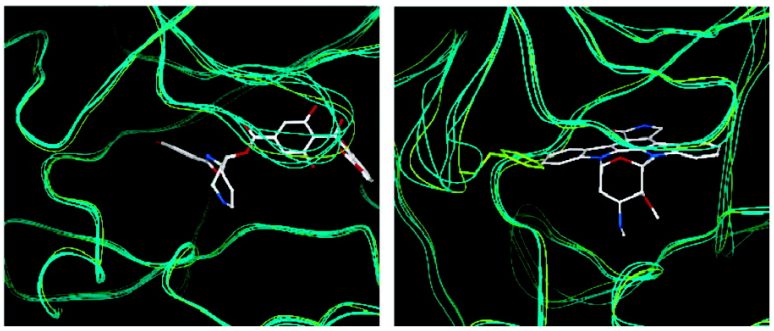
Article

Representing Receptor Flexibility in Ligand Docking through Relevant Normal Modes

Claudio N. Cavasotto, Julio A. Kovacs, and Ruben A. Abagyan

J. Am. Chem. Soc., **2005**, 127 (26), 9632-9640• DOI: 10.1021/ja042260c • Publication Date (Web): 14 June 2005



J|A|C|S
ARTICLES
Published on Web 06/14/2005

Representing Receptor Flexibility in Ligand Docking through Relevant Normal Modes

Claudio N. Cavasotto,*,† Julio A. Kovacs,‡ and Ruben A. Abagyan‡

Contribution from Molsoft LLC, 3366 North Torrey Pines Court, Suite 300, La Jolla, California 92037, and The Scripps Research Institute, 10550 North Torrey Pines Road, Mail TPC-28, La Jolla, California 92037

Received December 23, 2004; E-mail: claudio@molsoft.com

Abstract: Inspired by the current representation of the ligand−receptor binding process, a normal-mode-based methodology is presented to incorporate receptor flexibility in ligand docking and virtual screening. However, the systematic representation of the deformation space grows geometrically with the number of modes, and furthermore, midscale loop rearrangements like those found in protein kinase binding pockets cannot be accounted for with the first lowest-frequency modes. We thus introduced a measure of relevance of normal modes on a given region of interest and showed that only very few modes in the low-frequency range are necessary and sufficient to describe loop flexibility in cAMP-dependent protein kinase. We used this approach to generate an ensemble of representative receptor backbone conformations by perturbing the structure along a combination of *relevant* modes. Each ensemble conformation is complexed with known non-native binders to optimize the position of the binding-pocket side chains through a full flexible docking procedure. The multiple receptor conformations thus obtained are used in a small-scale virtual screening using receptor ensemble docking. We evaluated this algorithm on holo and apo structures of cAMP-dependent protein kinase that exhibit backbone rearrangements on two independent loop regions close to the binding pocket. Docking accuracy is improved, since the ligands considered in the virtual screening docked within 1.5 Å to at least one of the structures. The discrimination between binders and nonbinders is also enhanced, as shown by the improvement of the enrichment factor. This constitutes a new step toward the systematic integration of flexible ligand−flexible receptor docking tools in structure-based drug discovery.

1. Introduction

Computer-aided drug discovery through ligand docking-based virtual screening is already a key component in the lengthy and costly process of developing new drugs.[1,2] The capability to correctly predict ligand−protein interactions is fundamental to any accurate docking algorithm and the necessary starting point for any reliable virtual screening protocol. Molecular flexibility is critical for a thorough understanding of the principles that govern ligand binding to proteins. Structural changes in the receptor upon ligand binding is a very common phenomenon;[3] hence, ignoring this effect might have a strong impact on ligand docking[4] and virtual screening. The implications of protein flexibility in drug discovery have been recently reviewed.[5,6] Dealing with receptor flexibility is in many cases crucial to accurately predict the orientation and interactions of ligands within the binding pocket.[7] So the big challenge ahead is to routinely incorporate flexibility considerations into structure-based drug discovery in an affordable computing time.

Early attempts to include protein flexibility in ligand docking, such as soft docking[8] and partial side-chain flexibility[9,10] among others, have been reviewed.[11−13] However, most of these methods do not include backbone rearrangements, and explicit sampling of side chains is an unsurmountable drawback in virtual screening of large chemical libraries.

The use of multiple receptor conformations (MRCs) (either experimental or in silico generated) in ligand docking seems probably the best choice to date.[11,13] However, questions about how this should be efficiently accomplished are open. An important advantage of this approach is that the structural space of the binding pocket can be represented in a virtual screening process, even in the case of loop displacements.[14] In the early days, experimental structures were used to derive average interaction grids that became targets for rigid-ligand docking.[15] The FlexE method incorporates flexibility by combining rigid-

† Molsoft LLC.
‡ The Scripps Research Institute.

(1) Amzel, L. M. *Curr. Opin. Biotechnol.* **1998**, *9*, 366−369.
(2) Shoichet, B. K.; McGovern, S. L.; Wei, B. Q.; Irwin, J. J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 439−446.
(3) Najmanovich, R.; Kuttner, J.; Sobolev, V.; Edelman, M. *Proteins: Struct., Funct. Bioinf.* **2000**, *39*, 261−268.
(4) Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. *J. Med. Chem.* **2004**, *47*, 45−55.
(5) Davis, A. M.; Teague, S. J. *Angew. Chem., Int. Ed.* **1999**, *38*, 736−749.
(6) Teague, S. J. *Nat. Rev. Drug Discov.* **2003**, *2*, 527−541.
(7) Cavasotto, C. N.; et al. *J. Med. Chem.* **2004**, *47*, 4360−4372.
(8) Jiang, F.; Kim, S. H. *J. Mol. Biol.* **1991**, *219*, 79−102.
(9) Leach, A. R. *J. Mol. Biol.* **1994**, *235*, 345−356.
(10) Jones, G.; Willett, P.; Glen, R. C. *J. Mol. Biol.* **1995**, *245*, 43−53.
(11) Carlson, H. A. *Curr. Pharm. Des.* **2002**, *8*, 1571−1578.
(12) Teodoro, M. L.; Kavraki, L. E. *Curr. Pharm. Des.* **2003**, *9*, 1635−1648.
(13) Cavasotto, C. N.; Orry, A. J. W.; Abagyan, R. A. *Curr. Comput.-Aided Drug Des.* **2005**, in press.
(14) Cavasotto, C. N.; Abagyan, R. *J. Mol. Biol.* **2004**, *337*, 209−225.

