JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Insights into the Conformational Switching Mechanism of the Human Vascular Endothelial Growth Factor Receptor Type 2 Kinase Domain

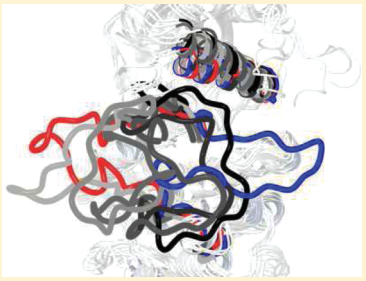
Matteo Chioccioli,[†] Simone Marsili,[‡] Claudia Bonaccini,[†] Piero Procacci,[‡] and Paola Gratteri*[,†]

[†]Laboratory of Molecular Modeling Cheminformatics and QSAR, Department of Pharmaceutical Sciences, University of Firenze, Via Ugo Schiff 6, I-50019 Sesto Fiorentino, Firenze Italy

[‡]Department of Chemistry "Ugo Schiff", University of Firenze, Via della Lastruccia 3-13, I-50019 Sesto Fiorentino, Firenze Italy

S *Supporting Information*

ABSTRACT: Human vascular endothelial growth factor receptor type 2 (h-VEFGR2) is a receptor tyrosine kinase involved in the angiogenesis process and regarded as an interesting target for the design of anticancer drugs. Its activation/inactivation mechanism is related to conformational changes in its cytoplasmatic kinase domain, involving first among all the αC-helix in N-lobe and the A-loop in C-lobe. Affinity of inhibitors for the active or inactive kinase form could dictate the open or closed conformation of the A-loop, thus making the different conformations of the kinase domain receptor (KDR) domain different drug targets in drug discovery. In this view, a detailed knowledge of the conformational landscape of KDR domain is of central relevance to rationalize the efficiency and selectivity of kinase inhibitors. Here, molecular dynamics simulations were used to gain insight into the conformational switching activity of the KDR domain and to identify intermediate conformations between the two limiting active and inactive conformations. Specific energy barriers have been selectively removed to induce, and hence highlight at the atomistic level, the regulation mechanism of the A-loop opening. The proposed strategy allowed to repeatedly observe the escape of the KDR domain from the DFG-out free energy basin and to identify rare intermediate conformations between the DFG-out and the DFG-in structures to be employed in a structure-based drug discovery process.



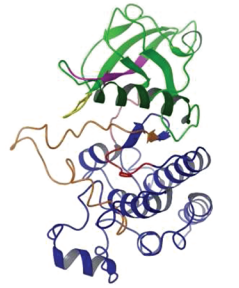
INTRODUCTION

Conformational changes in human vascular endothelial growth factor receptor type 2 (VEGFR2), and in general in tyrosine kinase receptors (RTK), characterize this class of proteins more than any other.[1] They play a key role in the activation/inactivation mechanism of protein kinases, involved in the cell signal transduction pathways and thus regarded as important target for cell proliferation, differentiation, migration, and metabolism and, therefore, in the treatment of cancer.[2] The majority of receptor kinases, included VEGFR2, consist of a single polypeptide chain spanning the cellular membrane. The extracellular portion represents the most variable region throughout the RTKs, while the more uniform cytoplasmatic domain (kinase domain receptor, KDR) is folded into two lobes, a smaller N-terminal lobe containing a functionally important glycine-rich nucleotide binding loop (P-loop or phosphate binding loop) and a larger C-terminal lobe. These lobes are joined by a segment, the hinge region, outlining a hydrophobic cleft where ATP binds. The C-lobe includes the catalytic domain and the so-called A-loop (activation loop), where the tyrosine residues undergoing phosphorylation are located. As a consequence of the high conformational mobility of these proteins, kinase domains can adopt at least two limiting conformations, the open and the closed state.[1,2] Structural and computational studies performed on kinase domains solved in the open and the closed states have made possible to highlight the main differences between these two conformations and to obtain an understanding of the complex conformational rearrangements occurring during the opening mechanism.[3−5] The transition from the closed to the open state involves a change in the orientation of both the A-loop in the C-terminal lobe and the αC-helix in the N-terminal lobe. The conformation taken by the A-loop affects, in turn, the conformation of a characteristic DFG motif, highly conserved among tyrosine kinases. In the so-called DFG-in active conformation the aspartate residue orients toward the ATP binding cleft, and the phenylalanine is buried in a hydrophobic pocket adjacent to the ATP site. In the DFG-out inactive conformation the flipping of the DFG motif relative to the active conformation causes the phenylalanine side chain to occupy the ATP binding cleft, thus uncovering the hydrophobic pocket. DFG-in/DFG-out states are also related to the position of the αC-helix, whose displacement can disrupt a salt bridge between a conserved glutamic acid and a catalytic lysine residue, critical for the kinase catalytic activity.[3,6] Previous computational studies[7,8] have suggested the possibility that the closed-to-open conformational transition in kinase domains

Received: October 26, 2011
Published: January 10, 2012

Table 1. Important VEGFR2/KDR Regions[a]

region		sequence	residue number
P-loop		GRGAFG	839-844
β3 strand		VAVKMLK	863-869
αC-helix		HSEHRALMSELKILIHI	874-890
hinge region		FCKFG	916-920
catalytic loop		HRDLAARN	1024-1031
A-loop		DFGLARDIYpKDPDYpVRKGDARLPLKWMAPE	1044-1073
	magnesium binding loop	DFGLA	1044-1048
	Activation segment	YpKDPDYp	1052-1057



N-lobe (light green) includes: P-loop (yellow), β3 strand (purple), αC-helix (dark green). Hinge region (pink) C-lobe (dark blue) includes: catalytic loop (red), A-loop (orange)

[a]Residue numbers as in 2OH4. Highly conserved residues in the kinase family are in red.

