BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2011,** *50,* 1755–1762 1755
DOI: 10.1021/bi101662j

Binding of a C-End Rule Peptide to the Neuropilin-1 Receptor: A Molecular Modeling Approach[†]

Nurit Haspel,[*,‡] David Zanuy,[*,§] Ruth Nussinov,[∥,⊥] Tambet Teesalu,[@] Erkki Ruoslahti,[@,▽] and Carlos Aleman[§,#]

[‡]*Department of Computer Science, The University of Massachusetts, Boston, Massachusetts 02125, United States,* [§]*Departament d'Enginyeria Química, ETSEIB, Universitat Politècnica de Catalunya, Avda. Diagonal 647, Barcelona E-08028, Spain,* [∥]*Center for Cancer Research Nanobiology Program, SAIC-Frederick, Inc., National Cancer Institute, Frederick, Maryland 21702, United States,* [⊥]*Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel,* [@]*Center for Nanomedicine, Sanford-Burnham Medical Research Institute at UCSB, 3119 Biology II Building, University of California, Santa Barbara, California 93106-9610, United States,* [▽]*Cancer Research Center, Sanford-Burnham Medical Research Institute, La Jolla, California 92037, United States, and* [#]*Center for Research in Nano-Engineering, Universitat Politècnica de Catalunya, Campus Sud, Edifici C', C/Pasqual i Vila s/n, 08028 barcelona, Spain*

Received October 15, 2010; Revised Manuscript Received January 18, 2011

ABSTRACT: Neuropilin-1 (NRP-1) is a receptor that plays an essential role in angiogenesis, vascular permeability, and nervous system development. Previous studies have shown that peptides with an N-terminal Arg, especially peptides with the four-residue consensus sequence R/K/XXR/K, bind to NRP-1 cell surfaces. Peptides containing such consensus sequences promote binding and internalization into cells, while blocking the C-terminal Arg (or Lys) prevents the internalization. In this study, we use molecular dynamics simulations to model the structural properties of the NRP-1 complex with a prototypic CendR peptide, RPAR. Our simulations show that RPAR binds NRP-1 through specific interactions of the RPAR C-terminus: three hydrogen bonds and a salt bridge anchor the ligand in the receptor pocket. The modeling results were used as the starting point for a systematic computational study of new RPAR analogues based on chemical modifications of their natural amino acids. Comparison of the structural properties of the new peptide−receptor complexes with the original organization suggests that some of the analogues can increase the binding affinity while reducing the natural sensitivity of RXXR to endogenous proteases.

Blood vessels are generally impermeable to macromolecules and supramolecular complexes. Increased permeability of blood vessels is associated with physiological angiogenesis during embryonic development (*1*). Vascular leakage is also a characteristic of tumor vessels (*2*). VEGF,[1] originally described as a tumor-secreted vascular permeability factor (*3*), is at least partially responsible for the leakiness of angiogenic vessels (*2*, *4*). VEGF was later independently cloned as an endothelial cell mitogen (*5*, *6*). Neuropilin-1 (NRP-1) is a hub receptor with multiple ligands that plays a central role in angiogenesis and cardiovascular development (*7*, *8*) and acts as a VEGF coreceptor during induction of vascular permeability. We have identified a class of peptides derived from phage display screens that induce cell internalization and tissue penetration via NRP-1 mediation (*9*). The peptides that promoted such a process upon binding to NRP-1 were found to have the R/K/XXR/K motif at their C-termini (known as the C-end rule), a pattern also shared by VEGF-A165 and by certain semaphorins (*9*). This four-residue motif contains an Arg or a Lys residue in the first and fourth position and any residue in the second and third positions. NRP-1 binding is particularly efficient when the RXXR C-end rule motif is present in tandem organization, such as RPARPAR, RGERPPR, and RVTPPR (*9*). Finally, cryptic R/K/XXR/K motifs embedded in a protein can be proteolytically activated by exogenous proteases to trigger vascular leakage and cell internalization (*9*). CendR peptides are particularly relevant to cancer drug delivery, as NRP-1 is frequently overexpressed in tumor cells. Hence, therapeutic strategies that include peptides with an R/K/XXR/K terminus may help in overcoming problems related to penetration of drugs into the tumor, because binding of such peptides to NRP-1 promotes extravasation and penetration of co-administered drugs to the tumor, as shown recently by Sugahara et al. (*10*).