10.1021/ja042260c CCC: $30.25

ligand docking to a combinatorial ensemble built from alternative discrete conformations of structures sharing a similar backbone trace.[16] A method that recombines multiple flexible regions into a discrete set of receptor conformations and that scales linearly with receptor flexibility has been recently presented and evaluated.[17] The ensemble of structures collected from molecular dynamimcs (MD) were also used to determine the correct geometry of a ligand−enzyme complex by docking and energy evaluation of selected complex structures.[18] In another interesting development, snapshots taken from MD simulations were employed to construct a receptor-based pharmacophore model of the HIV-1 integrase that was validated experimentally.[19] This method has been recently extended to the case of the apo HIV-1 protease.[20] Normal-mode analysis (NMA) has also been used to study the induced fit in HIV integrase,[21] and very recently a method of molecular dynamics and harmonic dynamics has been proposed and tested to study the docking of HIV-1 protease and its ligand.[22]

There is growing evidence about the relationship between pre-existing conformations of the receptor unbound state in equilibrium and structural changes upon ligand binding.[23−25] In those cases that cannot adjust to the rigid "lock-and-key" model,[26] ligand binding is seen as a combination of a conformational selection stage of partially fitting structures followed by minor structural rearrangements within the complex (induced-fit stage).[25] This led very recently to the formulation of the ligand binding process in terms of linear response theory, whereby the response of structures to ligand binding is predicted using the conformational ensemble of the unbound (unperturbed) state.[27] It has also been demonstrated that equilibrium conformations of a protein can be represented using low-frequency normal modes.[28−30]

On the basis of this evidence, we propose a low-frequency normal-mode-based algorithm to generate MRCs and thus incorporate receptor flexibility in ligand docking and virtual screening. In an attempt to drastically reduce the dimension of the conformational space that grows geometrically with the number of modes considered, we introduce a *measure of relevance* of normal modes on selected regions of interest known to be important for ligand binding. We found that very few modes are critical and sufficient to represent binding pocket plasticity in protein kinases, and we describe one algorithm to find those modes. Perturbation along the *relevant* modes followed by full flexible docking of known ligands to optimize positioning of side chains is then used to generate an ensemble of de novo alternative conformations. The structurally different ligand-binding pockets thus generated were used as starting points for a receptor ensemble docking (RED). We tested this methodology in both holo and uncomplexed structures of the cAMP-dependent protein kinase (cAPK, PKA). A small-scale virtual screening showed a significantly better discrimination of binders from nonbinders, which was evident by the improvement in the enrichment factors. Importantly, our procedure did not make custom-fit pockets that only accommodated the ligand used in the optimization. This constitutes a new advance in the challenging task of taking into account protein flexibility in structure-based drug discovery.

2. Results and Discussion

2.1. Foundations of the Normal-Mode-Based Approach in Receptor Ensemble Docking. It has been shown that ignoring protein flexibility in cAPK and other protein kinases is the reason ligands belonging to certain chemical spaces fail to dock correctly using the rigid receptor approach.[14] Our goal is to show that even when only one crystal structure (apo or holo) and a few binders are known, alternative structural conformations can be obtained by distortion along normal modes followed by full flexible docking of non-native ligands. These two steps are inspired in the ligand-binding process outlined in the Introduction: by perturbing along normal modes, we intend to represent equilibrium conformations of the receptor, while docking of known binders with flexible side chains is a way to generate structural rearrangements within the binding pocket (induced-fit stage). The lowest-energy complex should correspond to the structural conformer to be found experimentally. The receptor structures thus obtained could be used in a small-scale virtual screening using RED. We acknowledge that side-chain flexibility and other local motions might be not completely uncoupled from the docking process. In the future we plan to couple continuous changes along relevant modes with the docking step. In this paper, we are making the first step toward that direction by taking into account partial coupling through the side-chain optimization stage with known non-native binders.

We should note, however, that in certain cases like protein kinases, where binding pocket plasticity upon ligand binding is mainly concentrated in the gly-rich and C-term loops, any attempt to reproduce this sort of movements with normal modes should take into account the following:

(a) the first very low-frequency modes are associated with very low-energy large-scale dynamics, and thus do not represent the more localized and intermediate-amplitude loop rearrangement;

(b) the representation of the conformational space grows geometrically with the number of normal modes considered; and

(c) inclusion of low-energy modes not relevant to the region of interest will add unnecessary noise to the energy calculation of the system.

For this reason, we introduced the notion of *measure of relevance* for each mode on a region of interest, thus limiting the number of modes to be used in generating MRCs.

The main steps of our methodology can be summarized as follows:

(15) Knegtel, R. M.; Kuntz, I. D.; Oshiro, C. M. *J. Mol. Biol.* **1997**, *266*, 424−40.
(16) Claussen, H.; Buning, C.; Rarey, M.; Lengauer, T. *J. Mol. Biol.* **2001**, *308*, 377−395.
(17) Wei, B. Q.; Weaver, L. H.; Ferrari, A. M.; Matthews, B. W.; Shoichet, B. K. *J. Mol. Biol.* **2004**, *337*, 1161−1182.
(18) Lin, J. H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. *J. Am. Chem. Soc.* **2002**, *124*, 5632−5633.
(19) Carlson, H. A.; Masukawa, K. M.; Rubins, K.; Bushman, F. D.; Jorgensen, W. L.; Lins, R. D.; Briggs, J. M.; McCammon, J. A. *J. Med. Chem.* **2000**, *43*, 2100−2014.
(20) Meagher, K. L.; Carlson, H. A. *J. Am. Chem. Soc.* **2004**, *126*, 13276−13281.
(21) Keserû, G. M.; Kolossváry, I. *J. Am. Chem. Soc.* **2001**, *123*, 12708−12709.
(22) Tatsumi, R.; Fukunishi, Y.; Nakamura, H. *J. Comput. Chem.* **2004**, *25*, 1995−2005.
(23) Ma, B.; Kumar, S.; Tsai, C. J.; Nussinov, R. *Protein Eng.* **1999**, *12*, 713−720.
(24) Tsai, C. J.; Kumar, S.; Ma, B.; Nussinov, R. *Protein Sci.* **1999**, *8*, 1181−1190.
(25) Bosshard, H. R. *News Physiol. Sci.* **2001**, *16*, 171−173.
(26) Fischer, E. *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2984−2993.
(27) Ikeguchi, M.; Ueno, J.; Sato, M.; Kidera, A. *Phys. Rev. Lett.* **2005**, *94*, 078102.
(28) Noguti, T.; Gō, N. *Nature* **1982**, *296*, 776−778.
(29) Noguti, T.; Gō, N. *Biopolymers* **1985**, *24*, 527−546.
(30) Hayward, S.; Gō, N. *Annu. Rev. Phys. Chem.* **1995**, *46*, 223−250.

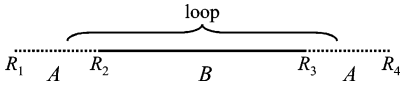


Figure 1. Notation used in the definition of the *measure of relevance* $\rho(n)$.

1. Determination of the *relevant* normal modes necessary to represent binding pocket flexibility (section 2.2).

2. Generation of a de novo ensemble of MRCs by perturbing the structure along the *relevant* modes (section 2.3).

3. Side-chain optimization by complexing the receptor conformational ensemble with known binders, followed by global-energy minimization using a flexible ligand−flexible side chains approach (section 2.4).

4. Receptor ensemble docking (RED) against the generated MRCs using a flexible ligand−grid receptor docking procedure, combination of the screening results, and keeping the best rank per compound (this is the merging−shrinking procedure that has been already described and validated[14]) (section 2.5).

We tested our methodology on the apo (PDB entry 1JLU) and holo (PDB entry 1FMO) structures of cAPK. After an ensemble of diverse backbone conformations was generated through distortion along relevant normal modes, the alternative structures generated from 1JLU and 1FMO were complexed with non-native binders staurosporine and balanol, respectively. It should be noted that these ligands failed to dock using the rigid receptor approach due to structural changes upon ligand binding. Side-chain conformations were then optimized through global-energy minimization. Ligands in their lowest energy state exhibited an RMSD of 1.2 and 0.8 Å (1FMO complexed with balanol and 1JLU complexed with staurosporine, respectively) when compared to their co-crystalized poses. Importantly, these conformations were selected solely on the basis of energy, with an energy-based discrimination $\Delta G > 10$ kcal/mol between the best energy conformations and the first structurally diverse one. The alternative receptor conformations were then used in a flexible ligand−rigid receptor virtual screening, and the results thus obtained were combined with those from the original PDB structures. RMSDs of seeded known binders together with enrichment factors were significantly improved by using this methodology.

