Article

# Exploring Protein Kinase Conformation Using Swarm-Enhanced Sampling Molecular Dynamics

Alessio Atzori,[†,‖] Neil J. Bruce,[†,‖,⊥] Kepa K. Burusco,[†] Berthold Wroblowski,[‡] Pascal Bonnet,[§] and Richard A. Bryce*[,†]

[†]Manchester Pharmacy School, University of Manchester, Oxford Road, Manchester M13 9PT, U.K.

[‡]Janssen Research & Development, a division of Janssen Pharmaceutica N.V., Turnhoutseweg 30, 2340 Beerse, Belgium

[§]Structural Bioinformatics & Chemoinformatics, Institut de Chimie Organique et Analytique (ICOA), UMR CNRS-Université d'Orléans 7311, Université d'Orléans, Rue de Chartres, F-45067 Orléans Cedex 02, France

S *Supporting Information*
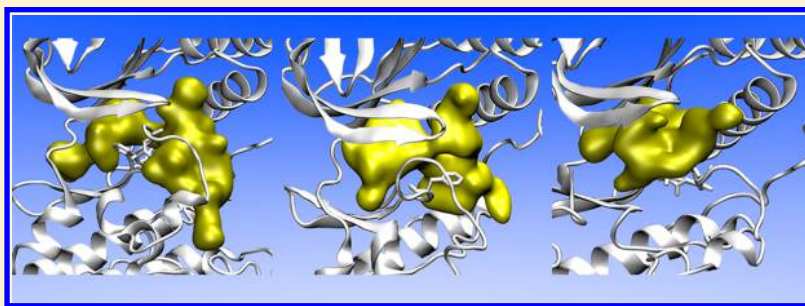
ABSTRACT: Protein plasticity, while often linked to biological function, also provides opportunities for rational design of selective and potent inhibitors of their function. The application of computational methods to the prediction of concealed protein concavities is challenging, as the motions involved can be significant and occur over long time scales. Here we introduce the swarm-enhanced sampling molecular dynamics (sesMD) method as a tool to improve sampling of conformational landscapes. In this approach, a swarm of replica simulations interact cooperatively via a set of pairwise potentials incorporating attractive and repulsive components. We apply the sesMD approach to explore the conformations of the DFG motif in the protein p38α mitogen-activated protein kinase. In contrast to multiple MD simulations, sesMD trajectories sample a range of DFG conformations, some of which map onto existing crystal structures. Simulated structures intermediate between the DFG-in and DFG-out conformations are predicted to have druggable pockets of interest for structure-based ligand design.

■ INTRODUCTION

Crystal structures are able to provide molecular-level detail on the interactions of ligands with their target receptors. Recently, valuable additional information has been mined from computing the time evolution of an initial protein or protein−ligand crystal structure using molecular dynamics (MD) simulations. For example, MD simulations of HIV integrase discovered a previously unknown part of the binding site that was able to explain experimental binding and mutagenesis data. This finding contributed to the successful development of the drug raltegravir, now a frontline drug in AIDS treatment.[1] Similarly, potential design opportunities from MD have been found, for example, for neuraminidase[2] and in recent work on *Trypanosoma cruzi* trans-sialidase.[3−5]

However, because of their rough free energy surfaces, observing nontrivial changes in protein structure, such as the opening and closing of pockets, loops, and grooves, is not commonly possible by MD without the application of biasing potentials.[6] These biasing methods can be equilibrium (e.g., umbrella sampling)[7] or nonequilibrium (e.g., steered MD) approaches.[8] Unfortunately, the manner in which these biasing

potentials or forces should be applied is not always apparent. Specific biased MD approaches range in their complexity of application and computational overhead and include active-site pressurization,[9] locally enhanced sampling,[10] accelerated molecular dynamics,[11] and replica-exchange-based approaches[12] as well as the group formed by local elevation,[13] conformational flooding,[14] and metadynamics methods.[15,16]

Here we propose an approach for exploring conformational space that utilizes a swarm-based MD scheme. This multicopy MD algorithm builds on the SWARM-MD method, which was first proposed by Huber and van Gunsteren[17] and more recently refined by us for the prediction of biomolecular structure.[18] The SWARM-MD approach involves creating multiple copies of the molecular system (i.e. a swarm of systems) and then simulating the dynamics of each copy in parallel. The simulations of these copies mimic cooperative swarm behavior through the addition of attractive forces that act on each member of the swarm, driving their trajectories

toward the mean trajectory of the entire swarm. In this way, the additional attractive force experienced by a given replica within the swarm is directly proportional to its distance from the average swarm structure. We demonstrated that our modified version of SWARM-MD is a simple and effective approach for the optimization of peptide and protein structures.[18] For example, during 20 independent 40 ns MD simulations of the Trp-cage miniprotein, none were able to fold to within a backbone root-mean-square deviation (RMSD) of 1.5 Å of the NMR conformation. By contrast, of 20 SWARM-MD replicas, 16 folded to a backbone RMSD below 1.5 Å within 40 ns.[18]

In this work, we move from using a swarm of MD replicas to optimize biomolecular structure to instead harnessing a swarm to explore the conformational landscape. To promote barrier crossing into alternative energy minima, we introduce a set of attractive and repulsive pair potentials that act between swarm copies. We refer to this method subsequently as swarm-enhanced sampling MD (sesMD).

We evaluate the ability of this sesMD approach to enhance the conformational exploration of a protein known to exhibit different structures, the anticancer target p38α mitogen-activated protein (MAP) kinase. Specifically, we focus on sampling of its Asp168-Phe169-Gly170 (DFG) motif, which can be characterized by different orientations of its phenyl-alanine side chain: in the "DFG-out" conformation, Phe169 points away from the nearby αC helix and projects into the ATP binding pocket, exposing an additional hydrophobic cavity, sometimes called the allosteric or deep pocket. In the "DFG-in" conformation, access to the deep pocket is hindered by Phe169 (Figure 1).[19] While active and inactive states of
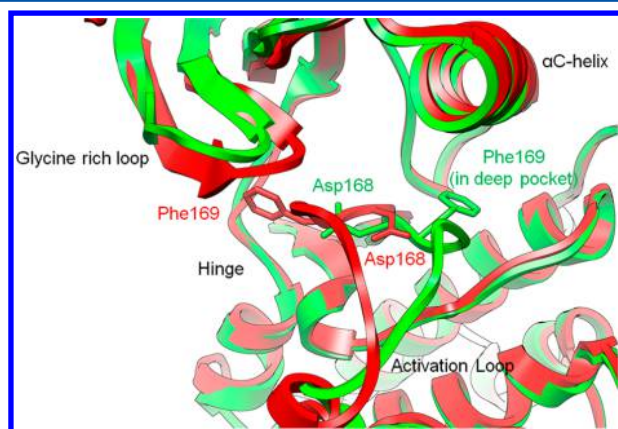

Figure 1. Superposition of the DFG-in (PDB 1P38, green) and DFG-out (PDB 1WBT, red) X-ray crystal structures of p38α MAP kinase. Backbone atoms are shown in cartoon representation and residues Asp168 and Phe169 in stick representation.

protein kinases can exhibit DFG-in conformations, the DFG-out conformation is ordinarily associated only with the inactive state. Thus, the movement of the phenylalanine side chain to the DFG-out conformation creates this deep pocket, a cavity that has been exploited in the design of inhibitors with good specificity, i.e., by type-II kinase inhibitors such as imatinib.[20] Experimental[21−24] and computational studies[25−32] have found that there is a significant energetic barrier between kinase DFG-in and DFG-out states. For example, molecular dynamics simulations have required strategies such as studying kinase mutants with modified kinetics,[28] massively distributed MD simulations,[29] or biasing methods.[25−27,30−32]

Here we apply sesMD to both DFG-out and DFG-in crystallographic conformations of p38α MAP kinase.[33,34] We assess the ability of sesMD to explore alternative conformations of the kinase that could be targeted subsequently by structure-based drug design approaches.

■ METHODS

sesMD. The SWARM-MD method described by Huber and van Gunsteren[17] involves the attraction of swarm members to the mean of the swarm via an extra potential, V^{swarm}, given by

$$V^{\text{swarm}}(\{\boldsymbol{\varphi}^{\alpha}\}) = \sum_{\alpha}^{M} A \, \exp[-B d_{\alpha}^{\text{RMS,av}}(\boldsymbol{\varphi}^{\alpha})] \tag{1}$$

where $d_{\alpha}^{\text{RMS,av}}$ is the root-mean-square dihedral angle difference of swarm member α from the average location of the swarm, $\boldsymbol{\varphi}^{\alpha}$ is the set of dihedral angles of swarm member α to which the swarm potential is applied, M is the number of members in the swarm, and A and B are parameters that govern the strength and range, respectively, of the attraction between each member and the swarm average.

