Computational Chemistry

DOI: 10.1002/anie.201103264

# Conformational Selection versus Induced Fit in Kinases: The Case of PI3K-γ**

*Marco D'Abramo, Obdulia Rabal, Julen Oyarzabal,* and Francesco Luigi Gervasio**

*In memory of Jacques Monod*

Molecular recognition is crucial for a multitude of fundamental biological processes, for example enzyme activation and inhibition,[1,2] and protein folding.[3,4] The increasing availability of high-resolution structures of ligands bound to receptors has changed our understanding of molecular recognition from a static concept, in which interactions are considered to be a rigid lock and key, to a dynamic one, in which both the ligand and its target can assume different, yet not always complementary, shapes. Within this dynamic framework, two limiting mechanisms have been proposed: conformational selection[5,6] and induced fit.[7] The first theory implies that there is a significant overlap between the conformational space that is occupied by the bound and unbound forms of the target, and that the role of the ligand is to stabilize certain conformations that are accessible to the unbound form. In the induced-fit hypothesis, the ligand induces the protein to explore regions of the conformational space that are virtually inaccessible to the unbound form.

Whereas direct observation of target dynamics has found a significant role for conformational selection,[8,9] induced-fit effects have been shown to be important in the binding pocket of a ionotropic glutamate receptor.[10] The idea that both mechanisms play a role in ligand/target binding is gaining new ground.[11] Similar to enzymatic catalysis, the interplay between the two mechanisms may be regulated by the different time scales that are involved and by allosteric effects.[12,13] Herein, we have tested the various hypotheses in the pharmacologically relevant case of the phosphoinositide 3-kinases (PI3K). We performed four long, all-atom molecular dynamics (MD) simulations (each lasting 1 μs or more), extensive bias-exchange metadynamics calculations,[14] and analyzed the conformational space that is spanned by the available crystal structures. The finding that the PI3K-signaling pathway is often deregulated in cancer has fueled an increasing interest in designing selective and potent inhibitors of PI3K kinases[15,16]

There are four different isoforms of class I PI3Ks. These isoforms have distinct substrate specificities and different roles in discrete cellular processes. Given their high sequence retention (95% or more) and a nearly identical adenosine triphosphate (ATP) binding pocket, it is not surprising that a potent and selective inhibitor for only one of the isoforms had remained elusive for a long time. The first selective inhibitor for the delta isoform was reported in 2003.[17] Recently, the crystal structures of various PI3K isoforms in complexes with different inhibitors were published. These structures showed a significant conformational variability in the binding cavity.[18–20] This variability is particularly prominent in the oncogenic target PI3K-γ, which makes it a valid and interesting system with which to explore conformational selection versus the induced-fit hypothesis (Figure 1; see also Figures S1 and S2 in the Supporting Information).

We explored the conformational space of the apo form (1E8Y) and the L64- cocrystallized holo form (3IBE) of PI3K-γ by performing two fully solvated MD simulations that lasted 1.5 μs and 1 μs, respectively. The dynamics were analyzed by principal component analysis (PCA) of the full MD trajectory and considered first the Cα coordinates, and then only the atoms defining the binding pocket (Table S2 in the Supporting Information). The global Cα PCA vectors of the apo and the holo forms are similar, as shown by the overlap of their covariance matrices (ca. 0.4). The first two vectors describe functionally relevant, hinge-like motions and the relative rotations of the N-terminal and C-terminal lobe. The local binding pocket vectors describe the opening and closing of the cavity (Figure S3 in the Supporting Information), but the overlap of the covariance matrices of the apo and the holo forms is smaller (ca. 0.2). The free energy surfaces (FESs) shown in Figure 2 are obtained by projecting MD trajectories of the apo form (Figure 2a,b) and the holo form (Figure 2c,d) on the first two Cα coordinates and binding pocket vectors of the apo-form trajectory.