occurs according to a conserved two-step mechanism in which the activation loop opens first, followed by the rotation of the αC-helix. In this context, a possible intermediate conformational state has also been identified, which exhibits structural features in between those of the closed and open states.[7−9]

Due to the involvement of kinases as targets in the treatment of cancer, their structural plasticity has been increasingly attracting researchers' attention, particularly in view of identifying small molecules as inhibitors targeting their different conformational states. Kinase inhibition can be achieved by targeting the ATP binding site in the active state or by locking kinase in an inactive state, thus various conformations of the KDR domain, included possible intermediate states, have to be taken into account in designing RTK inhibitors.[10−12]

Molecular dynamics (MD) simulations are a powerful tool for studying conformational rearrangements and the time-dependent behavior and evolution of molecular systems. Large biological molecules, such as proteins, are characterized by rough free energy landscapes with multiple minima separated by barriers that are larger than thermal energies[13,14] and that make inadequate the sampling performed with traditional MD simulations. Moreover, in large kinase domains, the sampling of the transition from the closed to the open state of the flexible A-loop (or vice versa) via standard MD simulations is further complicated by the high level of cooperativity shown in the KDR regulation mechanism. As an example, in a recent standard MD simulation of the kinase domain of the Abl protein, the DFG-in/DFG-out transition was observed only once in 2 μs of standard MD simulation, and it was induced by introducing an ad hoc mutation "to further increase the extent of motion of the nearby helix αC".[9] The problems connected to the roughness and complication of the free energy landscape in proteins can be partially overcome performing nonstandard MD simulations using accelerated sampling schemes. In previous works, kinases have been simulated at high temperature[15] or, using an addictive biasing potential, built to flatten the free energy profile along some selected collective variables.[7]

The present study has been intended for sampling the wide conformational variety of the kinase domain KDR of VEGFR2 taken as a representative component of RTKs. In fact, to the best of our knowledge, a comprehensive (structural or computational) study on the VEGFR2/KDR domain, shedding light into its activation mechanism at the atomistic level, has not been previously reported.

A protocol has been set up whereby only specific energy barriers, i.e., those involved in the A-loop opening mechanism, have been smoothed to accelerate sampling. First a "core" of residues has been identified whose dynamics is correlated to the opening transition of the KDR. Then, the solvated protein has been simulated using a modified potential, in which the intracore interactions have been weakened with respect to the original force field, eventually leading to a smoothed free energy landscape and to an accelerated sampling.

This softening procedure has been applied to the torsional and nonbonded interactions only, leaving unaltered the fastest bonded interactions (stretching, bending, and improper torsional terms).

By means of these 'altered dynamics' simulations, the escape of the VEGFR2/KDR domain from the DFG-out free energy basin has been repeatedly observed, and rare, intermediate conformations between the DFG-out and the DFG-in structures have been identified. The employed technique does not require the a priori definition of a reaction coordinate for the opening transition of the kinase. Rather, specific energy barriers can be selectively targeted in a hierarchical sequence of simulations to induce, and hence highlight at the atomistic level, the regulation mechanism of the A-loop opening.

The paper is organized as follows: In the first part of the Results Section we will discuss the identification of a particular set of residues that are important for the stability of the closed and open configurations (the core residues) from the analysis of the results of standard MD simulations. In the second part of the Results Section, we will present the results from a set of simulations performed using the modified potential (involving the core residues identified in the previous section) tailored to

dx.doi.org/10.1021/ci200513a | J. Chem. Inf. Model. 2012, 52, 483−491

accelerate the sampling of the opening transition of the VEGFR2/KDR, and we will propose a possible intermediate metastable state along the opening path. The Discussion and Conclusions Section contains conclusive remarks and discussion.

RESULTS

Identification of the Core Residues. Tailoring our procedure to the VEGFR2 kinase domain requires information about which parts are most rigid and which parts are more flexible. Interconversion of KDR inactive/active states is mainly related to conformational changes of some of its structural elements (Table 1), first among all the αC-helix in the N-lobe and the A-loop in the C-lobe.

The investigation of the opening/closure mechanism in KDR domain was carried out starting from the two limiting conformations of A-loop, i.e., the closed/DFG-out and the open/DFG-in conformations of the KDR domain A-loop. In an early stage of the study, we focused on the full characterization of these two limiting conformations prepared as described in the Experimental Section.

The two conformers were simulated for 10 ns at constant temperature and pressure (T = 300 K, P = 1 atm) using the AMBER ff03 force field[16] and the TIP3P explicit water model.[17] The simulations were performed with the ORAC software.[14,18]

The systems maintained their starting (closed or open) conformation throughout the simulations as shown by root-mean-squared deviation (RMSD) analysis (Figure S1, Supporting Information), although residues belonging to the A-loop exhibit larger fluctuations with respect to the rest of the protein. A rearrangement was also observed for the αC-helix in the N-terminal lobe which, though sharing the same initial conformation in the two simulations, undergoes to a partial and opposite rotation adopting two different final orientations (Figure 1).

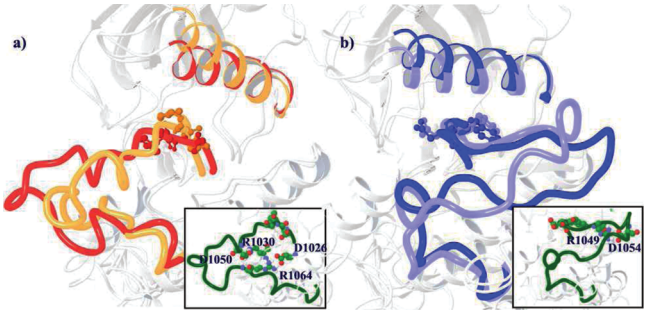


Figure 1. Ribbon representation of (a) the closed DFG-out and (b) the open DFG-in conformations of the KDR domain at the beginning (αC-helix and A-loop red and dark blue, respectively) and after 10 ns MD simulation (αC-helix and A-loop orange and light blue, respectively). Insets show the interactions between residues that act as constraints to the free movements of the A-loop.

