JCTC Journal of Chemical Theory and Computation

Letter

pubs.acs.org/JCTC

Assessing the Performance of Metadynamics and Path Variables in Predicting the Binding Free Energies of p38 Inhibitors

G. Saladino,[†] L. Gauthier,[‡] M. Bianciotto,[‡] and F. L. Gervasio*,[†]

[†]Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre (CNIO), c/Melchor Fernandez Almagro 3, 28029, Madrid, Spain

[‡]Structure, Design, Informatics, Lead Generation to Candidate Realization, Sanofi R&D, 195 route d'Espagne, Toulouse, France

ABSTRACT: The accurate yet efficient evaluation of the free energy profiles of ligand-target association is a long sought goal in rational drug design. Methods that calculate the free energy along realistic association pathways, such as metadynamics, have been shown to provide reliable profiles, while accounting properly for solvation and target flexibility. However, these approaches usually require prohibitive computational resources and expert human intervention. Here, we show how multiple walkers metadynamics, when performed with optimal path collective variables (PCV), provides in a predetermined amount of computer time an accurate set of free energy profiles for a series of p38 inhibitors. The chosen test set, spanning a wide range of activity, is a challenging benchmark, both for computational methods and for human intuition, as the correct order for the binding affinity cannot be easily guessed. An excellent ranking of the ligands was obtained with minimal human assistance, an important step toward a fully automated pharmaceutical work-flow.

INTRODUCTION

Free energy methods have been used for three decades to predict the binding free energy of drug-like ligands to pharmaceutical targets. End-point methods like free energy perturbation (FEP),[1] linear interaction energy (LIE),[2] and MM-PBSA[3] give excellent estimates of the relative binding free energy[4−7] and improve the scoring when coupled with docking.[8−10] Methods based on a physical path, especially umbrella sampling[11] and more recently metadynamics,[12,13] are in general more computationally expensive but have the advantage of predicting the full free energy profile along the whole binding pathway. This is advantageous because it enables estimation of the kinetics of binding that is being recognized as an important parameter in vivo.[14] Moreover, if important intermediate states are present, as for instance in the HIV integrase case,[15] physical path-based methods are able to describe them in detail, disclosing alternative routes for drug design. Umbrella sampling has been so far the most used free energy technique for the study of single drug binding.[16−19] However, with umbrella sampling, it is difficult to reconstruct free energy landscapes with a dimensionality higher than two, complicating the task of describing complex biomolecular phenomena such as drug binding. Metadynamics, conversely, is able to reconstruct multidimensional free energy profiles, taking into account the relevant degrees of freedom.

This greater flexibility underlies the increasing acceptance of metadynamics as the method of choice to study complex molecular recognition phenomena.[20−27] Unfortunately, metadynamics, as with umbrella sampling and other collective variables (CV) based techniques, requires a careful choice of the few collective variables along which the dynamic of the system is biased and the multidimensional free energy surface (FES) will be projected. These CVs must describe all of the slow degrees of freedom involved in the process for an accurate reconstruction of a meaningful FES.[13] As discussed in ref 25, the identification of a proper set of CVs able to properly describe binding and unbinding events is a rather challenging task. Intuitive guesses, such as for example the receptor−ligand distance, that might be considered a good choice at first have been proven to lead to convergence and/or efficiency issues, due to the large volume of the unbound state. Such choices negatively affect the efficiency of these techniques, as several trial simulations might be needed to identify the correct set of CVs to be used. Path Collective Variables (PCV)[28] coupled with metadynamics (PCV-metaD) not only solves the problem of finding an efficient CV to describe drug binding but also simplifies and standardizes the initial setup of the simulation. PCVs are based on the definition of an optimal, physically meaningful pathway in the form of a discrete series of PDBs, so that the simulations can then be efficiently guided along a suitable path. This PCV-metaD approach was pioneered in ranking a small congeneric series of CDK2 inhibitors.[25]

Still, the convergence of the free energy profiles, even with PCV, can be so slow to make the approach unpractical. To decrease the wall clock time and still reach convergence, there are two possibilities: PTmetaD[29] or multiple-walkers metadynamics.[30] The latter has several advantages in a pharmaceutical work-flow: walkers can be added or subtracted at will, making the use of temporarily available computer resources simple and efficient; the number of walkers do not depend on the size of the system, as for PTmetaD. The disadvantage with respect to PTmetaD is that the sampling of slow transverse variables is less efficient.[29]