Here we adapt SWARM-MD from an optimization approach to a method that permits enhanced coverage of conformational space by enabling multiple MD simulations of a molecular system to interact in a coupled fashion to surmount high energy barriers between conformations. We therefore introduce the following swarm-enhanced sampling (ses) potential V^{ses}:

$$V^{\text{ses}}(\{\boldsymbol{\varphi}^{\alpha}\}) = \sum_{\alpha}^{M} \sum_{\beta > \alpha}^{M} \{A \, \exp[-B d_{\alpha\beta}^{\text{RMS}}(\boldsymbol{\varphi}^{\alpha}, \boldsymbol{\varphi}^{\beta})]$$
$$+ C \, \exp[-D d_{\alpha\beta}^{\text{RMS}}(\boldsymbol{\varphi}^{\alpha}, \boldsymbol{\varphi}^{\beta})]\} \tag{2}$$

where $A-D$ are suitably calibrated parameters for attractive (A, B) and repulsive (C, D) terms and $d_{\alpha\beta}^{\text{RMS}}(\boldsymbol{\varphi}^{\alpha}, \boldsymbol{\varphi}^{\beta})$ is the root-mean-square dihedral angle distance between swarm members α and β. Distinct from eq 1, the first term in eq 2 is a *pairwise* attractive potential acting between pairs of swarm replicas. This potential is balanced by a repulsive exponential potential (the second term in eq 2). In this scheme, the values of A and C must be negative and positive respectively, while both B and D should be positive. The potential seeks to increase the spread of conformations explored while promoting transitions between wells. The resulting dynamics is therefore a combination of (i) the ses potential V^{ses}, which is applied to a selected set of dihedral angles and acts between pairs of replicas, and (ii) the interatomic potentials of the molecular mechanics force field, which act individually on each member of the swarm.

As a simple illustration of sesMD, we consider the effect of introducing V^{ses} on sampling of the conformational space of two solutes, pentane and alanine dipeptide (Ace-Ala-NMe). First, eight unbiased MD simulations of 10 ns were performed on each solute in explicit aqueous solvent (for computational details, see the Supporting Information). Sampling of low-energy regions for pentane (Figure 2a) and alanine dipeptide (Figure 2c) remains fairly localized, with only minor excursions into alternative wells on the landscapes. Alternatively, coupling together the eight replicas via the ses potential enables exploration of alternative low-energy regions during 10 ns of sesMD: for pentane, much greater coverage of symmetry-related gauche conformations is found (Figure 2b), while for alanine dipeptide, sampling of other regions of the Ramachan-dran landscape is achieved (Figure 2d). However, we note that
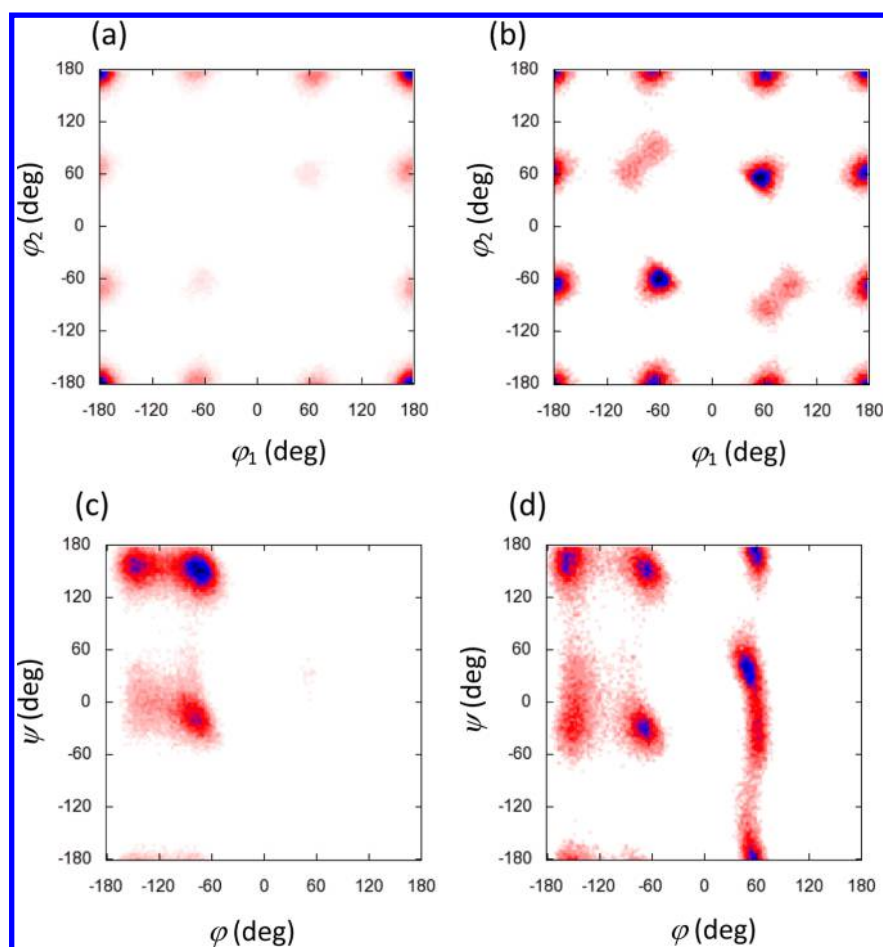
Figure 2. (a, b) Distributions of pentane conformations in water defined by the central C−C dihedral angles φ_1 and φ_2 during (a) eight independent MD simulations of 10 ns and (b) an eight-replica sesMD simulation of 10 ns. (c, d) Distributions of alanine dipeptide conformations in water as a function of the backbone dihedral angles ϕ and ψ during (c) eight independent MD simulations of 10 ns and (d) an eight-replica sesMD simulation of 10 ns. Blue regions indicate areas of highest density.

because of the presence of the V^{ses} term acting between replicas, the configurations sampled by sesMD are not drawn from a Boltzmann distribution. Nevertheless, the underlying potential energy surface topologies appear to be reasonably preserved via sesMD sampling (Figure 2b,d). Furthermore, as a result of the well-defined nature of the overall swarm Hamiltonian, it is formally possible to recover Boltzmann-weighted properties from the sesMD method where required (for more details, see the Supporting Information).

Computational Details. As starting points for molecular dynamics simulations of p38α MAP kinase, the crystallographic structures 1P38[33] and 1WBT[34] were used, representing DFG-in and DFG-out conformations of the protein, respectively. The structures were prepared for MD using the *xleap* program from the AMBER 11 suite.[35] The systems were modeled using the AMBER ff99SB force field[36] and solvated with ~13000 TIP3P water molecules. Counterions were added in order to neutralize the net charge of the systems. The SHAKE algorithm[37] constrained bonds between hydrogen and heavy atoms. A 2 fs time step was used. Simulations were performed in the *NPT* ensemble with a Langevin thermostat using a collision frequency of 2 ps^{-1}. A Berendsen isotropic barostat was used to maintain an average pressure of 1 atm. The particle mesh Ewald method[38] was used for computing long-range electrostatic interactions and a 12 Å cutoff for van der Waals' contributions.

For sesMD, the ses potential described in eq 2 was applied to the φ and ψ backbone dihedral angles of the DFG motif (Asp168, Phe169, and Gly170). For sesMD simulations starting from the DFG-in 1P38 crystal structure, the sesMD parameters (A, B) and (C, D) were taken as $(-150.0$ kcal/mol, 0.2 rad$^{-1})$ and $(150.0$ kcal/mol, 0.8 rad$^{-1})$ respectively. For sesMD simulations initiated from the DFG-out 1WBT structure, the parameters (A, B) and (C, D) adopted were $(-250.0$ kcal/mol, 0.2 rad$^{-1})$ and $(375.0$ kcal/mol, 0.8 rad$^{-1})$ respectively. The two ses potentials therefore differ in their dependence on $d_{\alpha\beta}^{RMS}$ (Figure 1S in the Supporting Information). The choice of parameters for the two structures is to some degree a reflection of the local potential energy landscape of the DFG loop within the two kinase states of DFG-in and DFG-out: a stronger ses potential was applied to the DFG-out structure, whereas the same strength of potential led to distortion of peptide bonds within the DFG loop for the DFG-in structure, apparently as a result of greater steric constraints of the occupied deep pocket within this structure.