Despite the central role of VEGF, the structure of NRP-1 in a complex with an R/KXXR/K peptide has not yet been elucidated from a crystal structure. However, the structure of the binding domains with a peptide mimic of the VEGF terminus was recently reported (*11*). Such a peptide, called tuftsin (an immunostimulatory tetrapeptide, TKPR), has a sequence very similar to the C-terminus of VEGF-A165, DKPRR, and competed with

[†]This research was supported in part by the National Science Foundation through TeraGrid resources provided by the Texas Advanced Computing Center (TACC) under Grant TG-MCB100025 (N. H.). Financial support from Generalitat de Catalunya is gratefully acknowledged (Research Group 2009 SGR 925; XRQTC; ICREA Academia prize for excellence in research to C.A.). This project has been funded in part by the National Cancer Institute, National Institutes of Health (NIH), via Contract HHSN261200800001E. This research was also supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

*To whom correspondence should be addressed. N.H.: e-mail, nurit.haspel@umb.edu; phone, (617) 287-6414. D.Z.: e-mail, david.zanuy@upc.edu; fax, +34.93.401.7150; phone, +34.93.405.4447.

[1]Abbreviations: NRP-1, neuropilin-1; RPAR, Arg-Pro-Ala-Arg peptide; VEGF, vascular permeability factor; CendR, C-end rule; NVT, constant volume, temperature, and number of moles; NPT, constant pressure, temperature, and number of moles; MD, molecular dynamics; rmsd, root-mean-square deviation; PDB, Protein Data Bank.

© 2011 American Chemical Society
Published on Web 01/20/2011

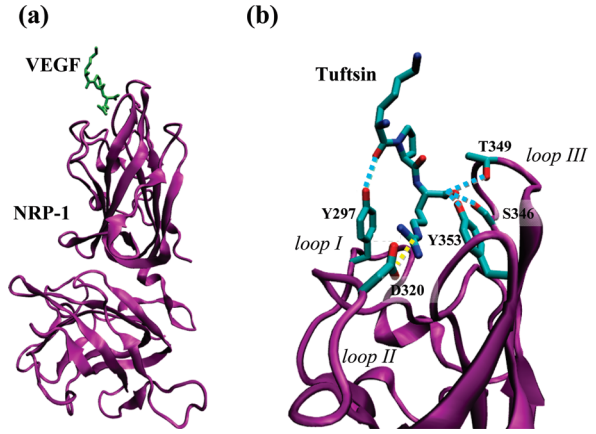

FIGURE 1: (a) Schematic illustration of the NRP-1−tuftsin complex (PDB entry 2ORZ) used to model the NRP-1−RPAR complex. NRP-1 is colored purple and tuftsin green. (b) Focused view of the binding site. NRP-1 interaction loops and the specific residues participating in the binding are labeled. Specific interactions between loop residues and NRP-1 are drawn. Hydrogen bonds are represented as blue dashed lines, while salt bridges are represented as yellow dashed lines.

VEGF for binding to NRP-1 (*12*). Both features demonstrated that the NRP-1−tuftsin complex could represent an accurate structural template for the binding of any active peptide. On the basis of those findings, in this work we have employed computer simulations to model the RPAR interaction with NRP-1 binding domain, using the NRP-1−tuftsin crystal structure (*11*) as a template. Figure 1a shows the overall NRP-1−tuftsin complex, while Figure 1b focuses on the binding region. Tuftsin binds to NRP-1 through a network of interactions that mainly clusters in the receptor segments called interaction loops. Thus, the interacting side chains that belong to the receptor are located in the three interaction loops identified in Figure 1b. The strongest and most relevant of such interactions is a salt bridge between tuftsin's C-terminal Arg and Asp-320 of NRP-1. Other hydrogen bonding interactions involve the Lys backbone of tuftsin and the side chain of Tyr-297 in NRP-1, and the carboxylate oxygen atoms of tuftsin's C-terminal Arg, which interact simultaneously with the side chains of Ser-346, Thr-349, and Tyr-353 of NRP-1. As we will describe, most of these interactions are also featured by the RPAR segment and NRP-1.