In this paper, we systematically apply PCV-metaD in a totally unsupervised manner to the difficult case of a well-known series of p38 inhibitors. The mitogen activated protein kinase (MAPK) p38 is an interesting target as it plays a crucial role in several signaling mechanisms and regulates a wide range of

Published: March 26, 2012

   dx.doi.org/10.1021/ct3001377 | J. Chem. Theory Comput. 2012, 8, 1165−1170

cellular processes, ranging from response to stress to migration and proliferation.[31] The development of p38 inhibitors is particularly important for fields like oncology, as it was suggested that p38 might act as a tumor suppressor, down-regulating the cell cycle progression and inducing apoptosis.[32] The selected series of inhibitors has been previously scored with docking algorithms,[33] energy grid methods,[34] implicit solvent methods such as MM-PB(GB)SA,[8] and thermodynamic integration.[33] The chosen test set provides not only a larger and more diverse congeneric series of compounds spanning a wide range of activity, with IC_{50} between 36 nM and 1.9 μM, but also a challenging benchmark, as the correct order of activity cannot be intuitively guessed. It is noteworthy that these calculations have been performed during the evaluation of a tier-1 (national-scale) supercomputing center, in terms of cost, performance, and practical issues that might arise when outsourcing resource-intensive calculations on an external HPC center. As it was known in advance that no interactive shell would be available during the course of the evaluation, the process has been automated through scripts for metadynamics path generation, running setup, monitoring, conditional resubmission, and termination. This "blind" screening process which forbids any posthoc adaptation of the work-flow was meant to stress-test the suitability of PCV with metadynamics in a realistic context.

MATERIALS AND METHODS

To carry out our test, we performed parallel PCV-metadynamics simulations using a common initial guess path for eight ligands taken from the reported test set (see Figure 1 and Table 1). Since the access to the High Performance Computing (HPC) facility was bound to a limited amount of CPU hours, we were not able to screen all of the ligands reported in the original paper.[33] However, in order for our results to be comparable with previously reported ones, we carefully selected the ligands, including those giving the poorest score−IC_{50} correlations. All of the simulations were conducted at the same time, without human supervision, using the NAMD 2.7 package[35] combined with the PLUMED plugin,[36] implementing PCV-metaD and using the Amber99SB force field for the protein part and the General Amber Force Field (GAFF) for the ligands, in explicit TIP3P water. All of the ligands were parametrized using the antechamber package,[37] with AM1-BCC level charges. The initial p38 structure (PDB code 3FC1) was prepared using the Schrödinger Suite (2009 version). Missing residues (10−12, 27−32 in the flap loop, and 166−174 in the activation loop) were added using Schrödinger Prime using p38 sequence Q16539. The four conserved crystallographic water molecules were retained. The starting positions for the other ligands were generated by docking on this structure. Each of the systems was then prepared using a stepwise minimization and relaxation procedure before an equilibration of 4 ns in the NPT ensemble and 1 ns in the NVT ensemble.

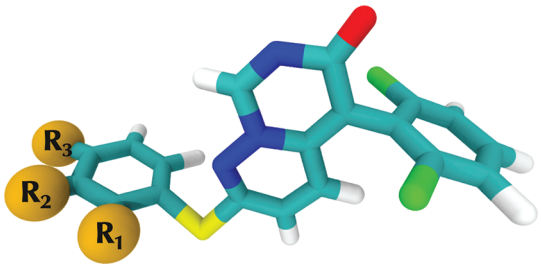

Figure 1. The common scaffold of the p38 inhibitors reported in ref 33 and used in the present study, based on a 1-(2,6-dichlorophenyl)-6-(phenyl)sulfanylpyrido[3,2-d]pyrimidin-2-one. The different moieties corresponding to R_1, R_2, and R_3 in the different ligands are reported in Table 1.

Table 1. Selected p38 Inhibitors Substituents Referring to the Common Scaffold in Figure 1

Nr.[a]	R_1	R_2	R_3	pIC_{50}
1	H	H	H	6.602
2	H	H	F	7.000
3	H	H	CH_3	5.854
4	H	Cl	Cl	6.097
5 (9)	H	Cl	F	6.301
6 (13)	CH_3	H	CH_3	6.577
7 (16)	Cl	H	F	7.444
8 (17)	F	H	F	8.046

[a]The numbering of ligands 1−4 is the same used in ref 33. For ligands 5−7, the corresponding numbers in ref 33 are given in parentheses. Ligand 8 is taken from ref 38 with the original number given in parentheses.