The convergence of the FESs of the apo form and the effect of the initial structure was checked by performing an additional 1 μs long MD simulation. This simulation started from a different crystal structure (3DBS) after the removal of the GD9 ligand, and the new FES was compared with the

[*] Dr. M. D'Abramo, Dr. F. L. Gervasio
Structural Biology and Biocomputing Programme
Spanish National Cancer Research Center (CNIO)
C/Melchor Fernandez Almagro 3, 28029 Madrid (Spain)
E-mail: flgervasio@cnio.es
Dr. O. Rabal, Dr. J. Oyarzabal
Experimental Therapeutics Programme
Spanish National Cancer Research Center (CNIO)
C/Melchor Fernandez Almagro 3, 28029 Madrid (Spain)
E-mail: julenoyarzabal@unav.es
Dr. O. Rabal, Dr. J. Oyarzabal
Current address: Small Molecule Discovery
Center for Applied Medical Research (CIMA)
University of Navarra, Pamplona (Spain)

[**] We acknowledge support from a Spanish Science and Innovation (MICINN) grant (BIO2010-20166, "AlteredDynamics") and the Barcelona Supercomputing Center for supercomputer time.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201103264.

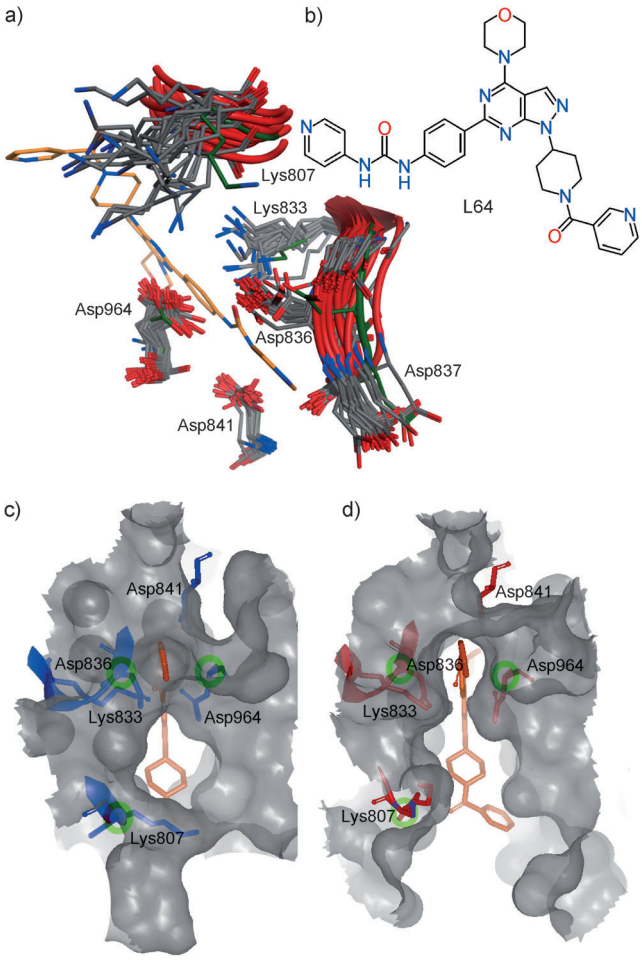


***Figure 1.*** a) Representation of the variability of the binding pocket. The lateral chains close to the binding pocket of the 26 available crystal structures (Table S1 in the Supporting Information) have been superimposed. The position of L64, the longest cocrystallized inhibitor available, is shown in orange. b) Chemical structure of L64. c) The position of L64 within the PI3K-γ cavity in the apo-form structure (1E8Y). The steric clash of L64 with the lateral chains is evident. d) The position of L64 in the cavity of the PI3K-γ holo-form structure (3IBE).

original (Figures S4 and S5 in the Supporting Information). In terms of the global displacements of the apo form, the FES has a broad minimum that is centered at (0,0) (Figure 2a). To link our results with the available experimental data, we projected the 26 crystal structures on these essential subspaces (Figure 2, gray squares). The spread of these projections on the surface hints at the dynamic variability (along the represented PCA vectors) of the available crystal structures. In both of the FESs, all of the crystal structures fall into partially populated regions, which shows that the conformations that were sampled by MD are compatible with those that are spanned by the crystal structures. The crystal structure of the apo form projects within the main minimum of the MD trajectory of the apo form (Figure 2a white rectangle). By comparing the FES of the apo form with that of the holo form (Figure 2a and Figure 2c, respectively), it is apparent that the position and depth of the main minimum is approximately the same. The holo form explores conformations that are slightly more closed, and for this reason the crystal structures of the apo and the holo forms are found in the energy minimum (Figure 2c).