2.2. Relevant Modes vs First Low-Frequency Modes. The NMA was performed on crystal structures 1FMO and 1JLU, based on a simplified spring model using C_α atoms only (see section 3 for details). This approximation, originally developed by Tirion[31] and extended by Hinsen[32] and Bahar,[33] was found to reproduce very well slow protein dynamics, while being very fast and less noisy than the all-atom calculation, and has been used in a number of different problems (see ref 34 and references therein). These are the reasons we chose a C_α model rather than an all-atom model. In fact, the computation time was only 12 min on a standard workstation, including the time needed for the computation of the deformability and mobility functions (used in the *measure of relevance*).

To select the smallest number of normal modes that are most "concentrated on", or "relevant" to the region of interest (not necessarily the first ones with lowest frequency), we introduced a *measure of relevance* $\rho(n)$ that expresses in relative terms how much each mode n is active on a specific region. We defined two adjacent regions along the chain: region A surrounds the ends of the loop, and region B includes the central part of the loop (Figure 1). The relevance of mode n on the loop is then defined as

$$\rho(n) = \frac{||\mathrm{d}_n||_{2,\mathrm{A}}}{||\mathrm{d}||_{2,\mathrm{A}}} + \frac{||\mathrm{d} - \mathrm{d}_n||_{2,\mathrm{B}}}{||\mathrm{d}||_{2,\mathrm{B}}} + \frac{||\mathrm{m}_n||_{2,\mathrm{B}}}{||\mathrm{m}||_{2,\mathrm{B}}} + \frac{||\mathrm{m} - \mathrm{m}_n||_{2,\mathrm{A}}}{||\mathrm{m}||_{2,\mathrm{A}}} \quad (1)$$

where m_n is defined in eq 12, d_n and d are defined in eq 14 of section 3, and

$$||\mathrm{d}||_{2,\mathrm{A}} = \left(\sum_{j\in\mathrm{A}} \mathrm{d}(j)^2\right)^{1/2}, \quad \text{etc.} \quad (2)$$

The first term in eq 1 represents modes that bend the chain near the ends of the loop. The second term tends to exclude those modes that distort the central part of the loop. The third term favors those modes that move the central part of the loop the most, while the last term avoids modes that would move the loop *and* the surroundings at the same time.

A few words might be in order here regarding the definition of relevance. While all four terms are reasonable, it is, in principle, not clear that all of them are needed in order to detect relevant modes. Hence, we performed a test consisting of dropping one term at a time and comparing the best-ranking modes—according to the various relevance measures obtained—with the modes furnished by a control. (The control is described in section 3.4.1.) The result of this test was that, except when the trial measure was the sum of the second and third terms, the overlap with the control modes worsened with respect to the overlap of the modes given through eq 1. On the other hand, one can easily think of examples of motions where the presence of the fourth term is essential. (In the test cases considered, this did not happen.) Therefore, one needs to include the first term as well.

In Figure 2, the distributions of mode frequencies are displayed. It should be noted that, due to the simplified model used, high-frequency modes from side-chain fast atomic motion are not present. To understand the two-bell shape of the distribution, we performed an analysis with augmented spring constants (data not shown). When $k \rightarrow \infty$ (no residue bond stretching or bending), the frequency distribution overlapped very well with the left-bell portion of Figure 2. Thus, the left-bell part corresponds to low-frequency distortions of the dihedral angles of the backbone.

Once the normal modes had been ranked by relevance, the first s modes (after which a significant decrease in ρ was observed) were selected (see Table 1 for details). Remarkably, in each of the receptors considered, less than 10 modes have significant contribution in the loops of interest, thus reducing dramatically the number of modes necessary to represent their conformational space. Furthermore, since the first relevant modes are numbers 144 (1FMO) and 166 (1JLU), it is evident that the first lowest-frequency modes cannot represent this type of midrange loop rearrangement, making it necessary to include modes with higher frequency.

2.3. De Novo Generation of an Ensemble of Structurally Diverse Backbone Conformations. With the s relevant modes, we formed a linear combination of them:

$$\mathbf{u} = \sum_{k=1}^{s} \alpha_k \mathbf{u}^{n_k} \quad (3)$$

Table 1. Identification of Relevant Normal Modes Responsible for Loop Rearrangement in the Binding Pocket of cAPK Kinase[a]

receptor	loop (R_1, R_2, R_3, R_4)[b]	relevant normal modes	s	ν	$\mathcal{N}$	κ	d_1 (Å)
1FMO	46, 51, 56, 60	144, 147, 559, 609, 610, 624, 627, 655, 660	9	2	512	5	2
1JLU	322, 325, 327, 330	166, 168, 176, 333, 336	5	3	243	3	3

[a] Shown are the receptor, loop parameters (see Figure 1), the relevant normal modes in the region of interest, the number s of normal modes used to make linear combinations, the number ν of coefficients used for each mode in the linear combination, the number $\mathcal{N}$ of linear combinations, the number κ of conformations (resulting after clustering of the $\mathcal{N}$ conformations) used in the docking, and the per-mode displacement limit. Each structure is deformed along the $\mathcal{N}$ directions, so that the resulting displacement of any C_α atom does not exceed $1.5d_1$. When near-duplicate conformations from these $\mathcal{N}$ are eliminated, κ conformations are obtained, which are then minimized and used as templates in receptor ensemble docking. [b] Actual residue numbers in the PDB structure.

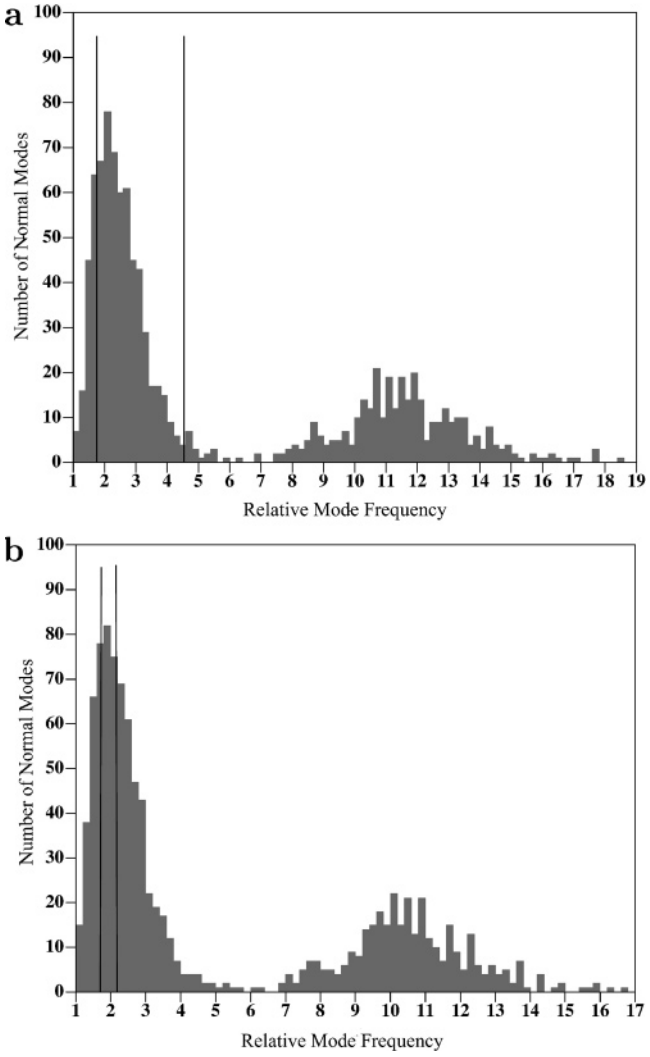


Figure 2. Distribution of relative normal-mode frequencies on crystal structures 1FMO (a) and 1JLU (b) based on a simplified spring model using C_α atoms only. Relative frequencies are calculated with respect to the first lowest-frequency mode. Vertical lines indicate the position of the lowest and highest relevant modes used to describe the rearrangement of the gly-rich loop (1FMO) and the C-term loop (1JLU). A total of nine and five relevant modes respectively were sufficient to represent loop plasticity.