We used the well-tempered formulation of metadynamics[39] enhanced with the multiple-walkers routine,[30] using a total of 10 walkers and a bias factor of 16. In order to comply with the "blind" nature of our screening, all simulations were arbitrarily stopped after a cumulative time of 240 ns NVT production phase at 300 K (10 walkers each running for 24 ns). To trace the path, the procedure of ref 28 was used, introducing the two variables $s(\mathbf{R})$ and $z(\mathbf{R})$:

$$s(\mathbf{R}) = \frac{1}{P-1}\left(\frac{\sum_{l=1}^{P} l\, e^{-\lambda \| \mathbf{S}(\mathbf{R}) - S(l) \|^2}}{\sum_{l=1}^{P} e^{-\lambda \| \mathbf{S}(\mathbf{R}) - \mathbf{S}(l) \|^2}} - 1\right)$$

and

$$z(\mathbf{R}) = -\frac{1}{\lambda}\ln\left(\sum_{l=1}^{P} e^{-\lambda \| \mathbf{S}(\mathbf{R}) - \mathbf{S}(l) \|^2}\right)$$

where the distance ||...|| between the current position $\mathbf{S}(\mathbf{R})$ and the reference frames of the path $\mathbf{S}(l)$ was calculated using a Mean Square Displacement (MSD) metric after alignment to the reference applying a roto-translation matrix (calculated with the Kearsley algorithm, as implemented in the PLUMED plugin[36]). Following the approach described in ref 25, exploratory metadynamics using suboptimal geometrical CVs was performed on the reference system (3FC1 structure) to identify a plausible exit pathway. In particular, the distance r between the centroid of the central heterocyclic core of the ligand and the C_α of binding site residues Lys_{71}, Leu_{85}, and Thr_{105} and the dihedral θ between the major inertia axes of the ligand and the protein (identified by the two central atoms of the heterocyclic core of the ligand and the centroids of two sets of C_α atoms in the stable α-helices 122−145 and 201−216) were used. A slow steering along the lowest energy path in the r,θ free energy map was perfomed to obtain an initial set of frames. This initial guess path was then optimized using the methodology described in ref 28. A total of 21 reference frames were selected to trace the path, with a mean interframe MSD of 2.16 Å^2. The C_α atoms of the β sheets in the hydrophobic cavity

dx.doi.org/10.1021/ct3001377 | J. Chem. Theory Comput. 2012, 8, 1165−1170

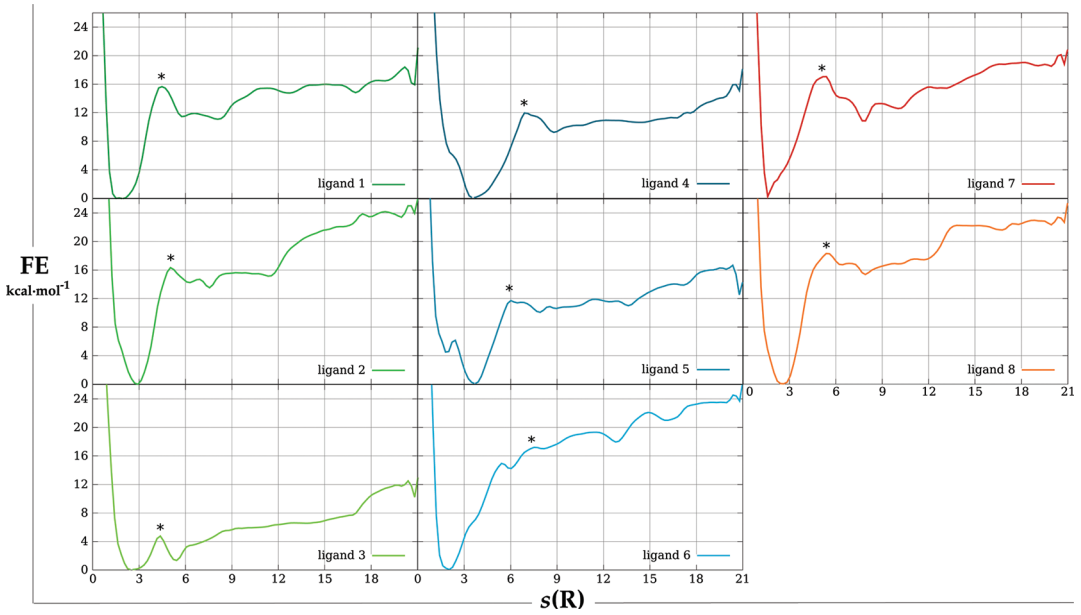

Figure 2. Free energy profiles for ligands 1−8 as a function of the $s(\mathbf{R})$ variable. In every profile, the automatically selected $s^{\ddagger}$ state is indicated with a * symbol.