In both simulations, the highly conserved DFG motif did not substantially change its starting orientation indicating that the time scale for DFG flip is much longer than a few nanoseconds. In the simulated trajectories, the closed A-loop conformation is stabilized by two salt bridges between the charged pairs D1050/R1030 and D1026/R1064 (Figure 1a, inset). The simulation of the open conformation shows a stable salt bridge between the side chains of R1049 and D1054 and a series of interactions between hydrophobic residues, also involving the DFG motif, as recently highlighted for other kinase domains.[3−5]

These results are not surprising since large conformational rearrangements, such as the opening/closing of the KDR domain, are too slow to occur spontaneously in fully atomistic MD simulations. Their long time scales originate from the ruggedness of the characteristic free energy landscape associated with this kind of transition, which is composed of multiple minima separated by relatively high free energy barriers. In particular, a complex event, such as the opening/closing of KDR domain, involves a sequence of smaller transitions over different energy barriers, including switching between torsional conformers and the rupture and formation of salt bridges and hydrophobic networks. Typical time scales for protein side chains and backbone movements ($\sim 10^{-9}$ and $\sim 10^{-5}$ s, respectively) entail a microsecond time scale for the process by which an inactive kinase is converted into an active kinase.[19]

In order to avoid the long waiting times inherent in conventional MD, barrier crossing was accelerated by weakening the interactions between a set of residues, here referred to as the core residues, which are central to the KDR opening mechanism, reducing the torsional and nonbonded energetic barriers slowing down conformational sampling.

From the previous MD simulations, it clearly emerged that the opening movement of the A-loop is hindered by two salt bridges involving two residues within the loop and two belonging to the rest of the protein, the R1030/D1050 and D1026/R1064 pairs. The two residues outside the A-loop, R1030 and D1026, and the A-loop residues (I1042-E1073) were thus chosen as the core of our system.

Figure S2 and Figure S3, Supporting Information report the root-mean-square fluctuations (RMSFs) averaged over the atoms of the residues in the range 1020−1090 obtained from the standard MD of the closed and open form of the KDR domain. The inspection of these RMSFs points out again that the reduced mobility of the A-loop in the closed state is due to the residues D1050 and R1064 involved in the salt bridges with R1030 and D1026, respectively. Thus, to weaken these salt-bridges, the two residues R1030 and D1026 were included in the core residues.

In the DFG-in open state, the region including the DFG motif is by far the less mobile part of the A-loop. The rigidity in this case is due to hydrophobic interactions involving the network of nonpolar residues surrounding F1045. To weaken such a network, the balance between solvent−core residues solvent−solvent interactions should be altered with a major scaling of the interaction potential.

Accelerated MD Simulations. The strategy followed in the present paper consists in selectively accelerating the dynamics of the VEGFR2/KDR A-loop, by weakening the interactions between the core residues identified in the previous section. A series of simulations were set up by softening the interactions involving atoms within the core residues using a modified potential V' which is related to the original potential energy function V by the relation:

$$V' = V + (\alpha - 1)V_c$$

Here, V_c denotes the potential collecting all the interactions between atoms belonging to the core residues and α is a factor smaller than one that quantifies the weakening of interactions between core residues. For $\alpha = 0$ the intracore interactions are

dx.doi.org/10.1021/ci200513a | J. Chem. Inf. Model. 2012, 52, 483−491

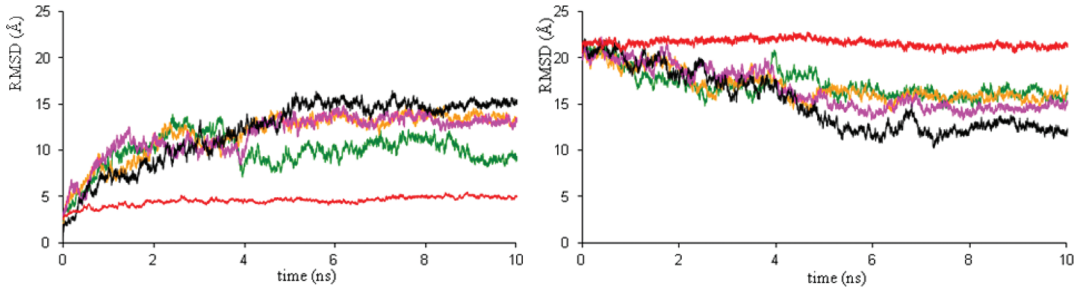


Figure 2. Ten ns accelerated MD simulations. RMSD values evaluated for the C_α atoms of the A-loop with respect to the corresponding atoms in the closed/DFG-out (left) and in the open/DFG-in (right) conformation as a function of the simulation time. Green, orange, purple, and black lines represent independent accelerated MD simulations in which the A-loop reaches an intermediate conformation; the red line represents an accelerated simulation in which the A-loop maintains the closed/DFG-out conformation.

totally removed. Accelerated dynamics of the kinase were simulated using a value of α= 0.1. Using an altered potential will lead to an altered sampling with respect to the original free energy landscape. It is instructive to consider the case in which the core includes all the residues in the system (protein and water molecules). In this case, from the point of view of conformational sampling, this approach is equivalent to simulate a temperature-induced conformational transition.[20−22] High-temperature simulations of kinase A-loop isomerization were reported in the literature[15] using harmonic restraints between atoms outside of the A-loop. However, the high-temperature conformational landscape of a kinase can be expected to differ substantially from the one in standard conditions. For example, the ability of water to stabilize hydrophobic interactions, that is crucial in protein folding and in many biological molecular recognition processes,[23,24] is partially lost at high temperature. With respect to a high-temperature dynamics, the "selective softening" of interactions we adopt here fully preserves the effect of hydrophobicity, while flattening energy barriers in the intracore potential energy landscape.