where each of the coefficients α_k takes on values between -1 and 1: $\alpha_k = \{-1, -1 + \Delta\alpha, ..., 1 - \Delta\alpha, 1\}$. The step size $\Delta\alpha$ (the same for all k) is given in terms of the number ν of conformations per mode: $\Delta\alpha = 2/(\nu - 1)$. This gives $\mathcal{N} = \nu^s$ possible linear combinations of the s modes. The number ν was selected in such a way that the total number of combinations $\mathcal{N} \lesssim 1000$.

Prior to making these linear combinations, each mode $\mathbf{u}^{n_k}$ is normalized so that $\max_{1\le j\le N}||\mathbf{u}_j^{n_k}|| = d_1$, where d_1 is a prescribed displacement limit. Now, the resulting linear combinations may exceed the limit d_1, so each of them (call it $\mathbf{u}$) is renormalized in the following way: let $M = \max_{1\le j\le N}||\mathbf{u}_j||$; if $M > d_1$, compute

$$f(M) = \frac{(d_2 - d_1)(1 - e^{-(M-d_1)/(d_2-d_1)}) + d_1}{M} \qquad (4)$$

and replace $\mathbf{u}$ by $f(M)\mathbf{u}$. The new $\mathbf{u}$ satisfies $\max_{1\le j\le N}||\mathbf{u}_j|| < d_2$. The second displacement limit d_2 is set to be 50% larger than d_1.

Next, each linear combination $\mathbf{u}$ is applied to the original structure by displacing all atoms of residue j by $\mathbf{u}_j$. Since deformation is performed in the Cartesian coordinate space, covalent geometry was restored by tethering and minimizing a 3D structure of ideal-geometry residues to each of the $\mathcal{N}$ generated structures. This procedure is consistent with the ECEPP/3 force field, which considers bond lengths and planar angles fixed. Minimization in the Cartesian coordinates space (or in the complete internal coordinates space) using for this purpose another force field would not be consistent. Once covalent geometry was restored, the structures were tethered to an identical copy and further energy-minimized in vacuo to avoid residue inter-locking and to relieve energy strain, while keeping to a minimum the deviations from the original $\mathcal{N}$ conformations. This was done in succesive steps where the tether weight was reduced from 50 kcal/mol Å^2 to 0, thus avoiding excesive deformation from highly strained regions.

Structures within each ensemble were then superimposed and clustered according to their backbone RMSD in the region of interest using a threshold of 0.8 Å, giving an acceptable balance between number of conformations and their structural diversity. After clustering, the ensemble was reduced from 512 to 5 conformations for 1FMO and from 243 to 3 conformations for 1JLU. The final numbers κ of conformations to be used in the docking, as well as other parameters, are listed in Table 1. Figure 3 displays the representative structures thus generated.

It should be clear now that few relevant modes ($\lesssim$10) are the natural way to represent intermediate-scale loop movements (for which selecting the first lowest-frequency modes will fail), thus avoiding the geometric explosion in the number of possible combinations of normal eigenvectors. It is also evident that the noise in the energy calculation of the system will be significantly reduced, since distortion was mostly localized in the region of interest.

2.4. Side-Chain Optimization within the Ligand-Binding Pocket. The optimized structures from the previous step (five

(31) Tirion, M. M. *Phys. Rev. Lett.* **1996**, *77*, 1905−1908.
(32) Hinsen, K. *Proteins: Struct., Funct. Bioinf.* **1998**, *33*, 417−429.
(33) Bahar, I.; Atilgan, A. R.; Erman, B. *Fold. Des.* **1997**, *2*, 173−181.
(34) Delarue, M.; Dumas, P. *Proc. Natl. Acad. Sci.* **2004**, *101*, 6957−6962.

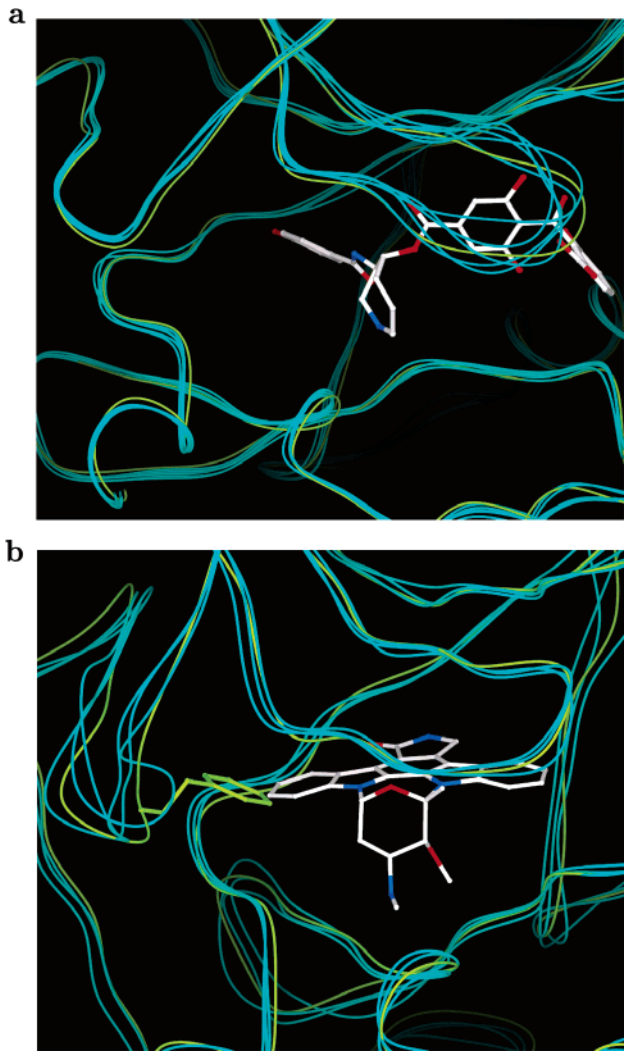


Figure 3. Representative multiple receptor conformations of cAPK structures (a) 1FMO and (b) 1JLU generated by perturbing the structure along a combination of relevant normal modes. Ligands that fail to dock under the rigid docking approximation to the source structures are displayed with white carbons. (a) Balanol (PDB 1BX6) fails to dock to 1FMO due to clashes with the gly-rich loop. (b) Staurosporine (PDB 1STC) fails to dock due to clashes with displayed F327. Note the ensemble of diverse backbone structures in the gly-rich loop (a) and in the loop bearing F327 (b). Color code: magenta, modeled structures; green yellow, source structures (1FMO and 1JLU); blue, nitrogen; red, oxygen.