were selected for the alignment, together with the C_α of the three longest α-helixes in the so-called C-lobe (approximately, residues 113−315). The MSD was calculated considering the non-hydrogen atoms of the ligand and the C_α of the loop sitting on top of the ligand (so-called *p-loop*) in the crystallographic structure and is known to play an important role in ligand binding. The same initial guess path was used for all eight ligands. Gaussians were added along both $s(\mathbf{R})$ and $z(\mathbf{R})$ every 5000 steps, with a height of 0.5 kcal/mol and a width of 0.3 CV units along each variable.

Following the approach first introduced in ref 33, we evaluate the predictive power of our "blind" PCV-metaD using not only the R^2 metrics, but also the Prediction Index (PI):

$$\mathrm{PI} = \frac{\sum_{j>i}\sum_i w_{ij}C_{ij}}{\sum_{j>i}\sum_i w_{ij}}$$

where

$$w_{ij} = |E_j - E_i|$$

and

$$C_{ij} = \begin{cases} 1 & \text{if } \dfrac{E_j - E_i}{P_j - P_i} < 0 \\ -1 & \text{if } \dfrac{E_j - E_i}{P_j - P_i} > 0 \\ 0 & \text{if } P_j - P_i = 0 \end{cases}$$

P_i and E_i are respectively the calculated and the experimental binding affinities for compound i. The function, developed by Pearlman et al., ranges from −1 to +1 with +1 meaning a totally exact prediction, −1 a completely wrong one, and 0 a random prediction. The term w_{ij} weights the results depending on their difference in experimental value, thus giving a higher reward/penalty when the method succeeds/fails to correctly score compounds with a large difference in affinity.

RESULTS

The free energy along the $s(\mathbf{R})$ and $z(\mathbf{R})$ variables was obtained from the history-dependent biasing potential of metadynamics and, for an easier comparison, projected along the $s(\mathbf{R})$ variable alone to obtain 1D profiles for all of the ligands (see Figure 2). The binding ΔG was computed integrating the bound ($s(\mathbf{R}) < 14$) and unbound ($s(\mathbf{R}) > 14$) regions of the profiles. Due to the rather flat nature of the profile once past the exit barrier ($s(\mathbf{R}) > s^{\ddagger}$), we found that the ΔG values were not significantly affected by the definition of the docked/undocked states.

Given that we started all of the simulations from the bound state and we stopped all of them after an arbitrary amount of time (i.e., 240 ns), the free energy profile of the bound state is much better converged than the profile on the unbound side. Taking the last 50 ns of the runs, the variation of the aligned free energy profiles in the bound state was less than 1 kcal/mol, while on the unbound state it was ∼2.5 kcal/mol, clearly due to the better sampling of the bound state. This is due to the different characteristics of the two states and to the nature of the metadynamics algorithm. Metadynamics is generally very efficient at flattening deep enthalpic minima, as the bound states, and much worse at dealing with flat entropic plateaus as the unbound states.[13] The use of a multiple-walkers algorithm surely helps the convergence of the free energy profile in the unbound case, explaining the observed performance. Still, to ensure full convergence in all cases, longer runs would be needed. For this reason, and following the same rationale of ref 25, we also calculated the free energy difference between the bound state and the transition state $\Delta G^{\ddagger}$ since in ref 22 it was shown that it often correlates well with the pIC_{50}. This correlation should not be too surprising as the k_{on} is generally, but not always,[40] similar for congeneric compounds. As the free energy profiles are complex and the relative free energy value of the plateau corresponding to the unbound state depends on the volume accessible to the unbound ligand, we define, in the following, $\Delta G^{\ddagger}$ to be the largest energy barrier in the profile, which corresponds always to the first barrier to the exit of the ligand from the cavity. In particular, it arises from the breaking

dx.doi.org/10.1021/ct3001377 | J. Chem. Theory Comput. 2012, 8, 1165−1170

of the native ligand−receptor contacts and the formation of the first solvation shells around the ligand. From a practical point of view, this state $s^{\ddagger}$ was found to be in the region $s(\mathbf{R}) \in [3,8]$ for all of the ligands, and the free energy $\Delta G^{\ddagger}$ was hence calculated with respect to the highest point in this window. The results are reported in Table 2.