Sixteen independent, 10 ns long accelerated MD simulations were performed starting from the closed, unphosphorylated, crystallographic conformation of the KDR domain. One more simulation was performed starting from the closed conformation in the phosphorylated state, in order to test the effect of the unusual phosphorylation[25] of the crystal structure (2OH4)[26] on the efficiency of sampling. Results from this simulation confirmed the importance of the dephosphorylation of the inactive conformation in order to allow the free movement of the A-loop and sampling of intermediate conformations (Figure S4 and Figure S5, Supporting Information).

Among the 16 accelerated simulations, 12 maintained the A-loop in a closed DFG-out conformation for the whole simulated time, while in the other 4, the system escaped from the closed conformation free energy basin, reaching an intermediate structure between the active and inactive conformations. The RMSD values for the four escaping trajectories (green, orange, purple, and black lines) and for the one taken as representative of the other 12 (red line) are shown in Figure 2. In all the escaping trajectories, the A-loop conformation gets closer to the open structure (Figure 2, right) while moving away from the initial closed structure (Figure 2, left).

In order to analyze the four escaping accelerated simulations, we performed a cluster analysis on the last 5 ns of the corresponding trajectories (Table S1, Supporting Information and Experimental Section). Comparing the most populated cluster of each simulation, we observe common structural features for the metastable intermediate states. Compared with the starting closed DFG-out conformation, these partially open conformations show, although to a different extent, a simultaneous rearrangement involving the A-loop and the α-C helix in the N-terminal lobe (Figure 3 and Figure S6, Supporting Information). In detail, while the A-loop partially opens in its C-terminal portion, the α-C helix undergoes a slight rotation away from the ATP binding site. This rearrangement comes with a shift in the position of the P-loop and with a partial rotation of the conserved DFG motif, which in three simulations (green, orange, and purple lines in Figure 2) is characterized only by a slight rotation of F1045. In the simulation showing the greatest opening of the A-loop (black line in Figure 2), both D1044 and F1045 undergo a simultaneous and partial rotation (Figure 4).

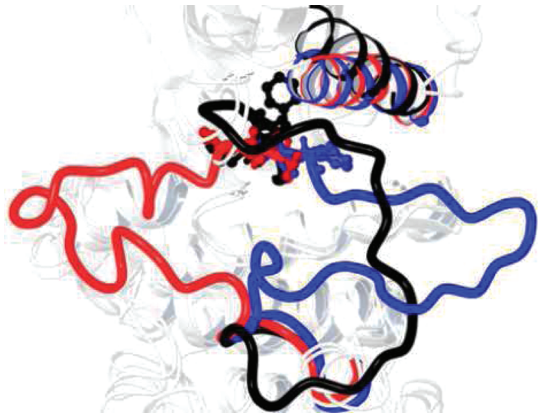


Figure 3. Movement of the A-loop observed for the 10 ns long accelerated MD simulation represented by the black line in Figure 2. The closed (red) and open (blue) states are reported as reference.

In this trajectory the intermediate state reached by the DFG motif is highly stabilized by a hydrophobic cleft formed by L887, V897, V912, and V914 which wraps and locks the side chain of F1045. Moreover, the polar interaction involving the conserved E883, belonging to the α-C helix, and the catalytic K866 contribute to further stabilize this metastable intermediate (Figure 5).

It is worth noting that in all the escaping trajectories, once the KDR domain reached the metastable intermediate state, it retained this conformation throughout the remaining simulated time, without reaching a fully opened configuration. The

dx.doi.org/10.1021/ci200513a | J. Chem. Inf. Model. 2012, 52, 483−491

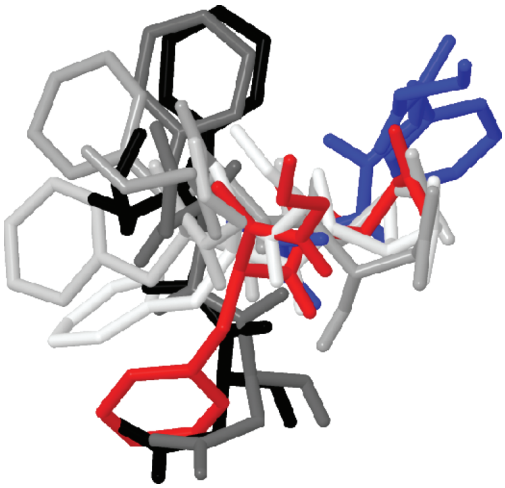


Figure 4. Snapshots taken from the trajectory represented by the black line in Figure 2. The snapshots are colored in chronological order on a gray scale. The X-ray solved DFG-out (red) and the "chimeric" homology-built DFG-in (blue) conformations are also shown as reference.

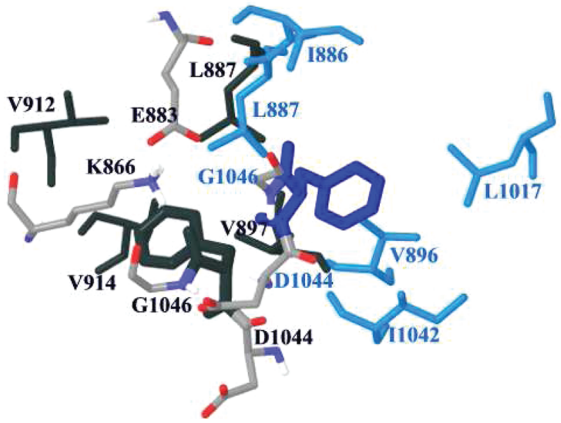


Figure 5. Comparison of the hydrophobic clefts surrounding the F1045 side chain in the intermediate metastable conformation (black) and in the "chimeric" homology-built DFG-in structure (blue).

sampling of the open DFG-in conformation is mainly prevented by the network of interactions which interferes with the complete DFG flip; the further rotation of the α-C helix and the subsequent breaking of the hydrophobic core seem to be the requirements to observe the DFG flip and the complete opening of the KDR domain. Accordingly, the A-loop adopts a shape very similar to the one observed in the open DFG-in conformation in its C-terminal region, while a reduced movement can be observed for the portion which includes the highly conserved DFG motif locked in the intermediate state (Figure 3).