For sesMD, 24 replicas of the kinase were employed. The initial configurations of the 24 replicas were generated as follows: the final structure of a 700 ps MD simulation at 300 K was taken as the initial configuration for 24 independent replica 500 ps simulations at 300 K with different initial velocities, and sesMD simulations were then performed using these 24 final configurations. During the first 500 ps of sesMD, the
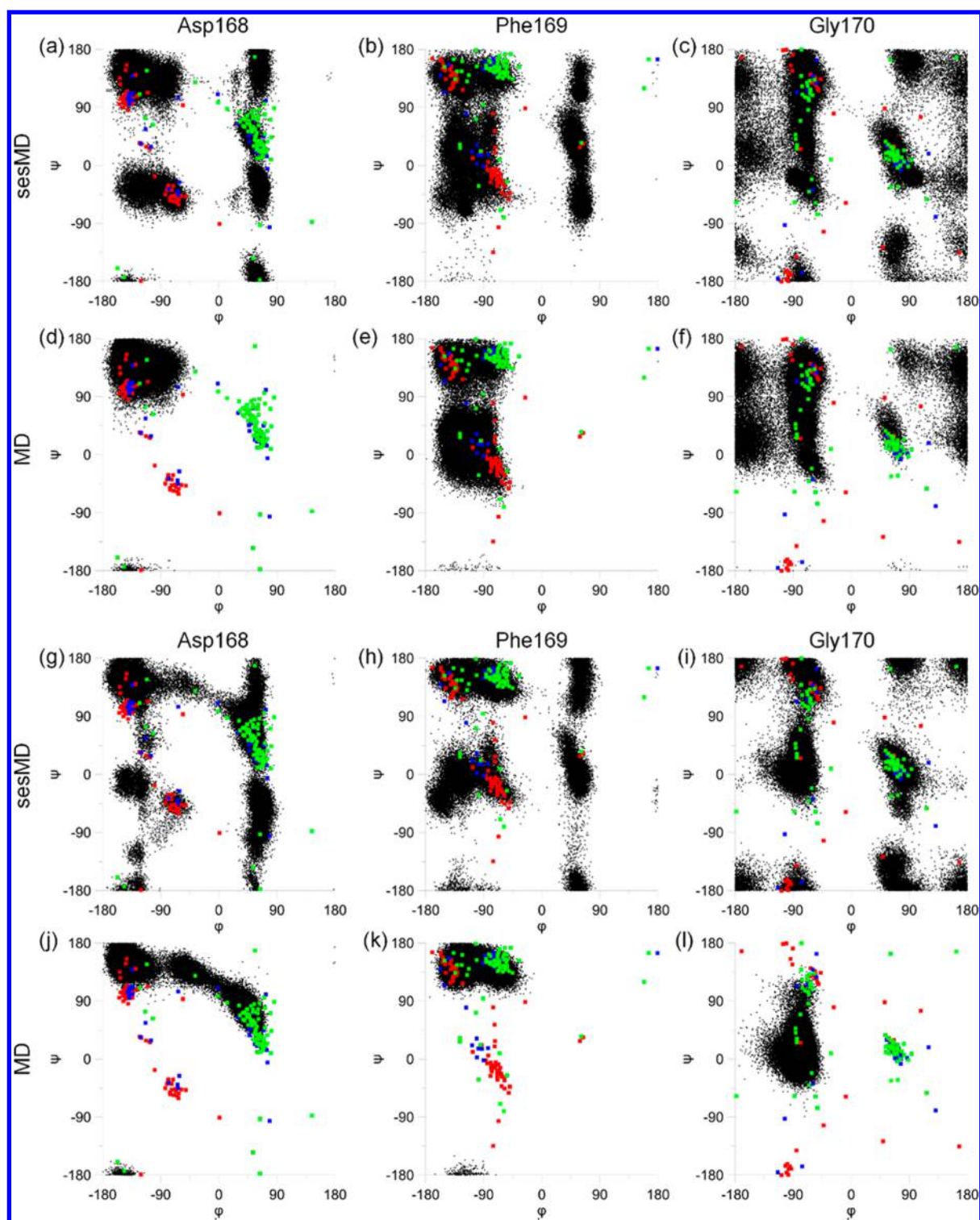
Figure 3. Ramachandran plots of the φ and ψ angles for each DFG residue sampled by 24 replicas of sesMD simulations (black) starting from p38α MAP kinase crystal structures 1WBT (a–c) and 1P38 (g–i) and by 24 independent MD simulations (d–f and j–l). Also shown are φ and ψ angles of 61 DFG-out (red), 74 DFG-in (green), and 33 unclassified (blue) p38α MAP kinase X-ray structures as defined in the MOE Kinase Explorer Database.

parameters A–D (eq 2) were increased linearly from zero to their final values, followed by 500 ps of sesMD at the final A–D values; these parameters were then reduced to zero over 500 ps, and 3.5 ns of unbiased MD was then performed. The 24 independent 5 ns MD simulations used for comparison with the sesMD simulations were initiated from the configurations

immediately prior to the first application of the ses potential detailed above. All of the MD and sesMD simulations were performed using a modified version of *sander* from the Amber 11 molecular simulation package.[35] Configurations were archived every 2 ps for analysis.
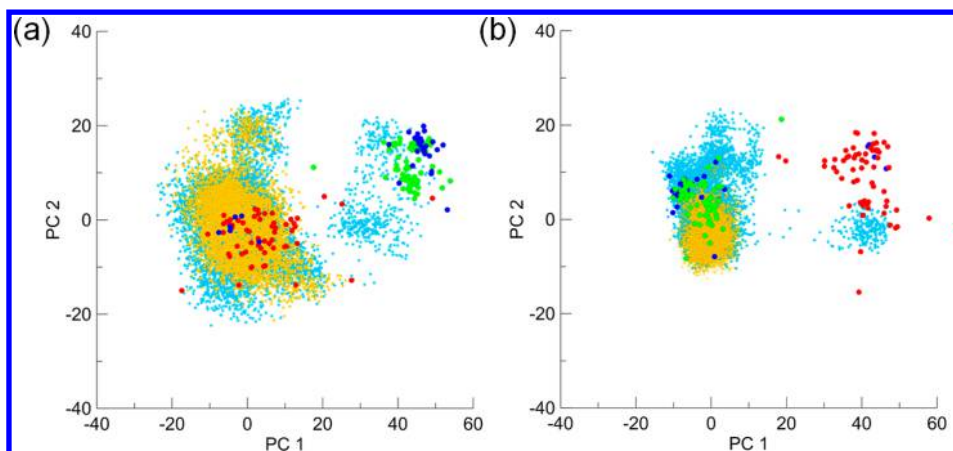
Figure 4. Projections of MD (yellow) and sesMD (cyan) ensembles of kinase and DFG-out (red), DFG-in (green), and unclassified (dark blue) p38α MAP kinase crystal structures onto the space defined by the top two principal components based on simulations starting from the (a) 1WBT and (b) 1P38 crystal structures.

For principal component analysis (PCA), conformations were taken from every 20 ps of each sesMD simulation replica's trajectory and fitted onto the Cα atoms of an average conformation calculated across all replicas in that simulation. With these fitted conformations, PCA of the atomic displacements of the DFG residues was performed using the *ptraj* module of Amber 11. Clustering analysis of the DFG residues was performed using the *kclust* algorithm from the MMTSB toolset[39] with a fixed clustering radius of 4.0 Å. Pocket detection and druggability analysis were performed using PocketAnalyzer^{PCA} [40] and Site Finder[41] from MOE.[42]

■ RESULTS

We evaluated the ability of sesMD to sample the conformational plasticity of unliganded p38α MAP kinase initiated from either a DFG-out crystal structure (1WBT) or a DFG-in crystal structure (1P38). As described in Methods, we applied the ses potential of eq 2 to the six backbone torsions linking the DFG loop amino acids, performing 5 ns of sesMD simulation using 24 replicas of the kinase in explicit solvent. These 5 ns of simulation comprise 1.5 ns under the influence of the ses potential followed by 3.5 ns of unbiased MD. In the following discussion, we compare these sesMD trajectories with 24 independent 5 ns MD simulations of the kinase in explicit solvent.

Sampling from DFG-out and DFG-in Crystal Structures. We consider the ability of MD and sesMD to sample the backbone φ and ψ torsions for the three residues of the key DFG motif of the kinase. As a reference, on the basis of the annotation of the MOE Kinase Explorer Database,[42] in Figure 3 we plot the φ and ψ angles of Asp168, Phe169 and Gly170 from the 74 DFG-in p38α MAP kinase crystal structures (green), the 61 DFG-out structures (red), and the 33 structures denoted as "unclassified" (blue). The DFG-in and DFG-out distributions of crystallographic φ and ψ angles are largely distinct for all three residues of the loop, but for Asp168 in particular (Figure 3a). Thus, the φ and ψ angles of Asp168 in the DFG-out 1WBT structure are (−139°, 111°), which sit within the DFG-out cluster around the β-sheet region (Figure 3a); in contrast, for the 1P38 DFG-in structure the φ and ψ angles are (32°, 68°), lying within the narrow left-handed α-helical region of DFG-in structures of the map (Figure 3a).

Unbiased MD simulations initiated from the 1WBT structure appear only to explore (φ, ψ) values in the locality of its initial backbone conformation (Figure 3d). However, application of the ses potential leads additionally to exploration of the DFG-out (φ, ψ) cluster in the right-handed α-helical region and notably around the DFG-in cluster as well (Figure 3a). Broader sampling by sesMD over MD is also observed for the backbone torsions of Phe169 (Figure 3b vs Figure 3e) and Gly170 (Figure 3c vs Figure 3f); generally, the crystal φ and ψ values map onto the conformations sampled by sesMD, although some simulated conformations do encompass regions of the Ramachandran map with few crystal structures (e.g., φ = 90° for Phe169 and φ = 180° for Gly170).

Interestingly, for MD simulations initiated from the 1P38 DFG-in structure, (φ, ψ) values of Asp168 for both DFG-in and DFG-out regions are explored (Figure 3j). However, once again, sesMD provides more complete coverage of (φ, ψ) space (Figure 3g−i), such that, for example, DFG-in (φ, ψ) values in the region of (−90°, −45°) are explored for Asp168 and Phe168 (Figure 3g,h) and in the vicinity of (90°, 0°) for Gly169 (Figure 3i).

To provide a broader view of the sampling of the DFG loop conformation, PCAs were performed on the heavy atoms of the DFG mainchain and sidechain atoms on the basis of the 168 X-ray crystal structures in combination with the 1WBT trajectory (Figure 4a) or the 1P38 trajectory (Figure 4b) from the MD or sesMD methods. The two top principal components from these PCAs account for 75% and 76% of the variance in atomic displacements of the DFG motif in the 1WBT and 1P38 simulations, respectively. Projection of the 168 X-ray crystal structures onto the axes defined by these first two principal components (Figure 4a,b) results in two distinct clusters, separating the DFG-out (red) and DFG-in (green) conformations. Conformations unclassified by the MOE Kinase Explorer Database (dark blue) to some degree occupy both clusters.