Table 2. Correlation between the Calculated ΔG and Experimental pIC_{50} before and after Correction[a]

Nr.	pIC_{50}	$\Delta G^{\ddagger}$	ΔG	$\Delta G^{\ddagger}_{corr}$	ΔG_{corr}	Γ_{corr}
8 (17)	8.046	18.294	20.792	16.054	18.552	2.24
7 (16)	7.444	17.016	15.402	14.866	13.252	2.15
2	7.000	16.363	19.659	13.973	17.269	2.39
1	6.602	14.498	14.203	12.818	12.523	1.68
6 (13)	6.577	17.303	18.583	15.013	16.293	2.29
5 (9)	6.301	11.636	10.961	9.586	8.911	2.05
4	6.097	11.943	10.561	9.903	8.521	2.04
3	5.854	4.645	6.500	2.605	4.460	2.04
PI		**0.87**	**0.89**	**0.87**	**0.89**	
R^2		*0.64*	*0.66*	*0.64*	*0.66*	

[a]All ΔG values are in kcal mol^{-1}.

As the PCVs allow the exploration of only a small region of the available unbound space, we also corrected our results according to

$$\Delta G_{corr} = -RT\ln\frac{\xi^{meta}_{bulk}V^{meta}_{bulk}V_{box}}{8\pi^2 V_{bulk}V_0} - RT\ln\frac{Z^{meta}_{site}}{Z^{meta}_{bulk}} = \underbrace{-RT\ln\frac{\xi^{meta}_{bulk}V^{meta}_{bulk}V_{box}}{8\pi^2 V_{bulk}V_0}}_{\Gamma_{corr}} + \Delta G_{meta}$$

as reported in ref 25, following ref 41. In the reported equation, V^{meta}_{bulk} represents the volume of the unbound state actually explored during the PCV-metaD run and ξ^{meta}_{bulk} is the fraction of the total possible orientations explored by the ligand in the unbound state. V_0 is the standard volume (1661 Å^3), and V_{bulk} is the bulk volume (i.e., $V_{box} - V_{protein}$). The results corrected for the unbound state volume are reported alongside the uncorrected ones in Table 2. The calculated contributions to the Γ_{corr} in Table 2 are reported in Table 3.

Table 3. Contributions to Γ_{corr} Calculated from the PCV-metaD runs

Nr.	V^{meta}_{bulk} (Å^3)	ξ^{meta}_{bulk} (rad^3)	Γ_{corr} (kcal/mol)
1	246.96	26.527	1.68
2	98.784	20.199	2.39
3	167.384	21.61	2.04
4	156.408	23.132	2.04
5	189.336	18.698	2.05
6	128.968	18.351	2.29
7	120.736	25.008	2.15
8	87.808	29.226	2.24

Without any sort of postprocessing, the relative ΔG values taken "as is" revealed an excellent correlation with the experimental values, resulting in an unprecedented PI of 0.89. For comparison, the PIs reported for various scoring methods are given in table Table 4.

The PCV-metaD result is then higher than the previously reported best score (0.85), registered for the thermodynamic integration method,[33] albeit at a much higher computational cost. While no standard deviation is available for the measured IC_{50}, Jorgensen et al.[38] suggested that an average 0.41 kcal/mol error on the pIC50 is to be expected, thus downplaying the incorrect ranking of ligand 4 with respect to 5, as its $\Delta(pIC_{50})$ (0.204 kcal/mol) presumably lies within the standard deviation itself. Interestingly, an impressive PI of 0.92 was reported for the p38 test set by Pearlman et al.[41] using the "FURMASA" method, based on MD-averaged grids and solvent accessible surface area calculations; cross-correlation analyses to attain the true prediction power of the method lowered the PI to 0.82. It is noteworthy that, considering the low cost of the approach, the "FURMASA" result is extremely encouraging.