To exploit the RPAR potential in targeted drug delivery, it may be advantageous to increase its half-life by enhancing its resistance to proteases, a goal that can be achieved by introducing single-residue chemical modifications. Such modifications should not alter the ability of RPAR to recognize and bind NRP-1. Thus, the interactions pattern featured by the wild-type segment bound to the receptor is used as a template that any candidate modification must preserve. In a recent study, we successfully used this strategy on CREKA (*13*), a linear pentapeptide that selectively binds to tumor vasculature (*14*), to increase the resistance of the peptide versus proteases. In the CREKA study, we found that replacement of a hydrogen atom with a methyl group at selected amide bonds results in peptide analogues that accumulate in tumors more effectively than the wild-type peptide (*13*). We then used the modified CREKA peptides to design nanoparticles that specifically blocked tumor circulation (*13*). In the work presented here, we have studied possible chemical modifications that could be introduced into RPAR. To discard those that might have greater chances to fail, we have evaluated their potential in silico by assessing their possible interactions, using a simplified model of the NRP-1 binding domain. The modifications introduced into the RPAR peptide have mainly focused on changes in the chemical scaffold of the terminal Arg, because this amino acid plays a key role in binding the receptor. Specifically, backbone and side chain modifications have been examined in detail. We also changed the directionality of the last peptide bond into a retro-peptide bond (*15*), and in parallel, we inverted the chiralities of all the amino acids in the peptide.

Our results constitute a first step in a comprehensive study of this important peptide−receptor binding, which induces cell internalization and tissue penetration. Computer modeling studies may improve our understanding of tissue penetration mechanisms and aid in increasing the efficiency of the delivery of drugs into tumors using the R/K/XXR/K motif.

METHODS

Molecular Dynamics Simulations for Modeling NRP-1−RPAR Interaction. Calculations were performed using the NAMD package (*16*). All the atoms of the system were considered explicitly, and the energy was calculated using the AMBER ff03 force field (*17*, *18*). Water molecules were represented using the TIP3 model (*19*). Simulations were performed using the NVT ensemble in an orthorhombic simulation box. Periodic boundary conditions were applied using the nearest image convention. The box size was adjusted to fit the complex size, the initial box dimensions being 114.0 Å × 93.0 Å × 107.0 Å to ensure infinite dilution.

The system contained 111486 atoms, of which 5040 were protein atoms and the rest were ions and solvent atoms. The starting molecular structure was modeled on the basis of the tuftsin-bound complex (PDB entry 2ORZ). The charge of all potential titratable groups was fixed at values corresponding to neutral pH (i.e., all Asp and Lys side chains were represented in their negatively and positively charged forms, respectively). Simulations were performed considering an ionic strength of approximately 0.8% (w/w) (66 Cl^- and 57 Na^+ ions) at 298 K. The overall charge of the system was kept neutral for the use of particle mesh Ewald summation to calculate the electrostatic interactions (*20*) (i.e., counterions being Cl^- and Na^+).

The protocol for MD simulations was as follows. Before each simulation was run, the potential energy of each system was minimized using 5000 conjugate gradient steps. To ensure a uniform distribution of the ions in solution, the protein atoms were held fixed while the solvent was heated to 500 K. Next, we let the system equilibrate at 298 K with NPT conditions to allow the water box size to shrink to its final dimensions, preventing a low water density. After the water box size was stabilized at approximately 112 Å × 89 Å × 104 Å, we heated and equilibrated the system. The heating protocol included a 15 ps period during which the temperature of the system was increased from 0 K to the final temperature of 298 K plus a 500 ps equilibration period.