In the case of the cavity subspace (that is, the binding pocket space, Figure 2b,d), the small overlap in the covariance matrices suggests that the motions of the apo form are somewhat different from those of the holo form. The topology of the two FES landscapes is also different. The single minimum that is present in the apo form is split into two main minima and shifted up when the ligand is in the cavity. These results indicate that a conformational selection mechanism operates at the global level, as the PCA vectors are similar and the shape of the FES is almost unaffected by the presence of the ligand in the Cα space. In the binding pocket space, relevant changes in the FES topology and in the direction of the PCA vectors occur. As a result of these changes, some conformations (notably the crystal structure of the holo form) seem to be less accessible to the apo form. However, an increase in flexibility in the apo form could also have a similar effect on the FES. We calculated the total and per-residue root mean square fluctuations (RMSF) of the two forms (Table S3 and Figure S6 in the Supporting Information) and found the two to be equally flexible. These results raise the question of whether the changes in the accessible FES landscape are in accordance with a local induced-fit effect on the binding pocket. To answer this question we analyzed the structural rearrangements in more depth. In the crystal structures, the changes are mainly a result of the reorientation of lateral chains or the backbone in the binding pocket and lead to a significant increase in the average cavity volume (Figure 1a). In Figure 3a the average volume of the cavity that was extracted from the MD simulation of the apo form is projected on the FES of the binding pocket. It is clear that the conformations of the binding pocket in the main energy minimum of the apo form have small volumes. The average cavity volume in the MD simulations of the holo and the apo forms is 132 $A^3$ and 67 $A^3$, respectively. An analysis of the root-mean-square deviations (RMSD) values of the apo and the holo conformations that were obtained from the MD simulations that began from the average holo-form structure (Figure S7 in the Supporting Information) revealed one cluster in the case of the Cα space and two clusters in the case of the binding pocket space. Taken together, these observations might be compatible with a local ligand-induced conformational-shift effect.

To test this hypothesis, different metrics that capture critical differences in the binding pocket (Figure 1c,d) must be considered. We calculated the FES as a function of two carefully chosen order parameters that describe the motions of the binding pocket well: the distance between the centers of mass of the heavy atoms of two groups of residues (Asp836–Asp964 and Asp836–Lys807) that are most affected by the presence of the ligand in the cavity (Figure 1, Figure 3b). The resulting FESs of the apo and the holo forms are significantly different, which confirms the presence of a significant local induced fit effect.[10] Several different rotamers are accessible to the lateral chains in the binding pocket, and some of them are less favorable. Many of the ligands

  