Table 4. Prediction Index for Various Scoring Methods

scoring method	PI
PCV-metaD	0.89
thermodynamic integration (TI)[33]	0.85
FURMASA[42]	0.82
MC/FEP + JAWS[38]	0.62
OWFEG[34]	0.56
MM-PBSA (at best)[8]	0.51
Dock Energy Score[33]	0.25
ChemScore[33]	0.04
PLPScore[33]	−0.05

We calculated the error of the predicted ΔG relative to the theoretical values from the linear regression ΔG. The average error is about 2.2 kcal/mol for ΔG and 1.9 kcal/mol for $\Delta G^{\ddagger}$. Ligand number 6, the dimethyl substituted compound, was the one reporting the highest deviation from the expected value, with an error of 4.7 and 4.0 kcal/mol for ΔG and $\Delta G^{\ddagger}$, respectively. Interestingly, the dimethyl ligand has been reported as an outlier by most of the methods reported in Table 3.

Removal of ligand 6 from the deviation average lowers the mean error of both ΔG and $\Delta G^{\ddagger}$ of 0.3 kcal/mol. Repeating the linear regression and the prediction index calculation without ligand 6 resulted in an increase of both the correlation coefficient (0.74 and 0.8 for ΔG and $\Delta G^{\ddagger}$, respectively) and the PI (0.96 and 0.98 for ΔG and $\Delta G^{\ddagger}$, respectively).

Analyses of the FES also revealed that, of all of the selected compounds, ligand 6 was the one reporting the worst convergence. We extended the metadynamics run for this ligand until a reasonable convergence was achieved (300 ns). Calculation of ΔG and $\Delta G^{\ddagger}$ using the newly converged FES led to results comparable with the ones obtained removing the dimethyl ligand: PI of 0.96 and 0.98 and R^2 of 0.77 and 0.75 for ΔG and $\Delta G^{\ddagger}$, respectively.

The results in Table 2 clearly suggest that PCV-metaD has the potential to be an optimal tool for lead optimization purposes. Minor corrections to the raw values resulted in an even higher PI of 0.98, compared to the already satisfactory 0.89 obtained before. Moreover, the data revealed a more than satisfactory quantitative correlation between the calculated ΔG and experimental pIC_{50} (Figure 3).

The linear regression of the data in Table 2, including correction to ligand 6, resulted in a determination coefficient R^2 of 0.77, suggesting that not only does PCV-metaD give an almost perfect ranking of the different ligands but it can also quantitatively distinguish between them.

Moreover, the availability of the full binding pathway, comprising all important intermediates and transition states,

dx.doi.org/10.1021/ct3001377 | J. Chem. Theory Comput. 2012, 8, 1165−1170

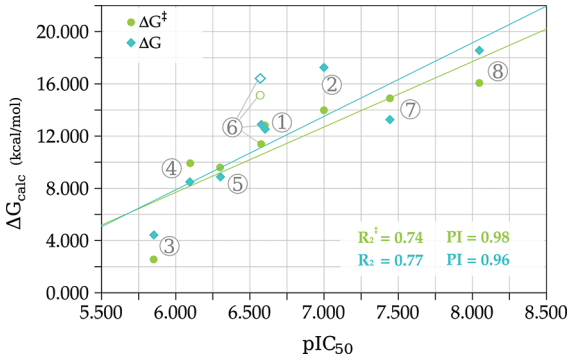

Figure 3. Linear regression of the data in Table 2 (after correction), except for ligand 6, for which the results obtained from the 300 ns run are used (the ΔG's at 240 ns are reported with a white-filled symbol). A satisfactory determination index R^2 of 0.77 was obtained, suggesting a semiquantitative correlation for the calculated ΔG.

delivers a wide range of supporting information, potentially important for the lead optimization process. If we compare the binding profiles of the strongest and weakest binders (ligands 8 and 3, respectively), we can observe an interesting difference: while both of them present a distinctive conformation at $s(\mathbf{R}) \sim 5$, this is a transition state for the first ligand and a stable secondary minimum for ligand 3. The analysis of the structures corresponding to this secondary minimum revealed that ligand 3 does not engage in the two strong hydrogen bonds with the hinge residues Met_{109} and Gly_{110}, which are typical of kinase inhibitors and whose population (38% and 41% in the deepest minimum) drops significantly (9% and less than 1%) in the secondary minimum. The increased flexibility allows the molecule to establish a range of alternative interactions and to assume a wider ensemble of binding poses. The hydrophobic interactions of the methyl moiety with the so-called "pocket I" hydrophobic region,[43] and in particular with the "gatekeepers" (residues Leu_{104} and Thr_{106}), are substituted at times by interactions with residues Tyr_{35}, Val_{38} in the Gly-rich loop, and Lys_{53}. When the interactions with Leu_{104} and Thr_{106} are lost, the other hydrophobic interactions are maximized, and the ligand assumes a "flipped" conformation, with the carbonyl group pointing outside of the cavity (Figure 4).