The stability of the 4 similar representative intermediate conformations, previously selected with the cluster analysis, was checked by performing 10 ns standard MD simulations at 300 K and 1 atm. These simulations were carried out with the Amber9 suite.[27] During the simulated time, the metastable states did not exhibit any important rearrangement in the A-loop, even if it is worth noting that in the three trajectories, starting from the structures with a reduced opening of the loop, the DFG motif rotates back to a DFG-out conformation. The analysis of the RMSD values revealed a greater stability for the A-loop conformation in which both D1044 and F1045 underwent a partial rotation; this simulation was thus extended up to 18 ns (Figure S7, Supporting Information). Then, on the last 5 ns of this simulation, we performed a cluster analysis (Table S1, Supporting Information) and selected a representative structure for the metastable intermediate state of the KDR domain. Apart from D1044 that returned to a position similar to that assumed in the DFG-out structure, the overall architecture of the KDR retained the previously reached intermediate conformation with the side chain of F1045 buried in the above-described hydrophobic network. The A-loop is in an intermediate conformation and adopts a shape very similar to that observed in the open conformation in its C-terminal region, while the α-C helix is slightly rotated away from the ATP binding site, with E883 involved in a salt bridge with K866. Finally, the superposition of the C-terminal lobe of this intermediate conformation with the corresponding region of the open and closed states led to the identification of a relative movement of the whole N-lobe with respect to the C-lobe (Figure S8, Supporting Information). In order to quantify this movement, the mean distance (D_{int} = 27.59 ± 0.43 Å) between the centroids of the N-lobe (residues from C815 to K918) and C-lobe (residues from F919 to D1169) in the different members of the most populated cluster (black trajectory in Figure S7, Supporting Information) was evaluated and compared to the corresponding distance (D_{out} = 26.74 Å) obtained for the experimental structure (pdb code 2OH4, closed conformation). This result, as well as the inspection of the Figure S8, Supporting Information, indicates that the rearrangement of the A-loop in the intermediate state is associated to a collective fluctuation of the whole domain, in which the two lobes move apart from each other.

The structure of the intermediate state was observed, in the accelerated dynamics starting from the closed state, as a consequence of the selective weakening of two salt-bridges with one partner involving an extra A-loop residue. The kinetics in the closed to intermediate transition is thus apparently regulated by the rare event of a concomitant breaking of the two salt bridges R1030/D1050 and D1026/R1064.

On the other hand, the intermediate state differs from the open DFG-in conformation in two main features: a slight movement of the N-lobe including the αC-helix with respect to the C-lobe and the entanglement of the F1045 residue within the hydrophobic network of residues flanking the hydrophobic pocket surrounding the same residue in the open state (Figure 5). Accelerated independent simulations, each lasting 10 ns, starting from the open DFG-in structure were also performed. Notwithstanding the potential weakening of the A-loop residues, the system stubbornly remains in the open state with the F1045 side chain tightly inserted in a network of hydrophobic residues, severely reducing the mobility of the A-loop at the DGF end (Figure S9 and S10, Supporting Information). The transition from the open DFG-in conformation to the intermediate state is seemingly regulated by a different kinetic mechanism involving cooperative motions rather than rare events. Such a mechanism could not be effectively accelerated using our scaling protocol that appears in such a case inadequate for promoting the disruption of such a hydrophobic network and hence, possibly, the evolution of the open structure to the intermediate state. A more aggressive scaling involving the solvent−solvent or solute−solvent interactions would be of course much more effective in releasing the F1045 side chain, but such a global modification

dx.doi.org/10.1021/ci200513a | J. Chem. Inf. Model. 2012, 52, 483−491

would alter dramatically the whole system, inducing uncontrolled and artificial conformational rearrangements.

DISCUSSION AND CONCLUSIONS

Proteins can fluctuate to produce conformations suited for ligand binding. Thus it is nowadays important in the drug design process to use methods that take into account the conformational changes occurring when protein and ligands form complexes as well as the variety of equilibrium states which can characterize a target.

The joint standard and altered MD simulations applied for sampling the low-energy states neighboring conformations of the kinase domain of VEGFR2 pointed out that the opening of the A-loop in the KDR domain is indeed a complex event that requires, besides the impressive conformational swap of the loop itself, a substantial reorganization of the secondary structure elements surrounding the loop. Such a complex opening/closing process is driven by a remarkable balance of concurrent electrostatic and hydrophobic interactions. In fact, the complete closed-to-open (or vice versa) A-loop transition seems mainly prevented by a hydrophobic barrier due to the positioning of the F1045 side chain of the DFG motif into a hydrophobic pocket and by a network of polar interactions involving the pairs K866/E883, R1030/D1050, and D1026/R1064. The switching in the electrostatic network was already pointed out for other kinases,[7,28,29] thus suggesting a common activation mechanism for the A-loop which takes place through the formation of metastable intermediates.[7−9,30−32] Unlike other nonstandard methodologies,[7] the accelarated MD used in the present study does not require the a priori definition of a reaction coordinate which always represents a critical aspect in the setting up of the computational procedure. The accelerated sampling is rather achieved by softening specific and localized free energy barriers to stimulate a temperature-induced conformational transition, while most of the degrees of freedom of the system remain in standard conditions. Such procedure allows to clearly identify the intermediate states between open and closed conformations in a few nanoseconds of altered dynamics. The computational procedure showed a strict correlation between the movement of the two N- and C-terminal lobes and the opening/closing mechanism of the A-loop, suggesting the idea that this cooperative movement is a mechanism for the control of the transition from the closed to the open conformation and vice versa. The major result emerging from this study is the identification of a stable, intermediate conformational state for the h-VEGFR2/KDR to use, together with the active and inactive conformations of the target, as a valuable platform for structure-based design of inhibitors for VEGFR2. Moreover, the presented procedure could take shape as a possible investigation strategy within protein kinases, making possible to overcome the time scale limitation of standard all-atoms MD simulations.