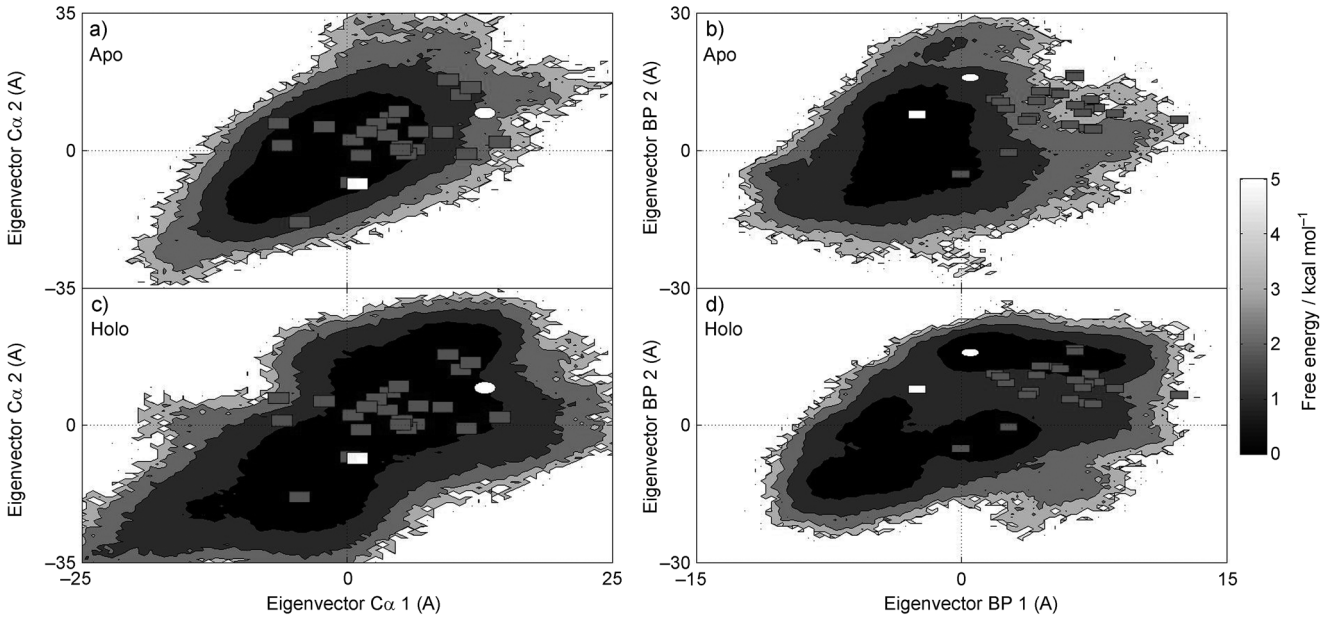


***Figure 2.*** a), b) FESs of the apo form of PI3K-γ. c), d) FESs of the L64-bound form of PI3K-γ. The FES in (a) and (c) are projected on the Cα eigenvectors space. The FES in (b) and (d) are projected on the binding pocket eigenvectors space. Each contour line corresponds to 1 kcal mol$^{-1}$. The projections of the 26 available PI3K-γ crystal structures are represented by gray squares with the exception of the apo form (1E8Y, white rectangle) and holo form (3IBE, white ellipse).

require a larger cavity, specific combinations of unfavorable orientations, and different backbone orientations to dock to the target. These conformations will rarely or never be visited by the apo form. Thus, similar to the recent findings for catalytically competent structures,[22] it is the entropic contribution that makes the ligand-accessible binding pocket conformations have high free energies.

To quantify the extent of conformational selection versus induced-fit effects, the $\rho_{sc}$[23] matrices of the trajectories of the apo form versus the holo form were also calculated. The $\rho_{sc}$ value measures structural deviations in a size-independent manner and can be used to compare Cα coordinates versus binding pocket atoms. The $\rho_{sc}$ values for Cα (Figure S8 in the Supporting Information) are almost always smaller than those of the binding pocket atoms (Figure S9 in the Supporting Information). This finding indicates that the structures of the apo and the holo forms that were sampled by MD simulations are quite similar at a global level, whereas larger differences are observed at the level of the binding pocket. By comparing the extent of the fluctuations in the residues in the binding pocket, it was confirmed that Ile963 and the residues that were chosen to calculate the FES are strongly affected by the presence of the ligand. In agreement with the calculated FES, two of the residues (Asp836 and Asp964), which are involved in the opening of the back pocket, have higher fluctuations in the simulations of the holo form. This result suggests that the conformational states that are compatible with the opening of the back pocket are partially inhibited in the absence of the ligand. The flexibility of Ile963 is reduced because of a direct interaction with the ligand (Table S4 and Figure S10 in the Supporting Information).

To investigate the entry of L64 into the binding pocket, we simulated an "artificial" induced-fit event by selecting 100 conformations from the 1 μs long MD simulation that started from 3DBS, docking L64 to them, and performing an additional 1 μs long MD simulation on the best pose. The RMSD that quantify the structural deviations in the binding pocket of the conformations that were sampled by MD simulations of the docked and apo forms from the 26 experimental crystal structures are reported in Figure S11 in the Supporting Information. It is apparent that, although the unbound trajectory samples conformations that are close to one of the bound crystal structures from time-to-time, it is in the trajectory of the bound form that regions with a very low RMSD value (with respect to the crystals) are located. By analyzing the trajectory that starts from the docked pose, it is also evident that the network of interactions between the protein and the ligand corresponds to that of the crystal structure, even if the initial docked pose is significantly different from the real pose (Figure S12 in the Supporting Information). Comparison of interaction patterns of the docked pose with those obtained from the real apo form 3IBE shows a fast convergence to a similar behavior (Table S4 and Figure S13 in the Supporting Information). Thus a MD simulation that starts from a non-optimal docked pose may evolve into the real interaction network in the case of L64. In contrast, the apo form by itself is unlikely to explore conformations in which the back pocket is open.