The same kinds of interactions are formed by ligand 8 in the transition state ($s(\mathbf{R}) \sim 5$). As the difluoro-phenyl group detaches from the "gatekeepers", interactions with Tyr_{35}, Val_{38}, and Lys_{53} are formed. However, in this case, no "flip" is observed, and the hydrogen bonds with Met_{109} and Gly_{110} are conserved. Moreover, as this compound lacks the terminal CH_3 group of ligand 3, the molecule needs to progress further toward the cavity exit to interact with the Gly-rich loop. This leads to the unfavorable solvation of the first dichloro-phenyl group that presumably contributes to rendering this conformation a transition state, rather than a secondary minimum as in ligand 3.

CONCLUSIONS

In the optimization phase of rational drug design, due to the complexity and flexibility of the targets and to the recognition of the importance of drug binding kinetics, there is an increasing need for automatic computational work-flows that are able to predict not only the binding free energy with a reasonable accuracy but also the full free energy profile along the reaction coordinate. Here, we use an unsupervised computational approach based on optimal path collective variables and multiple-walkers metadynamics to reconstruct the binding free energy profiles for a series of ligands binding to p38 that has been shown to be a meaningful and difficult test case. The results of our in-silico "blind" experiment suggest that the use of PCV and multiple walkers significantly improves the potential use of metadynamics in a drug-discovery work-flow, minimizing the system preparation phase and human intervention while providing reasonable free energy profile estimates in a set amount of computing time. The approach is more expensive than end-point methods such as TI[44] but much less expensive than other realistic path-based approaches.[19] Considering the important information deriving from the availability of the free energy profiles along drug association pathways, the approach proposed is ripe for a widespread use in the rational optimization phase of lead compounds.

AUTHOR INFORMATION

Corresponding Author

*E-mail: flgervasio@cnio.es.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge support by the Spanish Science and Innovation (MICINN) grant (BIO2010-20166, "AlteredDynamics"). L.G. and M.B. acknowledge D. Biehle for his insightful advice and technical support. The CEA-CCRT supercomputing center is acknowledged for the generous allocation of computer resources, and we warmly thank its

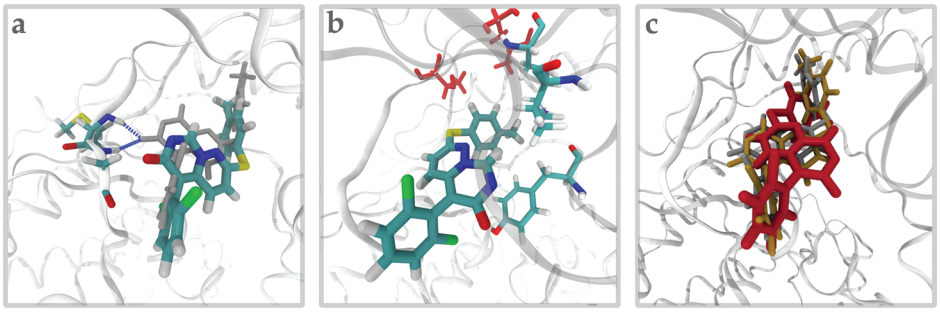

Figure 4. Conformations of ligand 3 in the deepest minimum (dark gray) and in the secondary minimum (a), where the two strong hydrogen bonds with residues Met_{109} and Gly_{110} are not observed. In the latter minimum, the molecule can adopt a "flipped" conformation (b, c) in which the methyl moiety interacts with residues Tyr_{35}, Val_{38}, and Lys_{53}.

representatives C. Ménaché and J.-N. Richet for their availability and support.