EXPERIMENTAL SECTION

Calculations were performed using a dual-core AMD Opteron four-processor LINUX workstation and the CINECA HPC computer IBM-SP6.[33]

Closed A-Loop Conformation. Among the X-ray solved structures of the VEGFR2/KDR domain in its A-loop closed conformation, only 2OH4,[26] and more recently also 2XIR, exhibit the entire activation segment solved, although the B-factors associated to its atom coordinates are very high. Thus, the crystal structure of the KDR domain of the human VEGFR2 in complex with a benzimidazole−urea inhibitor, refined at 2.05 Å resolution (pdb code 2OH4), was selected as working structure throughout this work.

2OH4 was submitted to the Schrodinger protein preparation procedure[34] in order to: (i) remove all nonprotein units, including the inhibitor and all the water molecules; (ii) select only one orientation for those residues (C860, C1043, L1067, E1156, numbering according to 2OH4 numbering system) with two alternative positions; (iii) rebuild residues having missing atoms (ranging from Y936 to Y994); and (iv) assign the most probable protonation state to the histidine residues according to the environment at pH of 7 (histidine tautomers with the hydrogen on the ε or δ nitrogen or double protonated). Moreover, the coordinates of the first residue (H814), not completely solved, were removed. Step (iii) was carried out starting from the knowledge that the removal of the residues 936−994 is necessary for crystallographic purposes but has no effect on the intrinsic kinase activity and that the region has a structural role in that it connects two α-helices in the C-terminal lobe and maintains the overall architecture of the KDR domain.[35] Thus the primary sequence of the KDR domain was modified through the introduction of a short loop containing five alanine residues connecting the two solved ends, by using the routine Build Loop of Swiss-PdbViewer[36] and verifying the results obtained according to a set of geometrical and energetic parameters provided by the routine Build Loop, such as a count of the clashes (number of bad contacts or H-bonds), an energy information (computed with a partial implementation of the GROMOS96 force field,[37] and a mean force potential value (computed from a "Sippl-like" mean force potential).[38] Finally, the two phosphotyrosines of the A-loop in the crystal structure 2OH4 (Yp1052 and Yp1057) were mutated into tyrosine residues (Y1052 and Y1057).

Open A-Loop Conformation. A homology model approach was used to obtain the open A-loop/DFG-in conformation of the KDR domain, whose experimental structure is not available. The primary sequences of the KDR domain and of the kinase domain of the insulin receptor (pdb code 1IR3)[39] were aligned using ClustalW[40] (BLOSUM62 similarity matrix) multiple global alignment, implemented in BioEdit.[41] This sequence alignment revealed a high degree of similarity for the A-loop in the two proteins: (i) the same number of residues between the two common end motifs, DFG and APE; (ii) the same position of the two phosphotyrosines (X in the primary sequences in Table S2, Supporting Information) of the KDR domain compared with two of the three phosphotyrosines of the insulin receptor; and (iii) the sequence identity in the residues at the beginning (VKI-DFG) and at the end (WMAPE) of the A-loop (Table S2, Supporting Information).

A "chimeric" structure of the KDR domain in its active conformation was obtained replacing the 2OH4 A-loop backbone coordinates by the corresponding coordinates of the crystal structure 1IR3. The remaining residues for the "chimeric" structure were taken from the KDR domain in the closed and DFG-out conformation (pdb code 2OH4), with the alanine loop inserted between residues Y936 and Y994. In detail, the homology model was generated first by inserting the A-loop of the crystal structure 1IR3, between the motif VKI (underlined in the Table S2, Supporting Information) and the motif APE, into the previously prepared structure of the KDR domain in its closed conformation, and then by modifying the

side chains of the different amino acids according to the KDR primary sequence using the program Swiss-PdbViewer[36] (MUTATE command). Similarly to the closed conformation of the KDR domain, we modified the two phosphotyrosines of the A-loop (Yp1052 and Yp1057) into tyrosine residues (Y1052 and Y1057). The reliability and the possible reorganization of the different structural elements in the built "chimeric" structure were investigated by using standard MD simulations.

MD Simulations. MD simulations were performed by using the program ORAC[14,18] and the Amber9 suite,[27] the AMBER ff03[16] force field, and the TIP3P water model.[17] The force field parameters for the two nonstandard phosphotyrosines (Yp1052 and Yp1057) were adapted from Homeyer et al.[42]

ORAC Standard MD Simulation. The phosphorylated and unphosphorylated open and closed structures were solvated by approximately 9100 water molecules in a rhombic dodecahedron box with periodic boundary conditions. The systems were initially equilibrated during a 100 ps NPT simulation at $T = 300$ K and $P = 1$ atm, where constant pressure was obtained using a modification of the Parrinello–Rahman Lagrangian,[43] and temperature control was achieved using a Nosé thermostat.[44] The cell side was equilibrated at an average value of 77 Å. The Ewald method with the smooth particle mesh algorithm[45] was used to compute electrostatic interactions. The grid spacing in each dimension of the direct lattice was ∼1.2 Å, whereas the Ewald convergence parameter was set to 0.43 $Å^{-1}$. Electroneutrality was enforced using, respectively, a uniform positive and negative charge density background equalizing the −1 e and +3 e excess negative and positive charges on the phosphorylated and unphosphorylated structures. A multiple time-step r-RESPA algorithm with a potential subdivision specifically tuned for proteins was used for integrating the equations of motion, with time-steps 9.0, 3.0, 1.5 fs for nonbonded and 0.75, 0.375 fs for bonded interactions.[18,46,47] The last configurations, obtained from these simulations performed on the unphosphorylated open (DFG-in) and closed (DFG-out) conformations of the target, were used as starting points for the production runs performed in the same conditions for 10 ns.