A residue-based cutoff was applied at 14.0 Å; i.e., the interaction was evaluated when two residues had any atoms separated by a shorter distance. A numerical integration time step of 1 fs was used for all the simulations. The nonbonded pair list was updated every five steps, and the trajectories were saved every 1000 steps (1 ps) for subsequent analysis. During the production runs, the system was held in a constant thermal bath at 298 K using thermal coupling. The simulation was run for a period of 40 ns. To ensure that the system reached equilibrium, the first 4 ns were excluded from the analyses, because during that time the

Table 1: NRP-1 Sequences Used To Build a Simplified Model of the Receptor Binding Pocket[a]

291	292	293	294	295	*296*	*297*	*298*	299	300	301			
I	T	A	S	S	*Q*	*Y*	*G*	T	D	W			
312	313	314	315	316	317	318	*319*	*320*	*321*	322	323		
E	N	G	W	T	P	G	*E*	*D*	*S*	Y	R		
343	344	*345*	*346*	*347*	*348*	*349*	*350*	*351*	*352*	*353*	*354*	355	356
G	A	*I*	*S*	*K*	*E*	*T*	*K*	*K*	*K*	*Y*	*Y*	V	K
411	412	413	414	415	416	417							
W	D	T	G	I	S	M							

[a]For each peptide segment, the original NRP-1 sequence position and the residue in single-letter code are specified. Residues that are italic define the three interaction loops cited in the text: residues 296−298, 319−321, and 345−354 constitute loops I, II, and III, respectively (*11*).

potential energy continued to increase slightly (see the next section). Analyses were conducted using NAMD (*16*), Visual Molecular Dynamics (VMD) (*21*), and MatLab.

Targeted Modifications of RPAR. (*i*) *A Simulation System Based upon a Simplified Receptor*. The results provided by the MD simulations of the wild-type complex allowed a topological description of the most relevant interactions of NRP-1−RPAR binding. This interaction map was used to build a simplified receptor model featuring the most important chemical groups. This model is represented by those receptor residues that enclose the whole binding cavity, including the three interaction loops that are listed in Table 1 (these loops are displayed in Figure 1 for the NRP-1−tuftsin complex). The use of a reduced model for the NRP-1 receptor allows us to screen for many candidates of RPAR mutants with a significant reduction in computational requirements. At the same time, this strategy allows us to overcome potential drawbacks that other approaches would present. If we were to freeze the less active parts of the receptor, computational resources would not be saved while the receptor would lose conformational flexibility, which is always an important factor for the formation and stabilization of any given ligand−receptor model. The model itself was then constructed by excision of small segments of the original protein that included the residues that were in direct or indirect contact with the RPAR peptide inside the receptor. For each excised segment, their edge residues (i.e., the actual length of each segment) were selected to completely cover the binding pocket space and to ensure charge neutrality of the system. Moreover, each peptide segment that defined part of the receptor model was anchored at both edges (N- and C-termini) via a harmonic potential to an inert surface, constituted by van der Waals spheres, similar to the strategy followed in ref *22*. The inert surface consisted of 100 spheres regularly distributed in a 47.5 Å × 47.5 Å square. Finally, the receptor model bound to each RPAR-modified peptide (see below) was solvated using a water cap radius of 23.75 Å with a total of 851 TIP3 water molecules (*19*).

(*ii*) *RPAR Analogues Studied*. Thirteen different RPAR analogues were built via replacement of the C-terminal Arg with a synthetic surrogate. The details of each molecular analogue and the rationale for its design are explained in Results and Discussion (Figure 4).

A chemical modification was considered to be potentially useful for further biomedical applications if it follows three conditions (listed from more to less relevant): (1) The docking in the binding site must maintain the interactions mapped for the wild type during the simulation. (2) The binding of the peptide should not cause a conformational rearrangement of the binding pocket (i.e., the rmsd with respect with the initial receptor topology should oscillate around the values observed for the wild type). (3) The energy for interaction between the residues of the receptor interacting with the RPAR analogue should be close to the values obtained for the wild-type RPAR.

(*iii*) *Simulation Protocol and Energy Parameters*. For each studied RPAR analogue, short MD simulations were performed using the AMBER9 package (*23*). Before each simulation was run, the potential energy of each system was minimized using 5000 conjugate gradient steps. The heating protocol included a 15 ps period udring which the temperature of the system was increased from 0 K to the final temperature of 298 K plus a 500 ps thermal equilibration. A residue-based cutoff was applied at 14.0 Å for the nonbonding interactions. A numerical integration time step of 2 fs was used for all the simulations. The nonbonded pair list was updated every five steps, and the trajectories were saved every 1000 steps (2 ps) for subsequent analysis. During the production runs, the system was held in a constant thermal bath at 298 K using thermal coupling. Each studied system, which includes a RPAR analogue, the receptor model, and the water cap, was simulated for a period of 5 ns.