for 1FMO and three for 1JLU) were complexed with non-co-crystalized ligands [1FMO with balanol (complex **A**) and 1JLU with staurosporine (Complex **B**)], and the global energy was minimized according to the double-energy scheme biased probability Monte Carlo (BPMC) method (see section 3.2). A solvation energy contribution was included by means of atomic solvation parameters,[35] together with an entropy term. This choice of energy terms was found to perform satisfactorily in earlier works.[7,14,36] Three atom layers of increasing size were considered: zone I, the ligand + side chains within 5.5 Å of the ligand-binding pocket; zone II, including side chains within 4.5 Å of zone I; and zone III, the rest of the receptor. Zone III was kept rigid during the global-energy optimization. Torsional variables associated with side chains in zone II were included in the local energy minimization step of the global-energy optimization, while only those associated with zone I were sampled in the Monte Carlo procedure. Zone II serves as a "buffer" to propagate disturbances in zone I. Since no significant rearrangement is expected within zone III, it is kept rigid during the Monte Carlo step for computing time reasons and relaxed through minimization at the end (see below). Initially, the positional and torsional variables of the ligand were randomly perturbed, and 10 independent parallel simulations were performed for each of the complexes. An upper limit of 3000 steps was set for the local energy minimization stage. The global-energy optimization stage was stopped after 3.5 million energy evaluations as used in similar studies.[14] During the simulations, conformational stacks of low-energy states[37] were collected and then merged. To further optimize ligand−receptor contacts, a relaxation step was then performed by a full local energy minimization of the complexes represented in the conformational stack.

2.4.1. Energy Re-evaluation of the Conformational Stack. Although the use of the atomic solvation approximation is fast, and it was found in this work and in others[7,14] that correct poses along with mis-docked ones are present in the conformational stack, the approximation is too crude to generally assign the lowest energy to the correct pose. To achieve this purpose, the electrostatic energy was recalculated by solving the Poisson equation and using the boundary element algorithm, along with an energy term proportional to the solvent-accessible surface to account for the nonpolar contribution to the solvation energy. This combination of energy terms was found to perform successfully with the ECEPP/3 force field.[7,14,38]

However, the total free energy of the system is not a good measure to discriminate the correct geometry of the complex (data not shown). The main reason for this is that although the normal modes along which the structure is distorted have high relevance in the regions of interest, they still have some residual contribution in regions far from the binding pocket. Moreover, since these regions are not global-energy-optimized, energy evaluation in those regions introduces noise that may preclude the correct complex geometry from being identified solely on the basis of total energy.

To avoid this energy noise and to rank structures solely on the basis of energy calculation, we considered the projection G_{proj} of the total free energy of our region of interest (zone II). Since the unbound state (free ligand and free receptor) is common to every conformation in the stack, we consider only the bound state in calculating G_{proj}. On the basis of the total free-energy expression, the projection of the free energy on zone II is naturally defined as

$$G_{proj} = E_{II}^{inter} + E_{II}^{tor} + G_{II}^{SASA} + G_{II}^{elec} - TS_{II} \qquad (5)$$

The first term on the right-hand side represents the interaction energy (van der Waals and hydrogen bond) of zone II with the rest of the system (including its self-energy). The second term represents the torsional energy within zone II, while the last term represents the configurational entropy of zone II. The third

(35) Abagyan, R. Protein structure prediction by global energy optimization. In *Computer simulation of biomolecular systems*, Vol. 3; van Gunsteren, W., Weiner, P., Wilkinson, A., Eds.; Kluwer/Escom: Dordrecht, The Netherlands, 1997.
(36) Cavasotto, C. N.; Orry, A. J. W.; Abagyan, R. A. *Proteins: Struct., Funct. Bioinf.* **2003**, *51*, 423−433.
(37) Abagyan, R.; Argos, P. *J. Mol. Biol.* **1992**, *225*, 519−532.
(38) Totrov, M.; Abagyan, R. *Biopolymers* **2001**, *60*, 124−133.

Table 2. Best-Energy Conformations of Protein−Ligand Complexes Generated through Distorsions along a Combination of Relevant Normal Modes and Fully Flexible Ligand Docking

complex	receptor	ligand	ligand RMSD (Å)[a]	ΔG^*_{proj} (kcal/mol)[b]
A	1FMO	balanol	1.2	0.0
			2.2	10.2
			2.5	18.8
B	1JLU	staurosporine	0.8	0.0
			3.5	11.6

[a] RMSD values are calculated with respect to the ligand native structure (for balanol, 1BX6; for staurosporine, 1STC). [b] ΔG_{proj} is the free-energy difference relative to the best-energy conformation.

term is calculated as $G_{II}^{SASA} = \gamma(SASA_{II})$, where $\gamma = 0.012$ kcal/(mol Å^2) is the surface tension and $SASA_{II}$ is the contribution of residues and ligand from zone II to the total solvent-accessible surface area (SASA).

To derive G_{II}^{elec}, we first consider the total electrostatic energy, expressed as

$$G^{elec} = \frac{1}{2}\sum_{i\neq j}\frac{q_i q_j}{\epsilon_{int}|\vec{r}_i - \vec{r}_j|} + \frac{1}{2}\sum_{i,j} q_i q_j \mathcal{G}^{RF}(\vec{r}_i,\vec{r}_j) \qquad (6)$$

where q_i and q_j are the charges associated with atomic centers at $\vec{r}_i$,$\vec{r}_j$, and $\mathcal{G}^{RF}(\vec{r}_i,\vec{r}_j)$ is the reaction field Green function matrix at $\vec{r}_i$,$\vec{r}_j$. The $\mathcal{G}^{RF}$ is independent of the charges and depends only on the geometry of the system and the dielectric constant at each point of space. It represents the electric potential due to the induced charges on the solvent boundary. Moreover, $\mathcal{G}^{RF}(\vec{r}_i,\vec{r}_j)$ corresponds to the value of the electric potential due to the reaction field at site $\vec{r}_i$, provided a unit charge is located at $\vec{r}_j$. By grouping indices in eq 6, it is clearly seen that G^{elec} can be split into four terms:

$$G^{elec} = G^{elec}_{II,II} + G^{elec}_{rest,rest} + G^{elec}_{II,rest} + G^{elec}_{rest,II} \qquad (7)$$

where "rest" refers to zone III. The first two terms represent the electrostatic self-energy of zone II and zone III, while the last two terms represent the interaction energy between zones II and III. These two terms have the same value, due to the symmetry of the Green function. We define the projection of G^{elec} onto zone II as

$$G^{elec}_{II} = G^{elec}_{II,II} + G^{elec}_{II,rest}$$

$$= \frac{1}{2}(G^{elec} + G^{elec}_{II,II} - G^{elec}_{rest,rest}) \qquad (8)$$

where eq 7 has been used.

The three G^{elec} terms on the right-hand side of the second eq 8 were calculated considering the complete set of charges, setting those of zone III to zero, and setting those of zone II to zero, respectively.

In Table 2 we show the RMSD of the top-ranking solutions for the systems considered, along with the relative free energy with respect to the best-ranking complex. It is clearly seen that in the cases considered, G_{proj} constitutes a satisfactory measure to discriminate the correct complex geometry solely on the basis of energy, with an energy gap >10 kcal/mol between the correct pose and the first misdocked one.