We directly simulated ligand binding by extensive bias-exchange free energy calculations. By using this method (see the Experimental Section for details) we observed several docking and undocking events and have reconstructed the major binding pathways. By using the unbiased replica we calculated the FES in the bi-dimensional space described by the protein–ligand distance and the principal eigenvectors (Figures S14 and S15 in the Supporting Information). A

  

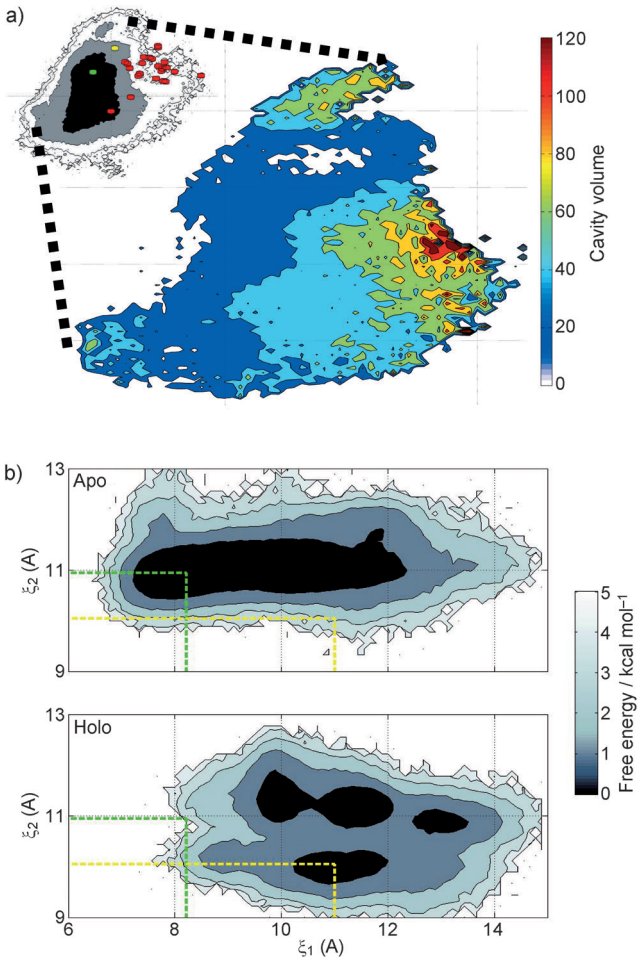


***Figure 3.*** a) Contour plot of the volume of the binding pocket cavity as a function of the binding pocket eigenvectors. The inset shows the corresponding FES for Figure 2b, cut at the 2 kcal mol$^{-1}$ contour level. b) FESs as a function of the distances between the center of mass of the Asp836–Asp964 (ξ1) and Asp836–Lys807 (ξ2) residues calculated from the trajectories of the apo and holo forms. The values that correspond to the crystal structures of the apo and holo forms are indicated by green and yellow dotted lines.

careful analysis of the main docking paths suggests that the binding is linked to conformational rearrangements in the binding pocket that are, in part, described by the first two binding pocket PCA vectors.

In summary, an analysis of very long trajectories of the apo form and ligand-bound form of PI3K-γ and free energy calculations show that the mechanism by which a ligand binds to this pharmaceutically important target can be best described as a combination of a long-range conformational selection that is complemented by a more localized ligand-induced conformational shift. Thus, modeling protein plasticity over different lengths and time scales during the ligand binding process is critical to understanding the mechanism of the process, and to rationalize the design of bioactive compounds.