It can also be seen that unbiased molecular dynamics simulations (yellow) explore only locally around the initial structures of the trajectories for either 1WBT (Figure 4a) or 1P38 simulations (Figure 4b). By contrast, sesMD simulations of 1WBT and 1P38 (cyan) explore more extensively, in both cases populating regions in the principal component space
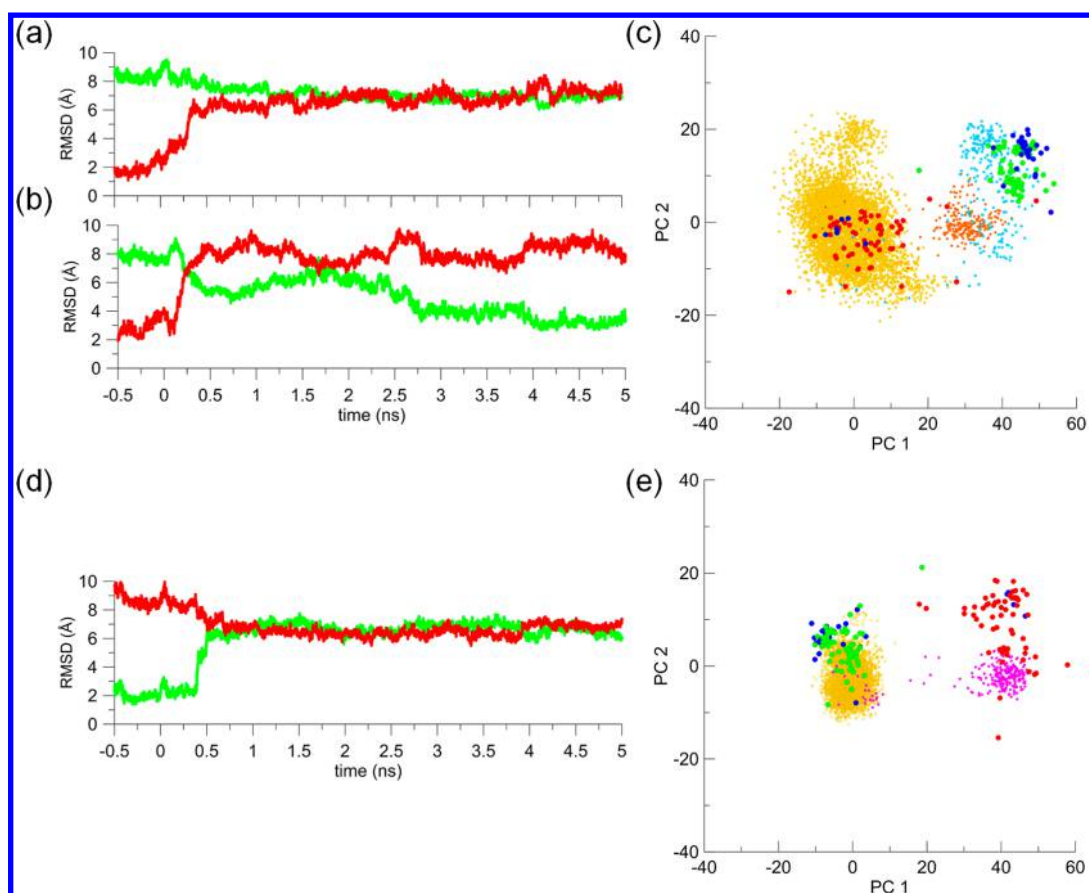
Figure 5. (a, b, d) Time series of heavy-atom RMSDs (in Å) of the DFG motif from sesMD simulations of p38α MAP kinase for (a) replica 1_{out}, (b) replica 2_{out}, and (d) replica 1_{in} calculated against the DFG-out 1WBT (red) and DFG-in 1P38 (green) crystal structures during pre-sesMD equilibration (-0.5 to 0.0 ns) and sesMD trajectories (0.0 to 5.0 ns). (c, e) Projections onto the first two principal components of sesMD configurations of (c) replicas 1_{out} (orange) and 2_{out} (cyan) and (e) replica 1_{in} (magenta), alongside PC values for 24 MD replicas (yellow) and DFG-in (green), DFG-out (red), and unclassified (dark blue) crystal structures.

around the initial DFG geometry and the opposing DFG orientation.

Further analysis of the sesMD trajectories indicates that the PC space most remote from the initial structures is sampled by a minority of its 24 replicas. This is shown by analysis of the RMSD of the atomic positions of the DFG-loop heavy atoms of each of the 24 replicas with respect to the 1WBT and 1P38 X-ray structures when aligned according to the overall protein conformation (Figures 2S and 3S in the Supporting Information). For the majority of the replicas in the 1WBT DFG-out simulation, the RMSD values are ~2−3 Å from this initial structure (Figure 2S), although for two replicas that we have labeled 1_{out} and 2_{out}, the RMSD relative to 1WBT rises to ~7 Å (Figure 5a,b respectively). Similarly, for sesMD simulations from 1P38, the majority of the conformations sample locally (Figure 3S); however, for one replica of the 1P38 simulation, which we have labeled 1_{in}, the RMSD rises to ~7 Å from 1P38 (Figure 5d). From a chart of the conformations sampled in the defined PC space by replicas 1_{out}, 2_{out}, and 1_{in}, we observe that they occupy the majority of the remotely sampled regions in the PCA (orange and cyan, Figure 5c; magenta, Figure 5e).

For all three of these simulation replicas, the largest change in RMSD occurs during the application of the ses potential rather than during the free MD phase of the simulations. More specifically, it occurs during the phase in which there is a gradual increase of the ses potential, i.e., during the interval

from 0 to 0.5 ns (Figure 5a,b,d). We may also examine the RMSDs of the DFG loops of these replicas with respect to their opposing crystal structure conformations. For replica 1_{out}, the RMSD remains high with respect to the 1P38 structure throughout the simulation, dropping only slightly from 8 to 6 Å (green, Figure 5a). A similar pattern is observed for replica 1_{in} of the 1P38 sesMD simulation (Figure 5d). Indeed, the high RMSDs of replicas 1_{out} and 1_{in} with respect to the DFG-in and DFG-out crystal structures is suggestive of the adoption of an intermediate DFG geometry. However, for replica 2_{out}, the RMSD drops over the course of the simulation to a value of ~3−4 Å (Figure 5b).

In order to obtain greater molecular detail of the conformational change in these replicas, we consider two inter-residue distances that have been suggested to be useful diagnostic measures of DFG-in and DFG-out conformations:[32] the distance between the side-chain centroids of Asp168 and Lys53 (neighboring the glycine-rich loop) and the distance between the side-chain centroids of Phe169 and Leu74 (located in the αC-helix). For the 1WBT DFG-out crystal structure, these distances are 13.5 and 17.8 Å respectively, and for the 1P38 DFG-in structure, the values are 11.8 and 6.9 Å respectively. Thus, from DFG-out to DFG-in, the considerable shortening of the Phe169−Leu74 distance and lesser reduction of the Asp168−Lys53 distance indicates that Phe169 now occupies the kinase deep pocket and Asp168 has correspondingly swung into the active site. For the first 2 ns of sesMD of

2769

dx.doi.org/10.1021/ci5003334 | J. Chem. Inf. Model. 2014, 54, 2764−2775

1WBT replicas 1_{out} and 2_{out}, we observe a transient deviation of the Asp168−Lys53 distance from its crystallographic value (Figure 6a) and a drop in the Phe169−Leu74 distance to ∼11
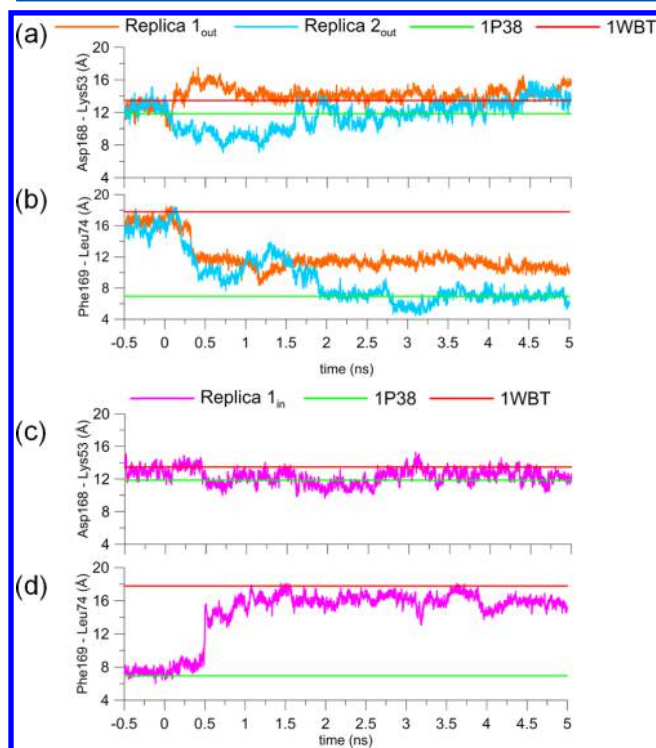

Figure 6. (a, c) Time series of the distance between the side-chain centroids of Asp168 and Lys53 for (a) replicas 1_{out} (orange) and 2_{out} (cyan) and (c) replica 1_{in} (magenta) of p38α MAP kinase from pre-sesMD equilibration (−0.5 to 0.0 ns) and sesMD trajectories (0.0 to 5.0 ns). (b, d) Time series of the distance between the side-chain centroids of Phe169 and Leu74 of (b) replicas 1_{out} (orange) and 2_{out} (cyan) and (d) 1_{in} (magenta) from pre-sesMD equilibration and sesMD trajectories. The X-ray distances for the 1WBT (red) and 1P38 (green) structures are also shown for reference.