ORAC Accelerated MD Simulations. Ten ns accelerated MD simulations were carried out on the unphosphorylated (16 simulations) and phosphorylated states of the previously equilibrated closed (DFG-out) conformation of the KDR domain. The potential weakening $V' = V + (\alpha - 1)V_c$, where V is the original potential energy function, was applied to the core residues (D1026, R1030, I1042-E1073), scaling by $\alpha = 0.1$ only the core residues potential terms (torsional and nonbonded interactions); V_c denotes the potential collecting all the interactions between atoms belonging to the core residues. Other interactions, including those between the core and the rest of the system, were not affected by scaling factors. All other experimental conditions were maintained as those of standard MD performed by means of ORAC.

Amber9 Standard MD Simulation. Representative unphosphorylated structures of the VEGFR2/KDR intermediate conformation, selected with a cluster analysis performed on the previous accelerated MD simulations, were immersed in a truncated octahedral box, whose edges were located 10 Å from the closest atom of the protein, containing about 10 000 water molecules. Then, 3 Cl^- counterions were added to the solvent bulk of the protein–water complexes to maintain neutrality on the system. Before starting the standard MD simulations, an energy minimization of the entire ensemble was performed by setting a convergence criterion on the gradient of 0.01 kcal mol^{-1} $Å^{-1}$. Then, water shells and counterions were equilibrated for 40 ps at 300 K, and subsequently, 10 ns of standard MD simulations, extended up to 18 ns for the structure showing the greatest opening of the A-loop, in an isothermal–isobaric ensemble were performed on each structure. In the production runs, the protein structures were simulated in periodic boundary conditions. The van der Waals and short-range electrostatic interactions were estimated within a 10 Å cutoff, whereas the long-range electrostatic interactions were assessed by using the particle mesh Ewald method,[45] with 1 Å charge grid spacing interpolated by fourth-order B-spline, and by setting the direct sum tolerance to 10^{-5}. Bonds involving hydrogen atoms were constrained by using the SHAKE algorithm[48] with a relative geometric tolerance for coordinate resetting of 0.00001 Å. Berendsen's coupling algorithms[49] were used to maintain constant temperature and pressure with the same scaling factor for both solvent and solutes and with the time constant for heat–bath coupling maintained at 1.5 ps. The pressure for the isothermal–isobaric ensemble was regulated by using a pressure relaxation time of 1 ps in the Berendsen's algorithm. The simulations of the solvated proteins were performed using a constant pressure of 1 atm and a constant temperature of 300 K. A time step of 2 fs was used in these simulations.

Cluster Analysis. Cluster analysis was performed with the quality threshold clustering algorithm[50] on the intramolecular distances between the C_α atoms belonging to the residues of the P-loop (G839-G844), the αC-helix in the N-lobe (H874–I890), and the A-loop (I1042-E1073). The distance between two structures a and b is defined as:

$$D = \max_i\{|d_i^b - d_i^a|\}$$

where the symbol d_i^a denotes the i-th distance C_α– C_α in the structure a, and the max function selects the distance that changes more in absolute value between the a and b structures. The distance threshold was set to 5 Å.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Further data. Supporting Figures 1–10, Supporting Tables 1–2 and details concerning standard and accelerated MD simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: paola.gratteri@unifi.it. Telephone: +39 055 4573702.

■ ACKNOWLEDGMENTS

The authors thank the Italian Ministero dell'Università e della ricerca for financial support. Ente Cassa di Risparmio di Firenze, Italy, is also gratefully acknowledged for a grant to M.C. We acknowledge the CINECA award no. HP10CSN0WG, 2010 for the availability of high-performance computing resources and support.

■ REFERENCES

(1) Huse, M.; Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **2002**, *109*, 275–282.