## *Experimental Section*

The simulations were performed by using GROMACS 4[24] with Amber99SB-ILDN*[25,26] including backbone corrections.[27] This improved version of Amber99SB has been shown to correctly reproduce the conformational ensemble of the bovine pancreatic trypsin inhibitor (BPTI) protein and various other miniproteins.[28] Three different initial crystal structures were used: the best-resolved apo form (1E8Y), the structure that is cocrystallized with the inhibitor L64 (3IBE), and the structure 3DBS. The 3DBS structure, which is co-crystallized with a different compound than L64, has a partially open binding pocket conformation, but does not have an accessible back pocket to fit the urea moiety of L64. Thus, it can be considered an "intermediate" between the apo form and the L64-bound forms. The structures were minimized, solvated, and equilibrated for 10 ns in the isothermal-isobaric (NPT) ensemble. Four canonical ensemble (NVT) MD simulations of 1.5 μs (1E8Y) and 1 μs (3IBE, 3DBS-apo, and 3DBS-docked) at 300 K were performed, with a time step of 2 fs and the velocity-rescaling algorithm.[29]

The essential dynamics analysis[31] was applied to the trajectories to extract the eigenvectors that represent the directions of the largest concerted protein motions. The overlap between the different sets of eigenvectors was calculated by using the method proposed by Hess.[32] The FESs as a function of the first two PCA vectors and the distance between the centers of mass of the heavy atoms of Asp836–Asp964 and Asp836–Lys807 were calculated by discarding the trajectories every 2 ps and applying the standard formula: $A(q) = -RT\ln(p(q))$, where $p(q)$ is the probability distribution that is obtained from the simulation. The convergence of the FESs that were obtained from the MD simulation of the apo form and the eventual effects as a result of starting from a different crystal structure were checked by comparison with those that were obtained by starting from 3DBS after removal of the ligand (Figures S4 and S5 in the Supporting Information). As intermolecular packing and lattice forces are expected to shift the structures from the energy minima,[21] all the structures were minimized with the force field before projecting them on the FES. The induced-fit effects were also checked by performing a 1 μs long MD simulation that started from a docked pose. 100 Snapshots of the 3DBS-apo MD simulation with the lowest RMSF value to the crystal structure 3IBE were used to re-dock L64 by using Gold 3.1[30] (Figure S12 in the Supporting Information). Indeed, during the long MD simulation the binding mode that is found in the crystal structure was reproduced better and better. Metadynamics-based bias-exchange simulations[14] were performed[33] with 6 collective variables (CVs): the first four eigenvectors of the binding pocket[34] and two protein–ligand distances. Each CV was run on a separate replica for more than 200 ns. The main free energy minima in the six-dimensional space were also determined.[14]

Received: May 12, 2011
Revised: October 19, 2011
Published online: December 1, 2011

**Keywords:** conformational selection · induced fit · kinases · molecular dynamics · molecular recognition

[1] A. Fersht, *Structure and Mechanism in Protein Sci: A Guide to Enzyme Catalysis and Protein Folding*, Freeman, New York, **1999**.
[2] G. Hammes, *Enzyme Catalysis and Regulation*, Academic Press, New York, **1982**.
[3] H. J. Dyson, P. E. Wright, *Curr. Opin. Struct. Biol.* **2002**, *12*, 54–60.
[4] A. Caflisch, *Trends Biotechnol.* **2003**, *21*, 423–425.
[5] J. Monod, J. Wyman, J. P. Changeux, *J. Mol. Biol.* **1965**, *12*, 88–118.

  