2.5. Small-Scale Virtual Screening Using RED: Improvement of RMSD Values and Enrichment Factors. Receptor structures **A** and **B** (extracted from complexes **A** and **B**, respectively) were used along with the experimental PDB structures to perform a small-scale virtual screening. A 1000-compound library of randomly selected molecules seeded with known cAPK binders was docked and scored to the MRCs using a flexible ligand−rigid receptor docking algorithm.[39−41] For each pair of receptor structures (1FMO and receptor **A**; 1JLU and receptor **B**), the screening results were sorted by rank and merged, and then the best rank for each compound was kept (merging−shrinking procedure). Results are reported in Tables 3 and 4.

Table 3. RMSD Values of Flexible Ligand−Rigid Receptor Docking on Multiple Receptor Conformations Using Experimental Structures and Those Generated through Our Relevant Normal-Mode-Based Algorithm[a]

compound	1FMO[b]	receptor **A**	1JLU[b]	receptor **B**
adenosine	0.4	0.6	0.7	4.0
balanol	8.5	1.1	1.6	4.0
staurosporine	9.8	7.0	10.1	0.8
H7	0.8	0.8	0.5	3.4
H8	0.8	1.1	0.8	3.5
H89	9.2	1.1	10.2	1.5

[a] RMSD values refer to the best-score solution and are reported in Å. [b] Taken from ref 14.

Table 4. Improvements in the Enrichment Factors (EF) Resulting from Performing a Small-Scale Virtual Screening on Multiple Receptor Conformations Using Experimental Structures (1FMO and 1JLU) and Those Generated through Our Relevant Normal-Mode-Based Algorithm (Receptors A and B)

receptor	EF top 1%	EF top 2%	EF top 10%	% binders with RMSD < 1.5 Å[a]
1FMO	16.7	16.7	5.0	50.0
receptor **A**	33.3	16.7	8.3	83.3
merged set	50.0	25.0	8.3	83.3
1JLU	16.7	8.3	3.3	66.7
receptor **B**	16.7	8.3	3.3	50.0
merged set	33.3	16.7	5.0	100.0

[a] RMSD values refer to the best-score solution.

In receptor **A** (generated from complex **A**), balanol had the best score. Remarkably, compound H89, which was not able to dock to 1FMO, appeared within the 0.5% top-ranking compounds, showing that this methodology does not necessarily create a custom-fit pocket. All of the compounds with the exception of staurosporine were docked within 1.5 Å, as shown in Table 3. The ligand docking RMSD values always refer to the top-scoring solution for that particular ligand. In Figure 4 we compare docked balanol and H89 with their co-crystalized poses. The enrichment factors of the merged set are larger than those of the individual conformations for the three cases considered (top 1%, 2%, and 10% ranking). The enrichment factors for receptor **A** structure are sligthly worse than those obtained with the IFREDA method,[14] but the RMSD values are comparable. Since the two receptor structures are similar but

(39) Abagyan, R.; Totrov, M.; Kuznetsov, D. *J. Comput. Chem.* **1994**, *15*, 488−506.
(40) Totrov, M.; Abagyan, R. Derivation of sensitive discrimination potential for virtual ligand screening. In *RECOMB '99: Proceedings of the Third Annual International Conference on Computational Molecular Biology*; Istrail, S., Pevzner, P., Waterman, M., Eds.; Associaton for Computer Machinery: New York/Lyon, France, 1999.
(41) Totrov, M.; Abagyan, R. Protein−ligand docking as an energy optimization problem. In *Drug−receptor thermodynamics: Introduction and experimental applications*; Raffa, R., Ed.; John Wiley and Sons: New York, 2001.

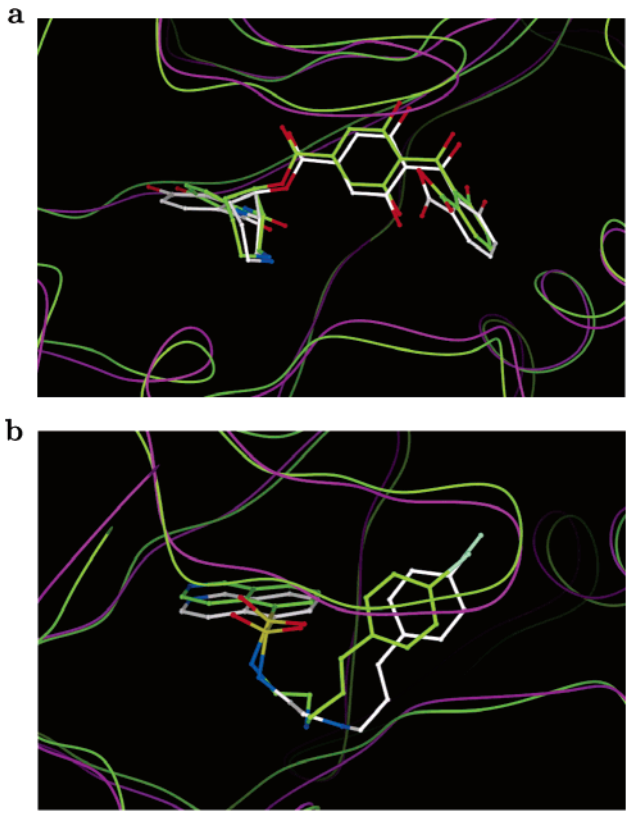


Figure 4. Flexible ligand−rigid receptor docking of compounds balanol and H89 to receptor **A** generated from 1FMO by (i) perturbation of the structure along relevant normal modes and (ii) fully flexible docking of balanol into the ligand-binding pocket. Receptor **A** is displayed in magenta superimposed onto the native receptors, displayed in green-yellow, of balanol (1BX6, a) and of H89 (1YDT, b). Carbon atoms of docked compounds are displayed in white. Color code: blue, nitrogen; red, oxygen; yellow, sulfur.

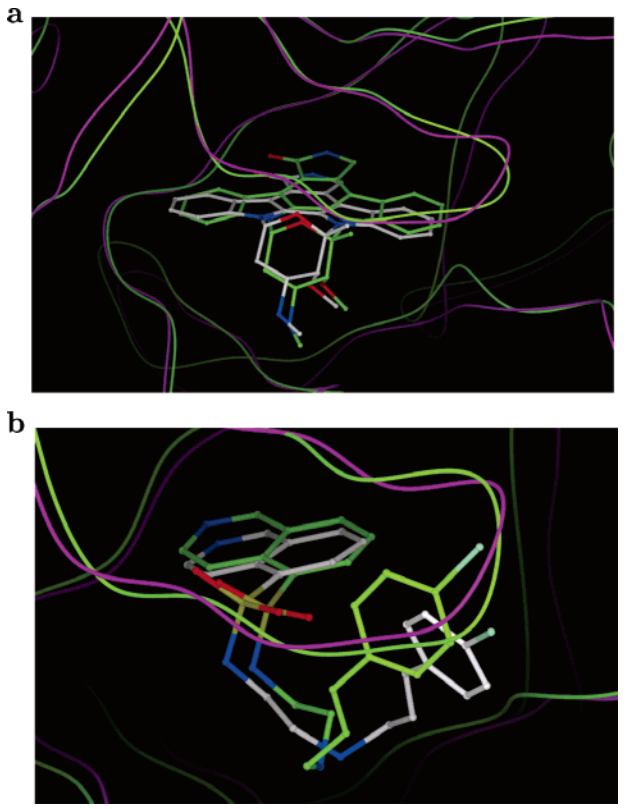


Figure 5. Flexible ligand−rigid receptor docking of compounds staurosporine and H89 to receptor **B** generated from 1JLU by (i) perturbation of the structure along relevant normal modes and (ii) fully flexible docking of staurosporine into the ligand-binding pocket. Receptor **B** is displayed in magenta superimposed onto the native receptors, displayed in green yellow, of staurosporine (1STC, a) and of H89 (1YDT, b). Carbon atoms of docked compounds are displayed in white. Color code: blue, nitrogen; red, oxygen; yellow, sulfur.

not identical, the differences in enrichment factors could be due to the scoring function.