(2) Hubbard, S. R.; Till, J. H. Protein tyrosine kinase structure and function. *Annu. Rev. Biochem.* **2000**, *69*, 373−398.
(3) 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.
(4) Kornev, A. P.; Taylor, S. S. Defining the conserved internal architecture of a protein kinase. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 440−444.
(5) Taylor, S. S.; Kornev, A. P. Protein kinases: Evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* **2011**, *36*, 65−77.
(6) Krishnamurty, R.; Maly, D. J. Biochemical Mechanisms of Resistance to Small-Molecule Protein Kinase Inhibitors. *ACS Chem. Biol.* **2010**, *5*, 121−138.
(7) 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.* **2009**, *131*, 244−250.
(8) Gan, W.; Yang, S.; Roux, B. Atomistic view of the conformational activation of Src kinase using the string method with swarms-of-trajectories. *Biophys. J.* **2009**, *97*, L08−L10.
(9) Shan, Y. B.; Seeliger, M. A.; Eastwood, M. P.; Frank, F.; Xu, H. F.; Jensen, M. O.; 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.
(10) Garuti, L.; Roberti, M.; Bottegoni, G. Non-ATP competitive protein kinase inhibitors. *Curr. Med. Chem.* **2010**, *17*, 2804−2821.
(11) Simard, J. R.; Getlik, M.; Grutter, C.; Schneider, R.; Wulfert, S.; Rauh, D. Fluorophore labeling of the glycine-rich loop as a method of identifying inhibitors that bind to active and inactive kinase conformations. *J. Am. Chem. Soc.* **2010**, *132*, 4152−4160.
(12) Zuccotto, F.; Ardini, E.; Casale, E.; Angiolini, M. Through the "gatekeeper door": Exploiting the active kinase conformation. *J. Med. Chem.* **2010**, *53*, 2681−2694.
(13) Kamberaj, H.; van der Vaart, A. Multiple scaling replica exchange for the conformational sampling of biomolecules in explicit water. *J. Chem. Phys.* **2007**, *127*, 234102.
(14) Marsili, S.; Signorini, G. F.; Chelli, R.; Marchi, M.; Procacci, P. ORAC: A Molecular Dynamics Simulation Program to Explore Free Energy Surfaces in Biomolecular Systems at the Atomistic Level. *J. Comput. Chem.* **2010**, *31*, 1106−1116.
(15) 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.
(16) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(17) Jorgensen, W. L.; Chandrashakar, J.; Madura, J.; Impey, R.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(18) Procacci, P.; Darden, T. A.; Paci, E.; Marchi, M. ORAC: A molecular dynamics program to simulate complex molecular systems with realistic electrostatic interactions. *J. Comput. Chem.* **1997**, *18*, 1848−1862.
(19) Volkman, B. F.; Lipson, D.; Wemmer, D. E.; Kern, D. Two-State Allosteric Behavior in a Single-Domain Signaling Protein. *Science* **2001**, *291*, 2429−2433.
(20) Paci, E.; Karplus, M. Unfolding proteins by external forces and temperature: The importance of topology and energetics. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6521−6526.
(21) Salimi, N. L.; Ho, B.; Agard, D. A. Unfolding simulations reveal the mechanism of extreme unfolding cooperativity in the kinetically stable alpha-lytic protease. *PLoS Comput. Biol.* **2010**, *6*, 2.
(22) Settanni, G.; Fersht, A. R. High temperature unfolding simulations of the TRPZ1 peptide. *Biophys. J.* **2008**, *94*, 4444−4453.
(23) Baron, R.; Set, P.; McCammon, J. A. Water in Cavity−Ligand Recognition. *J. Am. Chem. Soc.* **2010**, *132*, 12091−12097.
(24) Hummer, G. Molecular binding: Under water's influence. *Nat. Chem.* **2010**, *2*, 906−907.
(25) Roskoski, R. Jr VEGF receptor protein-tyrosine kinases: Structure and regulation. *Biochem. Biophys. Res. Commun.* **2008**, *375*, 287−291.
(26) Hasegawa, M.; Nishigaki, N.; Washio, Y.; Kano, K.; Harris, P. A.; Sato, H.; Mori, I.; West, R. I.; Shibahara, M.; Toyoda, H.; Wang, L.; Nolte, R. T.; Veal, J. M.; Cheung, M. Discovery of novel Benzimidazoles as potent inhibitors of TIE-2 and VEGFR-2 tyrosine kinase receptors. *J. Med. Chem.* **2007**, *50*, 4453−4470.
(27) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *Amber9 Suite*University of California, San Francisco: San Francisco, CA, 2006
(28) Banavali, N. K.; Roux, B. Anatomy of a structural pathway for activation of the catalytic domain of Src kinase Hck. *Proteins: Struct., Funct., Genet.* **2007**, *67*, 1096−1112.
(29) Ozkirimli, E.; Yadav, S. S.; Miller, W. T.; Post, C. B. An electrostatic network and long-range regulation of Src kinases. *Protein Sci.* **2008**, *17*, 1871−1880.
(30) Banavali, N. K.; Roux, B. Flexibility and charge asymmetry in the activation loop of Src tyrosine kinases. *Proteins* **2009**, *74*, 378−389.
(31) Yang, S.; Banavali, N. K.; Roux, B. Mapping the conformational transition in Src activation by cumulating the information from multiple molecular dynamics trajectories. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 3776−3781.
(32) Yang, S.; Roux, B. Src kinase conformational activation: Thermodynamics, pathways, and mechanisms. *PLoS Comput. Biol.* **2008**, *4*, 3.
(33) *CINECA, Scientific High Performance Computing*; http://www.cineca.it/en/hardware/ibm-sp65376-0.
(34) *Schrödinger, Protein Preaparation Wizard*; Schrödinger, LLC: New York. 2007.
(35) McTigue, M. A.; Wickersham, J. A.; Pinko, C.; Showalter, R. E.; Parast, C. V.; Tempczyk-Russell, A.; Gehring, M. R.; Mroczkowski, B.; Kan, C. C.; Villafranca, J. E.; Appelt, K. Crystal structure of the kinase domain of human vascular endothelial growth factor receptor 2: a key enzyme in angiogenesis. *Structure* **1999**, *7*, 319−330.
(36) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714−2723.
(37) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*: Vdf Hochschulverlag AG an der ETH Zürich: Zürich, Switzerland. 1996; pp 1−1042.
(38) Sippl, J. M. Calculation of Conformational Ensembles from Potentials of Mean Force: an approach to the knowledge based prediction of local structures in globular proteins. *J. Mol. Biol.* **1990**, *213*, 859−883.
(39) Hubbard, S. R. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J.* **1997**, *16*, 5572−5581.
(40) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **1994**, *22*, 4673−4680.
(41) Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser. (1979−2000)* **1999**, *41*, 95−98.
(42) Homeyer, N.; Horn, A. H. C.; Lanig, H.; Sticht, H. AMBER force-field parameters for phosphorylated amino acids in different

protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J. Mol. Model.* **2006**, *12*, 281−289.
(43) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, *45*, 1196−1199.
(44) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511−519.
(45) 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.
(46) Marchi, M.; Procacci, P. Coordinates scaling and multiple time step algorithms for simulation of solvated proteins in the NPT ensemble. *J. Chem. Phys.* **1998**, *109*, 5194−5202.
(47) Procacci, P.; Darden, T.; Marchi, M. A very fast molecular dynamics method to simulate biomolecular systems with realistic electrostatic interactions. *J. Phys. Chem.* **1996**, *100*, 10464−10468.
(48) 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.
(49) Berendsen, H. J. C.; Postma, J. P. M; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(50) Heyer, L. J.; Kruglyak, S.; Yooseph, S. Exploring expression data identification and analysis of coexpressed genes. *Genome Res.* **1999**, *9*, 1106−1115.