred-NOEs. In this work, we decided to discard the transferred-NOEs information and to use flexible ligands in the docking protocol. Our objective was to demonstrate that INPHARMA-NOEs alone are sufficient to discriminate between docking modes and bound ligand conformations, even in the absence of high quality receptor structures. The overall improvement of binding pose prediction (Table 1) over computational docking clearly demonstrates the potential of the technique.

CONCLUSIONS

In summary, we have introduced a new rescoring function of docking modes based on NMR-derived NOE data. By exploiting a criterion of consensus between the docking modes, the new algorithm, entitled INPHARMA-STRING, shows a remarkable performance of filtering false positives, even in the presence of sparse NMR data. Combined with extensive molecular modeling and simulation techniques, we show the impact of the proposed protocol also for induced-fit docking. This protocol should be widely applicable for protein−ligand docking and highlights the important role of NMR-based approaches in structure-based drug design.

ASSOCIATED CONTENT

Supporting Information

NOE intensities, evaluation of INPHARMA-score vs RMSD, analysis of ensemble docking, visualization of predicted binding modes, and more details on the methodological procedures are given. This information is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: carlomag@embl.de.

Author Contributions

⊥Lars Skjærven and Luca Codutti contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

T.C. and A.A. acknowledge support from the EMBL, DFG (grant CA 294/6-1) and the BMBF. L.S. and L.C. acknowledge support from the BMBF through grant 0315870B to T.C. We thank Adam Mazur (MPI, Göttingen) for software discussions and technical help, and Prof. Dirk Bossemeyer (DKFZ Heidelberg, Germany) for providing recombinant bovine PKA for crystallization.

ABBREVIATIONS

NMR, nuclear magnetic resonance; PKA, protein kinase A; STD, saturation transfer difference; MD, molecular dynamics; NOE, nuclear Overhauser effect; SBDD, structure-based drug design; SAR, structure activity relationship

REFERENCES

(1) Jorgensen, W. L. *Science* **2004**, *303*, 1813−1818.
(2) Jorgensen, W. L. *Acc. Chem. Res.* **2009**, *42*, 724−733.
(3) Jorgensen, W. L.; Bollini, M.; Thakur, V. V; Domaoal, R. A.; Spasov, K. A.; Anderson, K. S. *J. Am. Chem. Soc.* **2011**, *133*, 15686−15696.
(4) Davis, I. W.; Raha, K.; Head, M. S.; Baker, D. *Protein Sci.* **2009**, *18*, 1998−2002.
(5) Verdonk, M. L.; Mortenson, P. N.; Hall, R. J.; Hartshorn, M. J.; Murray, C. W. *J. Chem. Inf. Model.* **2008**, *48*, 2214−2225.
(6) Totrov, M.; Abagyan, R. *Curr. Opin. Struct. Biol.* **2008**, *18*, 178−184.