Å (Figure 6b). However, whereas replica 1_{out} remains at a Phe169−Leu74 distance of ∼11 Å, replica 2_{out} further decreases to a final value of ∼7 Å, the crystallographic value for the DFG-in structure (Figure 6b). For replica 1_{in} of 1P38, local conformational sampling is observed around the Asp168−Lys53 X-ray distance (Figure 6c). By contrast, a large change in the Phe169−Leu74 distance is observed to stabilize in proximity to the DFG-out X-ray value (Figure 6d).

We can visually trace the conformational motion of the DFG loop in the three replicas by considering snapshots from their trajectories (Figure 7). The DFG-out conformation of the 1WBT crystal structure (red) and the DFG-in pose of the 1P38 crystal structure (green) bracket conformations sampled over the sesMD simulations. These sesMD snapshots show the time progression over the 5 ns trajectory (from orange to yellow to cyan to blue). The trajectories for both replicas 1_{out} and 2_{out} (Figure 7a,b, respectively) show movement from their initial DFG-out structures (red) toward the opposing DFG-in conformations (green). However, whereas replica 1_{out} is arrested in an intermediate structure (Figure 7a), replica 2_{out} completes the transition to a DFG-in geometry with Phe169 occupying the deep pocket (Figure 7b) and mapping onto DFG-in orientations found in the apo (1P38) and liganded (2HVC and 3LGC)[43,44] DFG-in crystal structures (Figure 8).
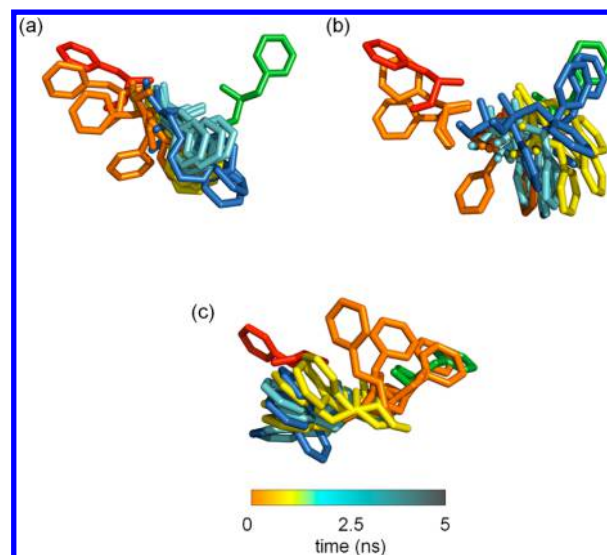

Figure 7. Conformation of Phe169 in selected snapshots from (a) replica 1_{out}, (b) replica 2_{out}, and (c) replica 1_{in} from sesMD simulations of p38α MAP kinase. The conformations from sesMD are color-coded according to the simulation time/stage: 0.5−1 ns (orange), 1−1.5 ns (yellow), 1.5−2 ns (cyan), 2−5 ns (blue). Also shown are the DFG-in 1P38 (green) and DFG-out 1WBT (red) crystal structures.
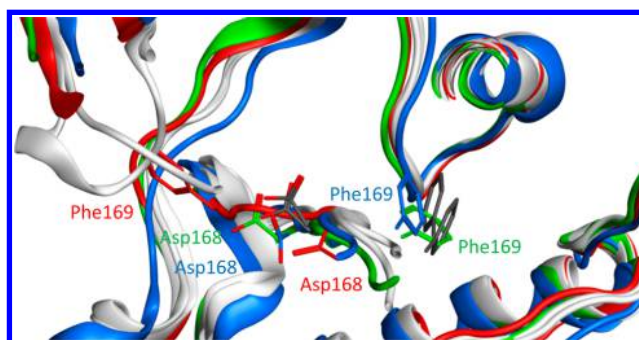

Figure 8. Final conformation from sesMD of replica 2_{out} showing the Asp168 and Phe169 side chains in the p38α MAP kinase active site (blue), compared with its initial pose from the DFG-out 1WBT X-ray structure (red). The DFG-in crystal structures 1P38 (green) and 2LGC and 3HVC (both gray) are also shown.

This structural transition is more completely illustrated by a movie of the trajectory (movie S1 in the Supporting Information). The DFG-out to DFG-in transition contrasts with a control simulation we ran from the same initial conditions of replica 2_{out} using 50 ns of unbiased MD. Here no transition is observed, but instead the kinase remains in a local DFG-out conformation (movie S2 in the Supporting Information).

For replica 1_{in}, a partial transition from DFG-in to DFG-out is observed, such that there is a significant shift in the Phe169 conformation (Figure 7c). However, Asp168 remains in a local well around its initial structure, maintaining a polar interaction with the backbone of Asn152. For replicas 1_{out}, 2_{out}, and 1_{in}, the largest movement of the DFG loop is associated with the period where the ses potential is increasing to its full strength over the interval 0.0 to 0.5 ns (orange in Figure 7). However, also particularly important for the replica 2_{out} simulation is the final 3.5 ns of nonbiased MD, during which the activation loop relaxes to a structure within the manifold of DFG-in loop conformations that are observed crystallographically (Figure 4S
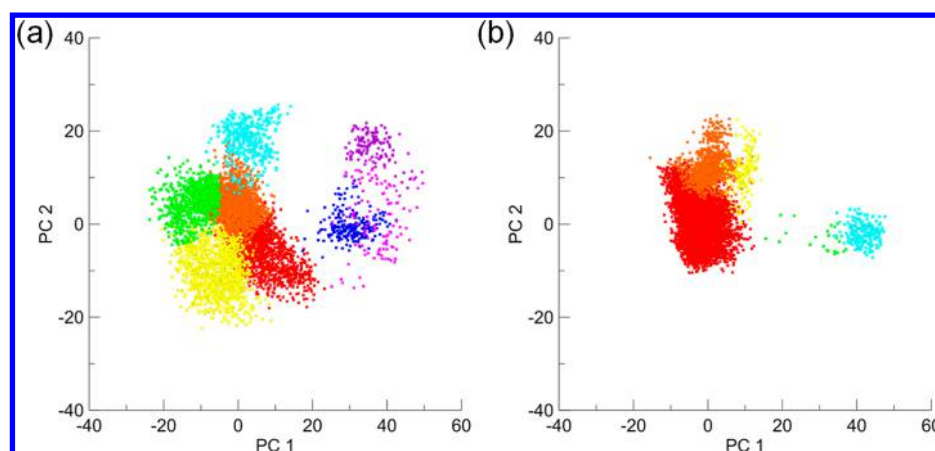
Figure 9. Projections onto the space defined by the first two principal components of the ensemble of 24 sesMD simulation replicas, initiated from the (a) 1WBT and (b) 1P38 crystal structures. The projections are color-coded according to cluster: $C1_{out}$ and $C1_{in}$ (red), $C2_{out}$ and $C2_{in}$ (orange), $C3_{out}$ and $C3_{in}$ (yellow), $C4_{out}$ and $C4_{in}$ (green), $C5_{out}$ and $C5_{in}$ (cyan), $C6_{out}$ (blue), $C7_{out}$ (magenta), and $C8_{out}$ (purple).

in the Supporting Information). A relaxation of the swarm potential was found to be similarly important in the final stages of folding of the Trp-cage miniprotein by SWARM-MD.[18]

Conformations and Druggability of Intermediate Structures of the DFG Loop. It is clear that a range of DFG loop conformations is sampled during sesMD of p38α MAP kinase (e.g., Figure 7), some of which may present opportunities for molecular design. To this end, we performed conformational clustering of the combined trajectories of all 24 sesMD replicas for the 1WBT and 1P38 simulations separately. Adopting a fixed clustering radius of 4.0 Å, we identified eight clusters of distinct conformations explored by the 1WBT sesMD replicas (labeled $C1_{out}$−$C8_{out}$) and five clusters of 1P38 sesMD conformations ($C1_{in}$−$C5_{in}$). From the location of these structures in PC space, we classify the clusters into DFG-out-like and DFG-in-like conformations for 1WBT (Figure 9a) and 1P38 simulations (Figure 9b). As might be expected, the most populated clusters are localized around the parent X-ray conformation. Thus, for the 1WBT sesMD simulation, clusters $C1_{out}$−$C4_{out}$, representing 85% of the total sesMD configurations (Table 1), occupy DFG-out-like conformations (left-hand side of Figures 9a and 10); however, the less populated $C5_{out}$−$C8_{out}$ clusters occupy a spectrum toward a DFG-in structure (right-hand side of Figures 9a and 10). For the 1P38 simulation, 80% of the conformations fall into a $C1_{in}$ DFG-in-like conformation (left-hand side of Figure 9b, right-hand side of Figure 10, and Table 1). Clusters $C2_{in}$ and $C3_{in}$, representing ~17% of the conformations sampled, are characterized by a displacement of the Phe169 side chain towards the αC helix. Clusters $C4_{in}$ and $C5_{in}$ group a smaller fraction of the ensemble, belonging only to replica 1_{in}, that adopt the intermediate state described previously (Figure 7c).