REFERENCES

(1) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
(2) Aqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994**, *7*, 385−391.
(3) Massova, I.; Kollman, P. A. *Perspect. Drug Discovery* **2000**, *18*, 113−135.
(4) Rizzo, R. C.; Toba, S.; Kuntz, I. D. *J. Med. Chem.* **2004**, *47*, 3065−74.
(5) Kuhn, B.; Kollman, P. A. *J. Med. Chem.* **2000**, *43*, 3786−3791.
(6) Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. *J. Am. Chem. Soc.* **2001**, *123*, 5221−5230.
(7) Masukawa, K. M.; Kollman, P. A.; Kuntz, I. D. *J. Med. Chem.* **2003**, *46*, 5628−37.
(8) Pearlman, D. A. *J. Med. Chem.* **2005**, *48*, 7796−7807.
(9) Lyne, P. D.; Lamb, M. L.; Saeh, J. C. *J. Med. Chem.* **2006**, *49*, 4805−4808.
(10) Guimaraes, C. R.; Cardozo, M. *J. Chem. Inf. Model.* **2008**, *48*, 958−970.
(11) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(12) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(13) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.
(14) Swinney, D. C. *Curr. Opin. Drug Discovery Dev.* **2009**, *12*, 31−9.
(15) Schames, J. R.; Henchman, R. H.; Siegel, J. S.; Sotriffer, C. A.; Ni, H.; McCammon, J. A. *J. Med. Chem.* **2004**, *47*, 1879−81.
(16) Bui, J. M.; Henchman, R. H.; McCammon, J. A. *Biophys. J.* **2003**, *85*, 2267−72.
(17) Deng, Y.; Roux, B. *J. Chem. Theory Comput.* **2006**, *2*, 1255−1273.
(18) Lee, M. S.; Olson, M. a. *Biophys. J.* **2006**, *90*, 864−77.
(19) Buch, I.; Harvey, M. J.; Giorgino, T.; Anderson, D. P.; De Fabritiis, G. *J. Chem. Inf. Model.* **2010**, *50*, 397−403.
(20) Gervasio, F. L.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2005**, *127*, 2600−7.
(21) Pietrucci, F.; Laio, A. *J. Chem. Theory Comput.* **2009**, *5*, 2197−2201.
(22) Masetti, M.; Cavalli, A.; Recanatini, M.; Gervasio, F. L. *J. Phys. Chem. B* **2009**, *113*, 4807−16.
(23) Provasi, D.; Bortolato, A.; Filizola, M. *Biochemistry* **2009**, *48*, 10020−10029.
(24) Limongelli, V.; Bonomi, M.; Marinelli, L.; Gervasio, F. L.; Cavalli, A.; Novellino, E.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 5411−6.
(25) Fidelak, J.; Juraszek, J.; Branduardi, D.; Bianciotto, M.; Gervasio, F. L. *J. Phys. Chem. B* **2010**, *114*, 9516−24.
(26) D'Abramo, M.; Rabal, O.; Oyarzabal, J.; Gervasio, F. L. *Angew. Chem., Int. Ed.* **2012**, *51*, 642−646.
(27) Lovera, S.; Sutto, L.; Boubeva, R.; Scapozza, L.; Dölker, N.; Gervasio, F. L. *J. Am. Chem. Soc.* **2012**, *134*, 2496−2499.
(28) Branduardi, D.; Gervasio, F. L.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 054103.
(29) Bussi, G.; Gervasio, F. L.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2006**, *128*, 13435.
(30) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. *J. Phys. Chem. B* **2006**, *110*, 3533−9.
(31) Cuadrado, A.; Nebreda, A. R. *Biochem. J.* **2010**, *429*, 403−17.
(32) Wagner, E. F.; Nebreda, A. R. *Nat. Rev. Cancer* **2009**, *9*, 537−49.
(33) Pearlman, D. A.; Charifson, P. S. *J. Med. Chem.* **2001**, *44*, 3417−3423.
(34) Pearlman, D. A.; Charifson, P. S. *J. Med. Chem.* **2001**, *44*, 502−511.
(35) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−802.
(36) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; et al. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.
(37) Wang, J.; Wang, W.; Kollman, P. a.; Case, D. a. *J. Mol. Graphics Modell.* **2006**, *25*, 247−60.
(38) Luccarelli, J.; Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2010**, *6*, 3850−3856.
(39) Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, 1−4.
(40) Sharp, S. Y.; et al. *Mol. Cancer Ther.* **2007**, *6*, 1198−211.
(41) Pearlman, D. a.; Rao, B. G.; Charifson, P. *Proteins* **2008**, *71*, 1519−38.
(42) Luccarelli, J.; Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2010**, *6*, 3850−3856.
(43) Fitzgerald, C. E.; Patel, S. B.; Becker, J. W.; Cameron, P. M.; Zaller, D.; Pikounis, V. B.; O'Keefe, S. J.; Scapin, G. *Nat. Struct. Biol.* **2003**, *10*, 764−9.
(44) Genheden, S.; Nilsson, I.; Ryde, U. *J. Chem. Inf. Model.* **2011**, *51*, 947−58.