Staurosporine and H89 cannot be docked to the apo structure 1JLU using the rigid receptor approach. However, in the alternative receptor conformation generated with the normal-mode-based procedure, and by docking staurosporine (receptor **B**), both compounds dock with an RMSD < 1.2 Å, while staurosporine has the best score (see Figure 5). Considering the merged set, it is interesting that 100% of the compounds are docked within 1.5 Å. The enrichment factors are comparable to those obtained with IFREDA,[14] and in the three cases considered, there is an improvement with respect to those of the individual receptors. However, the fact that all of the compounds are docked correctly but only half of them appear in the top 10% of the ranking list could be due to the scoring function. It should also be pointed out that, in the cases displayed in Table 4, there is no "dilution" of the top hits after the merging and shrinking procedure, since the enrichment factors of the merged set are larger than or equal to those corresponding to the individual structures.

3. Computational Methods

3.1. Complex Preparation. Receptor structures of 1FMO and 1JLU were taken from the Protein Data Bank, and hydrogens and missing heavy atoms were added. Peptide PKI-(5−24) was removed from 1FMO. The system was then subjected to a local minimization step. Asn and Gln residues were flipped whenever necessary to optimize hydrogen bonding. Polar hydrogens in the vicinity of the ligand-binding pocket were also optimized.

The 3D structures of cAPK ligands balanol (1BX6), staurosporine (1STC), adenosine (1FMO), and isoquinolinesulfonamides analogues H7 (1YDR), H8 (1YDS), and H89 (1YDT) were taken from their co-crystalized native structures (specified in parentheses), and their correct stereochemistry and formal charges were assigned. The staurosporine amino group and terminal amino substituents in H7, H8, and H89 were regarded as protonated, in agreement with an environment of pH = 7.4. Ligands were assigned MMFF[42] atom types and then subjected to global-energy optimization using the MMFF energy terms.

3.2. Energy Evaluation and Minimization. The molecular system is described in the internal coordinate space using a modified version of the ECEPP/3 force field.[43] Entropy[44] and solvation energy terms were added to the in vacuo energy. The stochastic global-energy optimization method consists of (i) random conformational change of free variables according to a predefined probability distribution as described in the biased probability Monte Carlo (BPMC) method;[44] (ii) local energy minimization of analytically differentiable terms, followed by total energy re-evaluation including nondifferentiable terms, like entropy and solvation energy; and (iii) acceptance or rejection on the basis of the Metropolis criterion[45] applied to the total energy. The temperature of the simulations was set to T = 600 K. The water

(42) Halgren, T. *J. Comput. Chem.* **1995**, *17*, 490−641.
(43) Némethy, G.; Gibson, K. D.; Palmer, K. A.; Yoon, C.; Paterlini, M. G.; Zagari, A.; Rumsey, S.; Scheraga, H. A. *J. Phys. Chem.* **1992**, *96*, 6472−6484.
(44) Abagyan, R.; Totrov, M. *J. Mol. Biol.* **1994**, *235*, 983−1002.
(45) Metropolis, N.; Rosenbluth, A.; Rosenbluth, N.; Teller, A.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

dielectric constant was set to $\epsilon_{water} = 78.5$. The internal dielectric constant was $\epsilon_{int} = 2R$ during the global-energy optimization process and $\epsilon_{int} = 2$ for electrostatic calculations when solving the Poisson equation.

3.3. Library Preparation and Flexible Ligand−Rigid Receptor Docking. A random chemical library was extracted from the Diverse Set of ChemBridge (ChemBridge, Inc., San Diego, CA). These compounds were prepared in a similar way as the native ligands. Key chemical descriptors of the random library (molecular weight, number of rotatable bonds, number of hydrogen bond donors, and number of hydrogen bond acceptors) overlap with those of cAPK ligands.[14] The random library thus extracted was seeded with cAPK ligands to get a docking library of 1000 compounds.

The flexible ligand−rigid receptor docking algorithm, as implemented in ICM,[39−41] consists of (i) representation of the receptor with six potential energy maps (three for van der Waals, electrostatic, hydrogen bond, and hydrophobic); (ii) global-energy minimization of the flexible ligand in the field of the receptor, so that both the self-energy of the ligand and its interaction with the receptor are optimized through this procedure; and (iii) assignment of a docking score to the best-energy conformation of each compound. The docking score takes into account the fit of the ligand within the binding pocket and additional terms for entropy loss, desolvation, hydrogen bonding, and hydrophobic effect.[40,41] Each virtual screening experiment was repeated four times, and the best score for each compound among the four was kept. The scoring function was not optimized for protein kinases. Computing time for ligand docking and scoring was ∼1−2 min on a 700 MHz processor (1 Mb RAM dual-processor node).

The enrichment factor (EF) for a library built with the n top compounds of the ranked library is defined as

$$\mathrm{EF}(n) = \frac{\mathrm{Hits}_n}{N_n} \Big/ \frac{\mathrm{Hits}_{\mathrm{total}}}{N_{\mathrm{total}}} \tag{9}$$

and expresses the relative change in the probability of finding a ligand in the focused library when compared to a random pick from the complete library.

3.4. Normal-Mode Analysis. 3.4.1. Protein Vibrational Analysis. We used a reduced C_α harmonic network protein model,[31] whereby the C_α atoms of the molecule are interconnected with springs defined as follows:

$$C_{ij} = 225\delta_{i,j-1} + \exp\left(-\frac{E_{ij} - E_1^{IJ}}{kT}\right) \qquad (1 \le i < j \le N) \tag{10}$$

where δ_{ab} is the Kronecker delta, i and j denote residue numbers, N is the total number of residues, k is the Boltzmann constant, T is the temperature (fixed $T = 600$ K), E_{ij} is the interaction energy between residues i and j (which was computed using the ICM algorithm[39]), I and J denote the residue types of residues i and j, respectively ($1 \le I, J \le 20$), and

$$E_1^{IJ} = \bar{E}^{IJ} - 1.65\sigma^{IJ} \tag{11}$$

where $\bar{E}^{IJ}$ and σ^{IJ} denote the mean and standard deviation of the distribution of nonzero energy values corresponding to the pair of types I,J.

The masses m_i of the pseudo-atoms, located at the C_α atom positions, were set to the total mass of the corresponding residues.

The units in eq 10 are kcal/(mol Å^2). The value $C = 225$ for the force constant between adjacent residues was obtained by concatenating springs corresponding to the C_α−C, C−N, and N−C_α bonds, whose force constants were taken from the MMFF force field.[42] However, the influence of varying C is minimal. Indeed, for values of C ranging from 25 to 10^6, there is a very good overlap (≥60%) between the 10 modes with highest relevance to the interesting loop and the 10 modes furnished by the control. The control consists of defining a $3N$-dimensional vector that approximately describes the motion of the interesting loop according to experimental data, expanding that vector in the normal-mode basis, and picking out the 10 modes with highest coefficients in the expansion. For smaller values of C the overlap decreases; for instance, for $C = 10$ the overlap is 40%, and for $C = 1$ the overlap is 30%. This is natural, since such low values make the model unrealistically soft; C should always be ≫1 since the distance between adjacent C_α atoms is constrained by a high energy barrier.