From inspection, we observe similarities in the transition structures sampled by the 1WBT sesMD simulation and crystal structures of p38α MAP kinase with intermediate DFG loop conformations. This is achieved despite the fact that in these crystal structures the kinase was complexed with ligand, whereas here we performed sesMD of the kinase in the absence of ligand. Good agreement is seen in the DFG conformations of 15 crystal structures of the kinase and that of replica 1_{out} (Figure 11a). For example, the sesMD conformations of Asp168 and Phe169 residues in this replica superimpose to within an RMSD of 0.4 Å with p38α MAP

Table 1. Populations (pop) and Average Volumes (V_{av}) of the Binding Pockets Identified by PocketAnalyzer[PCA] for Each Cluster, Along with Approximate Classification of Its DFG Conformation (DFG conf) and Propensity for Ligand Binding (PLB) of Structures Possessing the Lowest Distance from Each Cluster Centroid, As Calculated by Site Finder

cluster	DFG conf	pop (%)	V_{av} (Å^3)[a]	PLB
$C1_{out}$	out	25.2	1212 (347)	3.9
$C2_{out}$	out	23.9	1044 (255)	3.0
$C3_{out}$	out	18.8	1334 (294)	4.3
$C4_{out}$	out	17.6	1255 (225)	4.5
$C5_{out}$	intermed	6.6	866 (278)	3.8
$C6_{out}$	intermed	3.9	712 (277)	3.5
$C7_{out}$	in	1.9	1181 (294)	2.8
$C8_{out}$	in	2.0	1071 (291)	2.7
$C1_{in}$	in	79.6	880 (260)	3.0
$C2_{in}$	in	14.0	686 (211)	2.9
$C3_{in}$	in	3.5	955 (242)	4.0
$C4_{in}$	intermed	2.6	770 (183)	4.5
$C5_{in}$	intermed	0.4	1104 (213)	4.4
DFG-in X-ray av	in	–	736 (246)	–
DFG-out X-ray av	out	–	784 (316)	–
1P38	in	–	989	2.7
1WBT	out	–	1440	3.6

[a]Numbers in parentheses are standard deviations.

kinase in cocrystals with pyrrolotriazine compounds (PDB codes 3BV2 and 3BV3; Figure 11b).[45]

Furthermore, other p38α MAP kinase transitional structures from sesMD bear a resemblance to intermediates identified from a high-temperature restrained MD study of the kinase.[27] In that work, pseudo-DFG-in and pseudo-DFG-out transition structures of p38α MAP kinase were defined, both lying at RMSDs of ~6−8 Å from DFG-in and DFG-out X-ray structures.[27] These high-temperature conformations were characterized by a β-turn-like structure, with the side chain of Leu171 occupying the location normally occupied by Phe169 in a DFG-in (denoted pseudo-DFG-in) or DFG-out (pseudo-DFG-out) conformation. From the 1P38 simulation, we obtained sampling of structures akin to a pseudo-DFG-out conformation from sesMD, such that the side chain of Leu171 replaces the position previously occupied by the aromatic ring of Phe169 (Figure 11c); however, only a partial pseudo-DFG-in
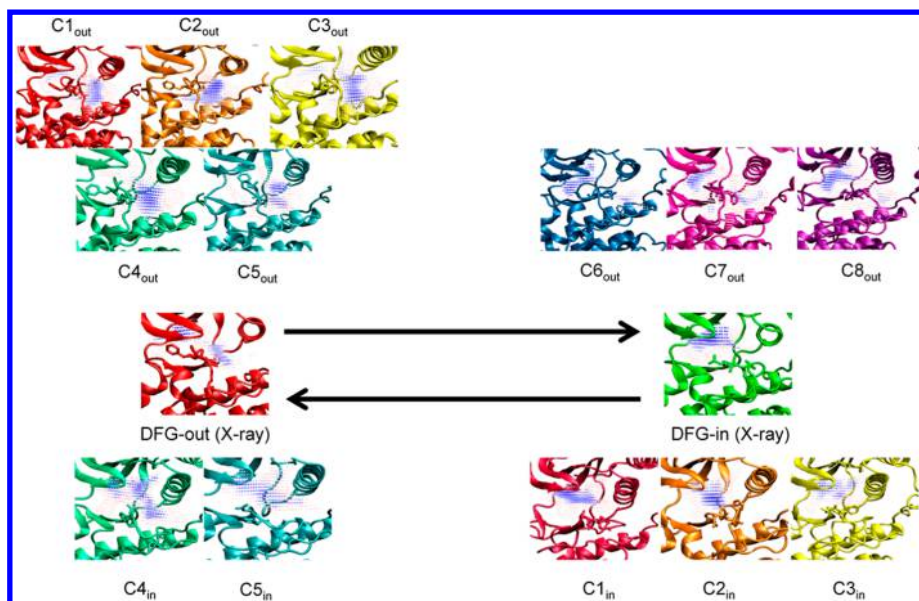
Figure 10. Average binding pocket topologies found in structures within conformational clusters identified from sesMD simulations of p38α MAP kinase (using PocketAnalyzer[PCA]), mapped onto structures nearest to the cluster centroid. Each structure is color-coded according to cluster (see the caption of Figure 9). The pocket is color-coded according to the weighted average (high frequency, blue; low frequency, red). As references, sets of DFG-in and DFG-out crystallographic structures were used.
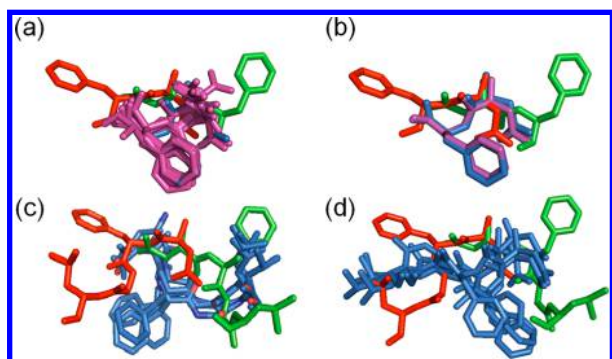

Figure 11. (a, b) Superpositions of a selected p38α MAP kinase conformation from sesMD replica 1_{out} (blue) with (a) 15 p38α MAP kinase X-ray structures (purple; for PDB codes, see Table 1S in the Supporting Information) and (b) X-ray structures 3BV2 and 3BV3 (purple). (c, d) Superpositions of selected intermediate DFG conformations (blue) from sesMD of (c) replica 1_{out} and (d) replica 1_{in}. For reference, DFG-in (1P38, green) and DFG-out (1WBT, red) X-ray structures are also shown.

structure was found from either the 1WBT or 1P38 simulation (Figure 11d).

In assessing the opportunities for structure-based design against p38α MAP kinase, it is useful to characterize the pockets formed by the clustered sesMD conformations. For reference, we first consider the average pocket shapes arising from the 74 DFG-in and 61 DFG-out crystal structures: the location of the additional deep pocket is clearly shown in X-ray DFG-out structures relative to DFG-in (Figure 10). However, on average the total pocket volumes in the DFG-out and DFG-in crystal structure sets are very similar, with values of 784 and 736 Å^3, respectively (Table 1). This hides considerable variability; for example, the volumes of pockets in 1WBT and 1P38 are 1440 and 989 Å^3, respectively (Table 1).

A range of pocket shapes (Figure 10) and volumes (Table 1) are exhibited by the conformational clusters of p38α MAP kinase from sesMD, i.e., $C1_{out}$–$C8_{out}$ and $C1_{in}$–$C5_{in}$. The

pockets in $C1_{out}$–$C4_{out}$ display a dual-lobed density, reflecting the volumes of the hinge and deep pockets, whereas structures representative of $C1_{in}$–$C3_{in}$ possess a single pocket in the hinge region (Figure 10). These pocket shapes are akin to those found in the parent X-ray structures (Figure 10). Interestingly, the $C5_{out}$ kinase conformation has a pocket shape distinct from that of the $C1_{out}$–$C4_{out}$ and X-ray structures. This cluster conformation shows an orientation of Phe169 comparable to that encountered on one of two transition pathways sampled via the aforementioned high-temperature MD study.[27] Clusters $C6_{out}$–$C8_{out}$ clearly show diminished deep pockets and expanded cavities in the hinge region typical of the average crystallographic DFG-in pocket (Figure 10). For 1P38 sesMD conformations $C4_{in}$ and $C5_{in}$, there is an opening up of the deep pocket region that is comparable to the topology of the DFG-out pocket (Figure 10). Indeed, cluster $C5_{in}$ has the largest average pocket volume of all the 1P38-derived clusters identified and closest in volume to that of the 1WBT structure (Table 1).

To quantify the potential of these pockets to bind small molecules, we performed an assessment of their druggability. Specifically, for the kinase structures closest to the cluster centroids, we computed a propensity for ligand binding (PLB) index[46] for pockets defined by Site Finder (we note that the kinase pocket topologies from Site Finder are broadly comparable to those obtained by Pocketanalyzer[PCA]). The PLB index is based on the specific amino acid composition of a pocket and has proven effective in predicting drug binding propensities for known protein structures.[46] Here, we first consider the PLB indices for 1WBT and 1P38 X-ray structures, which have values of 3.6 and 2.7, respectively (Table 1). These large positive PLB values indicate that both sites are predicted to be druggable, as expected; the DFG-out 1WBT structure has a larger index, in part a function of the larger pocket volume compared with 1P38 (Table 1).

This trend is also reproduced by the PLB scores predicted for the cluster centroid structures: sesMD structures based on

1WBT show a reduced PLB index by switching to a DFG-in structure, such that the PLP index of 3.9 for $C1_{out}$ decreases to 2.7 for $C8_{out}$ (Table 1); the opposite trend is observed for 1P38 sesMD structures progressing toward a DFG-out state (compare $C5_{in}$ with $C1_{in}$; Table 1). Interestingly, PLB indices greater than 4 are obtained from intermediate structures on either of these pathways, such as $C4_{in}$, $C5_{in}$ and $C5_{out}$ (Table 1). These structures could represent interesting starting points for structure-based inhibitor design.