(7) Cozzini, P.; Kellogg, G. E.; Spyrakis, F.; Abraham, D. J.; Costantino, G.; Emerson, A.; Fanelli, F.; Gohlke, H.; Kuhn, L. A.; Morris, G. M.; Orozco, M.; Pertinhez, T. A.; Rizzi, M.; Sotriffer, C. A. *J. Med. Chem.* **2008**, *51*, 6237−6255.
(8) Teodoro, M. L.; Phillips, G. N.; Kavraki, L. E. *J. Comput. Biol.* **2003**, *10*, 617−634.
(9) Knegtel, R. M.; Kuntz, I. D.; Oshiro, C. M. *J. Mol. Biol.* **1997**, *266*, 424−440.
(10) Barril, X.; Morley, S. D. *J. Med. Chem.* **2005**, *48*, 4432−4443.
(11) Carlson, H. A.; McCammon, J. A. *Mol. Pharmacol.* **2000**, *57*, 213−218.
(12) Amaro, R. E.; Baron, R.; McCammon, J. A. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 693−705.
(13) Ivetac, A.; McCammon, J. A. *Curr. Pharm. Des.* **2011**, *17*, 1663−1671.
(14) Meyer, B.; Peters, T. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 864−890.
(15) Carlomagno, T. *Annual Review of Biophysics and Biomolecular Structure* **2005**, *34*, 245−266.
(16) Hajduk, P. J.; Greer, J. *Nat. Rev. Drug Discovery* **2007**, *6*, 211−219.
(17) Riedinger, C.; Endicott, J. A.; Kemp, S. J.; Smyth, L. A.; Watson, A.; Valeur, E.; Golding, B. T.; Griffin, R. J.; Hardcastle, I. R.; Noble, M. E.; McDonnell, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 16038−16044.
(18) Stark, J.; Powers, R. *J. Am. Chem. Soc.* **2008**, *130*, 535−545.
(19) Wang, B.; Raha, K.; Merz, K. M. *J. Am. Chem. Soc.* **2004**, *126*, 11430−11431.
(20) Constantine, K. L.; Davis, M. E.; Metzler, W. J.; Mueller, L.; Claus, B. L. *J. Am. Chem. Soc.* **2006**, *128*, 7252−7263.
(21) Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1982**, *48*, 402−417.
(22) Ni, F.; Scheraga, H. A. *Acc. Chem. Res.* **1994**, *27*, 257−264.
(23) Li, D.; DeRose, E. F.; London, R. E. *J. Biomol. NMR* **1999**, *15*, 71−76.
(24) Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108−6117.
(25) Sánchez-Pedregal, V. M.; Reese, M.; Meiler, J.; Blommers, M. J. J.; Griesinger, C.; Carlomagno, T. *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 4172−4175.
(26) Orts, J.; Griesinger, C.; Carlomagno, T. *J. Magn. Reson.* **2009**, *200*, 64−73.
(27) Orts, J.; Tuma, J.; Reese, M.; Grimm, S. K.; Monecke, P.; Bartoschek, S.; Schiffer, A.; Wendt, K. U.; Griesinger, C.; Carlomagno, T. *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 7736−7740.
(28) Reese, M.; Sánchez-Pedregal, V. M.; Kubicek, K.; Meiler, J.; Blommers, M. J. J.; Griesinger, C.; Carlomagno, T. *Angew. Chem., Int. Ed. Engl.* **2007**, *46*, 1864−1868.
(29) Orts, J.; Bartoschek, S.; Griesinger, C.; Monecke, P.; Carlomagno, T. *J. Biomol. NMR* **2011**, *52*, 23−30.
(30) Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603−14.
(31) Wang, Y.-S.; Liu, D.; Wyss, D. F. *Magn. Reson. Chem.* **2004**, *42*, 485−489.
(32) Aguilar, J. A.; Nilsson, M.; Bodenhausen, G.; Morris, G. A. *Chem. Commun.* **2012**, *48*, 811−813.
(33) Engh, R. A.; Girod, A.; Kinzel, V.; Huber, R.; Bossemeyer, D. *J. Biol. Chem.* **1996**, *271*, 26157−26164.
(34) Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C. *Acta Crystallogr., Sect. D* **2002**, *58*, 899−907.
(35) Grant, B. J.; Rodrigues, A. P. C.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. D. *Bioinformatics* **2006**, *22*, 2695−6.
(36) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779−815.
(37) Jain, A. N. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 281−306.
(38) Case, D. A.; Darden, T. A.; T.E. Cheatham, I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*; University of California, San Francisco, 2012.
(39) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712−725.
(40) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(41) Rueda, M.; Bottegoni, G.; Abagyan, R. *J. Chem. Inf. Model.* **2010**, *50*, 186−193.
(42) Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T. L.; Homans, S. W.; Kessler, H.; Luchinat, C.; Meyer, B.; Oschkinat, H.; Peng, J.; Schwalbe, H.; Siegal, G. *Nat. Rev. Drug Discovery* **2008**, *7*, 738−45.
(43) Carlomagno, T. *Nat. Prod. Rep.* **2012**, *29*, 536−554.
(44) Boström, J.; Hogner, A.; Schmitt, S. *J. Med. Chem.* **2006**, *49*, 6716−6725.
(45) Frauenfelder, H.; McMahon, B. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4795−4797.
(46) Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450*, 964−972.
(47) Tompa, P. *Trends Biochem. Sci.* **2002**, *27*, 527−533.
(48) Davis, I. W.; Baker, D. *J. Mol. Biol.* **2009**, *385*, 381−392.
(49) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. *J. Med. Chem.* **2006**, *49*, 534−553.
(50) Bottegoni, G.; Kufareva, I.; Totrov, M.; Abagyan, R. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 311−325.
(51) Skjærven, L.; Reuter, N.; Martinez, A. *Future Medicinal Chemistry* **2011**, *3*, 2079−2100.
(52) Furnham, N.; Blundell, T. L.; DePristo, M. A.; Terwilliger, T. C. *Nat. Struct. Mol. Biol.* **2006**, *13*, 184−5 discussion 185.
(53) Abagyan, R.; Kufareva, I. *Methods Mol. Biol.* **2009**, *575*, 249−279.
(54) Seeliger, D.; De Groot, B. L. *PLoS Comput. Biol.* **2010**, *6*, e1000634.
(55) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.* **2006**, *49*, 5912−5931.
(56) Cavasotto, C. N.; Abagyan, R. A. *J. Mol. Biol.* **2004**, *337*, 209−225.
(57) Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. *J. Med. Chem.* **2004**, *47*, 45−55.
(58) Carlomagno, T.; Blommers, M. J.; Meiler, J.; Cuenoud, B.; Griesinger, C. *J. Am. Chem. Soc.* **2001**, *123*, 7364−7370.
(59) Carlomagno, T.; Sánchez, V. M.; Blommers, M. J. J.; Griesinger, C. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 2515−2517.