The resulting eigenvalue problem was solved by means of the FORTRAN subroutine DSYEVR in the LAPACK linear algebra library.[46]

3.4.2. Mobility and Deformability Functions. The mobility function is given by the classical formula for atomic fluctuations[47] (excluding the factor kT):

$$\mathrm{m}^2 = \sum_{n=7}^{3N}\left(\frac{||\mathbf{u}^n||}{\omega_n}\right)^2 \equiv \sum_{n=7}^{3N}\mathrm{m}_n^2 \tag{12}$$

where the ω_n are the vibrational frequencies, related to the eigenvalues by $\lambda_n = \omega_n^2$. The first six modes correspond to rigid motions, and hence have null eigenvalues.

As in a previous work,[48] we view the normal modes as vector fields over the molecule. For any such vector field $\mathbf{u}$, we can define the corresponding "conformal tensor" $S_\mathbf{u}$ with components

$$S_{kl} = \frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right) - \frac{1}{3}\delta_{kl}\,\mathrm{div}\,\mathbf{u} \qquad (1 \le k, l \le 3) \tag{13}$$

where div $\mathbf{u}$ denotes the *divergence* of $\mathbf{u}$: $\mathrm{div}\,\mathbf{u} = \sum_{k=1}^{3} \partial u_k/\partial x_k$.

Note: u_1, u_2, and u_3 stand for the three components of the vector field $\mathbf{u}$ as functions of the spatial coordinates x_1, x_2, and x_3. This should not be confused with an expression such as $\mathbf{u}_i$, which means "$\mathbf{u}$ at the ith residue" (a three-dimensional vector).

The tensor $S_\mathbf{u}$ describes how the vector field $\mathbf{u}$ affects (locally) the shape of the molecule.[48] And it does so in a "relative" way, since it involves only derivatives of $\mathbf{u}$. Thus, we call it the *relative conformal tensor*. Accordingly, the deformability measure given previously[48] is called here *relative deformability*.

In this work we want to define and use an "absolute" deformability measure, which takes into account not only the derivatives of the normal modes but also their amplitudes. The rationale for this is that high-frequency/low-amplitude modes should have a small contribution to the total deformability measure. Therefore, the (relative) scalar deformation function[48] corresponding to mode n (after scaling it by its thermal amplitude ω_n^{-1}), $||S_{\mathbf{u}^n/\omega_n}||$ (the "derivative" of the normal mode), should be multiplied by the wavelength l_n of that mode in order to get the corresponding "absolute" scalar deformation function. The wavelength is not known, but assuming, as a first-order approximation, that the speed of the vibrations is independent of their frequency, we can take (disregarding constant factors) $l_n = \omega_n^{-1}$. Thus, our *absolute deformability function* d:$M \to$ R is defined as

$$\mathrm{d}^2 = \sum_{n=7}^{3N}\left(\frac{||S_{\mathbf{u}^n}||}{\lambda_n}\right)^2 \equiv \sum_{n=7}^{3N}\mathrm{d}_n^2 \tag{14}$$

The method for the numerical computation of the partial derivatives has been previously described.[48]

4. Conclusions

Ligand binding to a receptor is currently best described as a selection process of partially fitting structures (conformer

(46) Anderson, E.; et al. *LAPACK Users' Guide*, 3rd ed.; Society for Industrial and Applied Mathematics: Philadelphia, PA, 1999.
(47) Brooks, B. R.; Janežič, D.; Karplus, M. *J. Comput. Chem.* **1995**, *16*, 1522−1542.
(48) Kovacs, J. A.; Chacón, P.; Abagyan, R. *Proteins: Struct., Funct., Bioinf.* **2004**, *56*, 661−668.

selection stage) followed by minor structural changes upon ligand binding (induced-fit stage). This concept inspired our methodology to incorporate receptor flexibility in ligand docking and virtual screening through a normal-mode-based algorithm to generate multiple receptor backbone conformations, followed by a flexible ligand−flexible side chain docking of non-native ligands. The alternative receptor conformations thus generated could be used for virtual screening using the receptor ensemble docking (RED) approach.

However, intermediate-scale loop motions like those found in the binding pocket of protein kinases cannot be represented by picking the first lowest-energy modes. Furthermore, adequate representation of the conformational space grows geometrically with the number of modes considered. To overcome these limitations, we introduced a measure of *relevance* that expresses how active a given mode is on a region of interest. In this way, we showed that very few normal modes ($\lesssim$10) are necessary to describe loop flexibility in protein kinases. Remarkably, the relevant modes are not those with first lowest frequency, but are among the low-frequency modes (see Figure 2). We validated our methodology on holo and apo structures of cAPK protein kinase, where loop rearrangement of $\sim$2 Å takes place. Alternative receptor conformations were generated by perturbing the structures along a combination of relevant modes, followed by a flexible ligand−flexible side chain docking of known non-native ligands in order to optimize side-chain conformations. In a second stage, the receptor conformations thus obtained were used as starting points for a virtual screening using RED.

The lowest-energy de novo ligand−receptor complexes generated with our procedure from 1FMO and balanol (complex **A**), and from 1JLU and staurosporine (complex **B**), exhibited a ligand RMSD within 1.2 Å compared to the experimental structures 1BX6 and 1STC, respectively, while their ΔG was lower than 10 kcal/mol with respect to the first misdocked structure (see Table 2). It should be emphasized that balanol and staurosporine cannot dock to 1FMO and 1JLU, respectively, in the rigid receptor approach. The small-scale virtual screening performed against the multiple receptor conformations 1FMO and receptor **A** (from complex **A**), 1JLU, and receptor **B** (from complex **B**) showed improved enrichment factors when compared to those obtained using a single receptor conformation. Altogether, this indicates that the structural diversity of the pocket was correctly represented. Moreover, most of the known ligands were able to dock within 1.5 Å to each of the multiple receptor ensembles, showing that our procedure does not necessarily make custom-fit pockets that accommodate only those ligands used in the generation of the ensemble structures. Thus, the alternative backbone conformations generated represent actual states of the system.

This methodology is applicable to both holo and apo structures. The use of known binders to optimize the position of side chains has the advantage of reproducing likely binding pockets, introducing however a limitation in the applicability of the method, since it cannot be used as it stands for an ab initio protein de-orphanization. This opens a large space for improvement of our method. Although side-chain flexibility and local motions might be not always uncoupled from the docking process, in this paper we do take into account partial coupling through the side-chain optimization stage with known non-native binders. We plan to couple continuous changes along relevant modes with the docking step in the future.

Virtual screening using RED is a new avenue toward the consideration of protein flexibility in computer-aided drug discovery. In many cases, receptor structural diversity can be successfully represented with a few receptor conformations, either experimentally or in silico generated. Screening against a few structures is computationally affordable, using actual receptor conformations, with the possibility of incorporating diversity from both side-chain and loop rearrangements.

Acknowledgment. We thank Dr. Andrew Orry for helpful discussions.

Supporting Information Available: Complete list of authors for refs 7 and 46. This material is available free of charge via the Internet at http://pubs.acs.org.

JA042260C