■ DISCUSSION AND CONCLUSIONS

From the preceding analysis, it is evident that application of swarm-enhanced sampling MD to the DFG motif of p38α MAP kinase provides a broader sampling of DFG conformational states compared with unbiased MD simulations. From sesMD of an initially DFG-out kinase conformation, we observe sampling of a range of loop conformations that includes DFG-in structures (Figure 8) and intermediate geometries (Figure 11a−c) resembling p38α MAP kinase conformations identified previously from crystallography[45] and computation.[27] From sesMD based on a DFG-in crystal structure, a range of DFG loop conformations is sampled, although to a lesser extent than for the simulation based on a DFG-out conformation. In particular, while a partial transition is observed for Phe169, Asp168 remains localized in an interaction with the backbone of Asn152. This hydrogen bond may point to the implication of the protonation state of Asp168 in conformational change, as suggested in work by Shan et al. on Abl kinase.[28] In both 1P38 and 1WBT sesMD simulations, we identify interesting intermediate pocket shapes, for example $C4_{in}$ and $C5_{out}$ pockets, that could be useful in identifying novel kinase inhibitors via virtual screening.

SesMD is a biased MD approach using multiple simulation replicas. While the concept of coupling replicas within a multicopy MD framework to improve sampling is not new,[6,10,47−52] in sesMD we have implemented a novel form of the potential between replicas, incorporating attractive and repulsive terms. This swarm of simulation replicas could be viewed as somewhat akin to a swarm of van der Waals particles, simultaneously exploring mutually exclusive regions of phase space. An alternative multicopy MD approach to enhanced sampling is the replica exchange method (REM),[12] where at periodic intervals neighboring replicas may be swapped between simulations. The key choice is in the proximity of the replicas in terms of temperature (for temperature REM) or potential (for Hamiltonian REM) such that efficient exchange can occur in accord with the requirement of detailed balance, without requiring too many replicas to span the desired range of temperature or potential. This has proven to be an issue for large explicitly solvated solutes, which require a large number of replicas to achieve this overlap. This problem can be mitigated by using a new implementation of the replica exchange with solute tempering (REST2) method,[53] where the replica Hamitonians are scaled in order to remove the dependence on the number of explicit water molecules. SesMD is not subject to the energy overlap requirement of REM. However, in its current implementation, sesMD exchanges coordinate/dihedral information between all replicas at each time step, in contrast to REM, which compares only neighboring replicas typically at every picosecond. As a result of this communication in sesMD, for a given time step, the decrease in speed of sesMD compared with the equivalent set of independent MD trajectories is ∼38%. However, for the 5 ns mixed sesMD/

MD protocol applied here to p38α MAP kinase, an overall decrease in speed of only 14% was found, while the degree of sampling within the computed 5 ns was markedly improved for sesMD relative to 24 independent MD trajectories of 5 ns or a single 50 ns trajectory.

From REM methods, as with the popular single-trajectory enhanced sampling approaches of accelerated MD[11] and metadynamics,[15,16] it is possible to obtain Boltzmann-weighted ensemble properties directly or indirectly. This is also formally the case for sesMD: because of the well-defined Hamiltonian of the overall swarm of trajectories, one can reweight the swarm of replicas according to the approach of Torrie and Valleau[7] (see the Supporting Information). We do not apply this approach here, where we scale the ses potential in seeking to drive conformational exploration of the kinase. Nevertheless, the general topology of the underlying unbiased potential energy surface appears to be preserved for sesMD of the model butane and alanine dipeptide systems (Figure 2b,d), and experimentally observed intermediate DFG conformations of the kinase are obtained during sesMD (Figure 11a,b).

We note that other biased MD approaches have been employed to study kinase conformational change: for example, targeted MD simulations have been used to study the structural reorganization of Abl kinase[54] and EGFR kinase.[31] However, distinct from sesMD, a priori knowledge of the alternative protein conformation is required for targeted MD. Another biased MD study performed nonequilibrium MD simulations of p38α MAP kinase in order to explore DFG conformation.[26] Steering forces were applied to the Asp168 and Phe169 side chains of the DFG loop; the direction and magnitude of the pulling forces required recalibration at different stages in the conformational change. As with this nonequilibrium MD approach and the metadynamics method, for sesMD a choice must be made as to the degrees of freedom to which to apply the biasing potential, although in the case of the kinase this is intuitive, i.e., DFG backbone torsions. Potentially, for a given protein crystal structure, the choice of the amino acid dihedral set to explore by sesMD could be guided by experimental insight, for example from active-site B-factors, or systematically mapped for active-site residues. Accordingly, it will be important to assess the level of transferability of ses parameters to other systems of interest, for small and larger molecules. We note that careful system-dependent parametrization is required for other biasing methods such as metadynamics and accelerated molecular dynamics. For the latter, to obtain suitable parameters, an initial estimate of the energy landscape to be explored is obtained from simulation in the absence of biasing potential prior to trialling short biased simulations. We anticipate that future work will provide insight into the required parameter ranges and protocols for optimal sesMD applications.

SesMD therefore constitutes a straightforward, parallel and promising approach to conformational sampling. Here we have applied sesMD to the generation of a range of possible alternative conformations of a protein based on a single crystal structure. Such a strategy applied to a therapeutic target protein could reveal novel protein concavities of potential use in providing new directions for structure-based inhibitor design.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Computational details for sesMD simulations of pentane and alanine dipeptide; a scheme for reweighting properties obtained

from sesMD simulations; and further information on sesMD of p38α MAP kinase (biasing potentials, trajectory of replicas, activation loop conformation). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*Tel: +44 161 275 8345. Fax: +44 161 275 2396. E-mail: r.a. bryce@manchester.ac.uk.

Present Address
⊥N.J.B.: Heidelberg Institute for Theoretical Studies (HITS gGmbH), Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany.

Author Contributions
∥A.A. and N.J.B. contributed equally to this work.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Schames, J. R.; Henchman, R. H.; Siegel, J. S.; Sotriffer, C. A.; Ni, H.; McCammon, J. A. Discovery of a Novel Binding Trench in HIV Integrase. *J. Med. Chem.* **2004**, *47*, 1879−1881.

(2) Landon, M. R.; Amaro, R. E.; Baron, R.; Ngan, C. H.; Ozonoff, D.; McCammon, J. A.; Vajda, S. Novel Druggable Hot Spots in Avian Influenza Neuraminidase H5N1 Revealed by Computational Solvent Mapping of a Reduced and Representative Receptor Ensemble. *Chem. Biol. Drug Des.* **2008**, *71*, 106−116.

(3) Mitchell, F. L.; Miles, S. M.; Neres, J.; Bichenkova, E. V.; Bryce, R. A. Tryptophan as a Molecular Shovel in the Glycosyl Transfer Activity of *Trypanosoma cruzi* Trans-Sialidase. *Biophys. J.* **2010**, *98*, L38−L40.

(4) Mitchell, F. L.; Neres, J.; Ramraj, A.; Raju, R. K.; Hillier, I. H.; Vincent, M. A.; Bryce, R. A. Insights into the Activity and Specificity of *Trypanosoma cruzi* Trans-Sialidase from Molecular Dynamics Simulations. *Biochemistry* **2013**, *52*, 3740−3751.

(5) Demir, O.; Roitberg, A. E. Modulation of Catalytic Function by Differential Plasticity of the Active Site: Case Study of *Trypanosoma cruzi* Trans-Sialidase and *Trypanosoma rangeli* Sialidase. *Biochemistry* **2009**, *48*, 3398−3406.

(6) Adcock, S. A.; McCammon, J. A. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* **2006**, *106*, 1589−1615.

(7) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(8) Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Schulten, K. Molecular Dynamics Study of Unbinding of the Avidin−Biotin Complex. *Biophys. J.* **1997**, *72*, 1568−1581.

(9) Withers, I. M.; Mazanetz, M. P.; Wang, H.; Fischer, P. M.; Laughton, C. A. Active Site Pressurization: A New Tool for Structure-Guided Drug Design and Other Studies of Protein Flexibility. *J. Chem. Inf. Model.* **2008**, *48*, 1448−1454.

(10) Simmerling, C. L.; Fox, T.; Kollman, P. A. Use of Locally Enhanced Sampling in Free Energy Calculations: Testing and Application to the α → β Anomerization of Glucose. *J. Am. Chem. Soc.* **1998**, *120*, 5771−5782.

(11) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919−11929.

(12) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(13) Huber, T.; Torda, A. E.; van Gunsteren, W. F. Local Elevation: A Method for Improving the Searching Properties of Molecular Dynamics Simulation. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695−708.

(14) Grubmuller, H. Predicting Slow Structural Transitions in Macromolecular Systems: Conformational Flooding. *Phys. Rev. E* **1995**, *52*, 2893−2906.

(15) Laio, A.; Gervasio, F. L. Metadynamics: A Method To Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science. *Rep. Prog. Phys.* **2008**, *71*, No. 126601.

(16) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.

(17) Huber, T.; van Gunsteren, W. F. SWARM-MD: Searching Conformational Space by Cooperative Molecular Dynamics. *J. Phys. Chem. A* **1998**, *102*, 5937−5943.