[6] C. J. Tsai, B. Ma, R. Nussinov, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9970 – 9972.
[7] D. E. Koshland, *Proc. Natl. Acad. Sci. USA* **1958**, *44*, 98 – 104.
[8] O. F. Lange, N.-A. Lakomek, C. Farès, G. F. Schröder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, B. L. de Groot, *Science* **2008**, *320*, 1471 – 1475.
[9] K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964 – 972.
[10] A. Y. Lau, B. Roux, *Nat. Struct. Mol. Biol.* **2011**, *18*, 283 – 288.
[11] G. G. Hammes, Y.-C. Chang, T. G. Oas, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13737 – 13741.
[12] D. D. Boehr, R. Nussinov, P. E. Wright, *Nat. Chem. Biol.* **2009**, *5*, 789 – 796.
[13] H.-X. Zhou, *Biophys. J.* **2010**, *98*, L15 – L17.
[14] S. Piana, A. Laio, *J. Phys. Chem. B* **2007**, *111*, 4553 – 4559.
[15] T. Crabbe, M. J. Welham, S. G. Ward, *Trends Biochem. Sci.* **2007**, *32*, 450 – 456.
[16] R. Williams, A. Berndt, S. Miller, W.-C. Hon, X. Zhang, *Biochem. Soc. Trans.* **2009**, *37*, 615 – 626.
[17] C. Sadhu, B. Masinovsky, K. Dick, C. G. Sowell, D. E. Staunton, *J. Immunol.* **2003**, *170*, 2647 – 2654.
[18] A. Berndt, S. Miller, O. Williams, D. D. Le, B. T. Houseman, J. I. Pacold, F. Gorrec, W.-C. Hon, P. Ren, Y. Liu, C. Rommel, P. Gaillard, T. Rückle, M. K. Schwarz, K. M. Shokat, J. P. Shaw, R. L. Williams, *Nat. Chem. Biol.* **2010**, *6*, 117 – 124.
[19] A. J. Folkes, K. Ahmadi, W. K. Alderton, S. Alix, S. J. Baker, G. Box, I. S. Chuckowree, P. A. Clarke, P. Depledge, S. A. Eccles, L. S. Friedman, A. Hayes, T. C. Hancox, A. Kugendradas, L. Lensun, P. Moore, A. G. Olivero, J. Pang, S. Patel, G. H. Pergl-Wilson, F. I. Raynaud, A. Robson, N. Saghir, L. Salphati, S. Sohal, M. H. Ultsch, M. Valenti, H. J. A. Wallweber, N. Chi Wan, C. Wiesmann, P. Workman, A. Zhyvoloup, M. J. Zvelebil, S. J. Shuttleworth, *J. Med. Chem.* **2008**, *51*, 5522 – 5532.
[20] A. Zask, J. C. Verheijen, K. Curran, J. Kaplan, D. J. Richard, P. Nowak, D. J. Malwitz, N. Brooijmans, J. Bard, K. Svenson, J. Lucas, L. Toral-Barza, W.-G. Zhang, I. Hollander, J. J. Gibbons, R. T. Abraham, S. Ayral-Kaloustian, T. S. Mansour, K. Yu, *J. Med. Chem.* **2009**, *52*, 5013 – 5016.
[21] W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hünenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. A. van der Vegt, H. B. Yu, *Angew. Chem.* **2006**, *118*, 4168; *Angew. Chem. Int. Ed.* **2006**, *45*, 4064.
[22] J. R. E. T. Pineda, D. Antoniou, S. D. Schwartz, *J. Phys. Chem. B* **2010**, *114*, 15985 – 15990.
[23] G. N. Maiorov, V. M. Crippen, *Proteins Struct. Funct. Genet.* **1995**, *22*, 273.
[24] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435 – 447.
[25] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins Struct. Funct. Genet.* **2006**, *65*, 712 – 725.
[26] S. Piana-Agostinetti, K. Lindorff-Larsen, P. Maragakis, M. P. Eastwood, R. O. Dror, D. E. Shaw, *Biophys. J.* **2009**, *96*, 406a.
[27] R. B. Best, N. V. Buchete, G. Hummer, *Biophys. J.* **2008**, *95*, L07 – L09.
[28] D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, W. Wriggers, *Science* **2010**, *330*, 341 – 346.
[29] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.
[30] G. Jones, P. Willett, R. C. Glen, *J. Mol. Biol.* **1995**, *245*, 43 – 53.
[31] A. Amadei, A. B. M. Linssen, H. J. C. Berendsen, *Proteins Struct. Funct. Genet.* **1993**, *17*, 412 – 425.
[32] B. Hess, *Phys. Rev. E* **2002**, *65*, 031910.
[33] A. Berteotti, A. Cavalli, D. Branduardi, F. L. Gervasio, M. Recanatini, M. Parrinello, *J. Am. Chem. Soc.* **2009**, *131*, 244 – 250.
[34] L. Sutto, M. D'Abramo, F. L. Gervasio, *J. Chem. Theory Comput.* **2011**, *6*, 3640 – 3646.

  