(18) Bruce, N. J.; Bryce, R. A. Ab Initio Protein Folding Using a Cooperative Swarm of Molecular Dynamics Trajectories. *J. Chem. Theory Comput.* **2010**, *6*, 1925−1930.

(19) Kornev, A. P.; Haste, N. M.; Taylor, S. S.; Ten Eyck, L. F. Surface Comparison of Active and Inactive Protein Kinases Identifies a Conserved Activation Mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17783−17788.

(20) Nagar, B.; Bornmann, W. G.; Pellicena, P.; Schindler, T.; Veach, D. R.; Miller, W. T.; Clarkson, B.; Kuriyan, J. Crystal Structures of the Kinase Domain of C-Abl in Complex with the Small Molecule Inhibitors PD173955 and Imatinib (STI-571). *Cancer Res.* **2002**, *62*, 4236−4243.

(21) Liu, Y.; Gray, N. S. Rational Design of Inhibitors That Bind to Inactive Kinase Conformations. *Nat. Chem. Biol.* **2006**, *2*, 358−364.

(22) Angell, R. M.; Angell, T. D.; Bamborough, P.; Bamford, M. J.; Chung, C.-w.; Cockerill, S. G.; Flack, S. S.; Jones, K. L.; Laine, D. I.; Longstaff, T. Biphenyl Amide P38 Kinase Inhibitors 4: DFG-In and DFG-Out Binding Modes. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4433−4437.

(23) Vogtherr, M.; Saxena, K.; Hoelder, S.; Grimme, S.; Betz, M.; Schieborr, U.; Pescatore, B.; Robin, M.; Delarbre, L.; Langer, T.; Wendt, K. U.; Schwalbe, H. NMR Characterization of Kinase P38 Dynamics in Free and Ligand-Bound Forms. *Angew. Chem., Int. Ed.* **2006**, *45*, 993−997.

(24) Bukhtiyarova, M.; Karpusas, M.; Northrop, K.; Namboodiri, H. V.; Springman, E. B. Mutagenesis of P38α MAP Kinase Establishes Key Roles of Phe169 in Function and Structural Dynamics and Reveals a Novel DFG-Out State. *Biochemistry* **2007**, *46*, 5687−5696.

(25) Berteotti, A.; Cavalli, A.; Branduardi, D.; Gervasio, F. L.; Recanatini, M.; Parrinello, M. Protein Conformational Transitions: The Closure Mechanism of a Kinase Explored by Atomistic Simulations. *J. Am. Chem. Soc.* **2008**, *131*, 244−250.

(26) Filomia, F.; De Rienzo, F.; Menziani, M. C. Insights into MAPK DFG Flip Mechanism by Accelerated Molecular Dynamics. *Bioorg. Med. Chem.* **2010**, *18*, 6805−6812.

(27) Frembgen-Kesner, T.; Elcock, A. H. Computational Sampling of a Cryptic Drug Binding Site in a Protein Receptor: Explicit Solvent Molecular Dynamics and Inhibitor Docking to P38 MAP Kinase. *J. Mol. Biol.* **2006**, *359*, 202−214.

(28) Shan, Y.; Seeliger, M. A.; Eastwood, M. P.; Frank, F.; Xu, H.; Jensen, M.; Dror, R. O.; Kuriyan, J.; Shaw, D. E. A Conserved Protonation-Dependent Switch Controls Drug Binding in the Abl Kinase. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 139−144.

(29) Shukla, D.; Meng, Y.; Roux, B.; Pande, V. S. Activation Pathway of Src Kinase Reveals Intermediate States as Targets for Drug Design. *Nat. Commun.* **2014**, *5*, No. 3397.

(30) Vashisth, H.; Maragliano, L.; Abrams, C. F. DFG-Flip in the Insulin Receptor Kinase Is Facilitated by a Helical Intermediate State of the Activation Loop. *Biophys. J.* **2012**, *102*, 1979−1987.

(31) Papakyriakou, A.; Vourloumis, D.; Tzortzatou-Stathopoulou, F.; Karpusas, M. Conformational Dynamics of the EGFR Kinase Domain Reveals Structural Features Involved in Activation. *Proteins* **2009**, *76*, 375−386.

(32) Lovera, S.; Sutto, L.; Boubeva, R.; Scapozza, L.; Dalker, N.; Gervasio, F. L. The Different Flexibility of C-Src and C-Abl Kinases Regulates the Accessibility of a Druggable Inactive Conformation. *J. Am. Chem. Soc.* **2012**, *134*, 2496−2499.

(33) Wang, Z.; Harkins, P. C.; Ulevitch, R. J.; Han, J.; Cobb, M. H.; Goldsmith, E. J. The Structure of Mitogen-Activated Protein Kinase P38 at 2.1 Å Resolution. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2327−2332.

(34) Gill, A. L.; Frederickson, M.; Cleasby, A.; Woodhead, S. J.; Carr, M. G.; Woodhead, A. J.; Walker, M. T.; Congreve, M. S.; Devine, L. A.; Tisi, D.; O'Reilly, M.; Seavers, L. C. A.; Davis, D. J.; Curry, J.; Anthony, R.; Padova, A.; Murray, C. W.; Carr, R. A. E.; Jhoti, H. Identification of Novel P38α MAP Kinase Inhibitors Using Fragment-Based Lead Generation. *J. Med. Chem.* **2005**, *48*, 414−426.

(35) Case, D. A.; Darden, T.; Cheatham, T. E., III; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M., Jr.; Roberts, B.; Wang, B.; Hayik, S. A.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, J.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T. K. P. A.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 11*; University of California: San Francisco, 2010.

(36) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712−725.

(37) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *N*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(38) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(39) Feig, M.; Karanicolas, J.; Brooks, C. L. MMTSB Tool Set: Enhanced Sampling and Multiscale Modeling Methods for Applications in Structural Biology. *J. Mol. Graphics Modell.* **2004**, *22*, 377−395.

(40) Craig, I. R.; Pfleger, C.; Gohlke, H.; Essex, J. W.; Spiegel, K. Pocket-Space Maps To Identify Novel Binding-Site Conformations in Proteins. *J. Chem. Inf. Model.* **2011**, *51*, 2666−2679.

(41) Liang, J.; Woodward, C.; Edelsbrunner, H. Anatomy of Protein Pockets and Cavities: Measurement of Binding Site Geometry and Implications for Ligand Design. *Protein Sci.* **1998**, *7*, 1884−1897.

(42) Chemical Computing Group MOE. Montreal, QC, Canada, 2003.

(43) Honndorf, V. S.; Coudevylle, N.; Laufer, S.; Becker, S.; Griesinger, C.; Habeck, M. Inferential NMR/X-ray-Based Structure Determination of a Dibenzo[*a,d*]Cycloheptenone Inhibitor−P38α MAP Kinase Complex in Solution. *Angew. Chem., Int. Ed.* **2012**, *51*, 2359−2362.

(44) Perry, J. J.; Harris, R. M.; Moiani, D.; Olson, A. J.; Tainer, J. A. P38α MAP Kinase C-Terminal Domain Binding Pocket Characterized by Crystallographic and Computational Analyses. *J. Mol. Biol.* **2009**, *391*, 1−11.

(45) Wrobleski, S. T.; Lin, S.; Hynes, J.; Wu, H.; Pitt, S.; Shen, D. R.; Zhang, R.; Gillooly, K. M.; Shuster, D. J.; McIntyre, K. W.; Doweyko, A. M.; Kish, K. F.; Tredup, J. A.; Duke, G. J.; Sack, J. S.; McKinnon, M.; Dodd, J.; Barrish, J. C.; Schieven, G. L.; Leftheris, K. Synthesis and SAR of New Pyrrolo[2,1-*f*][1,2,4]Triazines as Potent P38α MAP Kinase Inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2739−2744.

(46) Soga, S.; Shirai, H.; Kobori, M.; Hirayama, N. Use of Amino Acid Composition To Predict Ligand-Binding Sites. *J. Chem. Inf. Model.* **2007**, *47*, 400−406.

(47) Keasar, C.; Elber, R. Homology as a Tool in Optimization Problems: Structure Determination of 2D Heteropolymers. *J. Phys. Chem.* **1995**, *99*, 11550−11556.

(48) Stultz, C. M.; Karplus, M. MCSS Functionality Maps for a Flexible Protein. *Proteins: Struct., Funct., Genet.* **1999**, *37*, 512−529.

(49) Pan, A. C.; Sezer, D.; Roux, B. Finding Transition Pathways Using the String Method with Swarms of Trajectories. *J. Phys. Chem. B* **2008**, *112*, 3432−3440.

(50) Malevanets, A.; Wodak, S. J. Multiple Replica Repulsion Technique for Efficient Conformational Sampling of Biological Systems. *Biophys. J.* **2011**, *101*, 951−960.

(51) Shirts, M. R.; Pande, V. S. Mathematical Analysis of Coupled Parallel Simulations. *Phys. Rev. Lett.* **2001**, *86*, 4983−4987.

(52) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER6*;. University of California: San Fransisco, 1996.

(53) Wang, L.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J. Phys. Chem. B* **2011**, *115*, 9431−9438.

(54) Levinson, N. M.; Kuchment, O.; Shen, K.; Young, M. A.; Koldobskiy, M.; Karplus, M.; Cole, P. A.; Kuriyan, J. A Src-like Inactive Conformation in the Abl Tyrosine Kinase Domain. *PLoS Biol.* **2006**, *4*, 753−767.