All parameters were extrapolated from the AMBER ff03 force field (*18*) for each proposed synthetic amino acid with the exception of the electrostatic parameters that were explicitly computed. Atomic centered charges were calculated by fitting the (U)HF/6-31G(d) quantum mechanical and the Coulombic molecular electrostatic potentials (MEPs) to a large set of points placed outside the nuclear region. The electrostatic parameters derived at this level of theory are fully compatible with the current parameters of the AMBER force field (*17*). For each new amino acid analogue, the electrostatic parameters were fitted using as the main chain conformation the $C_{7,eq}$ conformation, which usually is the lowest-energy conformation in small model dipeptides.

RESULTS AND DISCUSSION

Analysis of the Simulated NRP-1−RPAR Complex. Figure 2a shows the temporal evolution of the overall potential energy of the complex through the MD simulation. The energy kept increasing at the beginning of the simulation and stabilized after 4 ns. Such an energy increase can be attributed to the initial position of the RPAR peptide in the receptor pocket, which was initially modeled according to the position of tuftsin (*11*). Accordingly, the first 4 ns was excluded from the numerical analyses below to ensure that the complex is properly equilibrated. Figure 2b displays the evolution of the ligand−protein complex rmsd, which was calculated with respect to the minimized structure. As one can see, the structure of the complex is very stable and the rmsd stays within the range of 1.2−1.25 Å from the initial conformation throughout the simulation. This stability is due to the combination of a good geometric match between NRP-1 and RPAR and a strong network of interactions that holds the complex tightly bound.

To further validate our results, we tested specific interactions that hold the RPAR molecule in its binding pockets. The default criteria of VMD (i.e., 3 Å distance and 160°) were used to define a hydrogen bond. We found the following interactions: (1) The C-terminal Arg of RPAR is held by a number of interactions, the strongest of which is a salt bridge with Asp-320 of NRP-1. Figure 2c shows the temporal evolution of the distance between the two charged centers of these two amino acids throughout the whole simulation. As one can see, the distance fluctuates between 3.4 and 3.5 Å, indicating a strong and stable salt bridge. (2) The

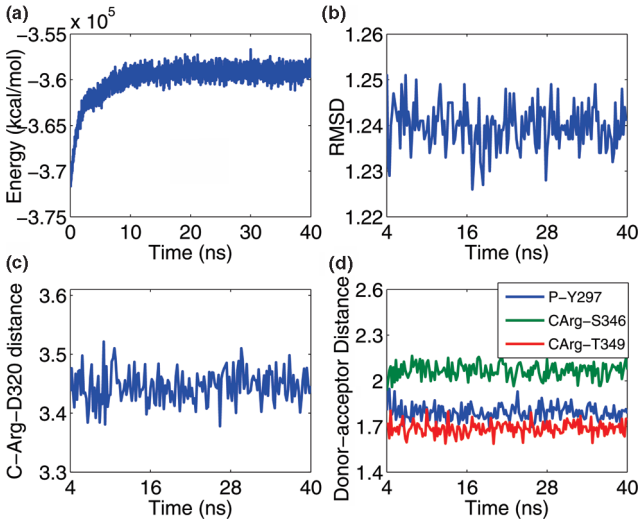

FIGURE 2: NRP-1−RPAR trajectory analysis. (a) Temporal evolution of the potential energy for the NRP−RPAR complex during the MD simulation. (b) Temporal evolution of the rmsd of the NRP-1−RPAR complex during the MD simulation. The rmsd was calculated with respect to the minimized structure and was measured considering the backbone atoms only, to reduce noise. (c) Temporal evolution of the distance between Asp-74 of NRP-1 and the C-terminal Arg of RPAR. The distance is that between the charged centers of these two amino acids. (d) Distances between the donor and acceptor centers of the three conserved hydrogen bonds between NRP-1 and RPAR. See the legend for details. CArg refers to the C-terminal Arg residue of RPAR.

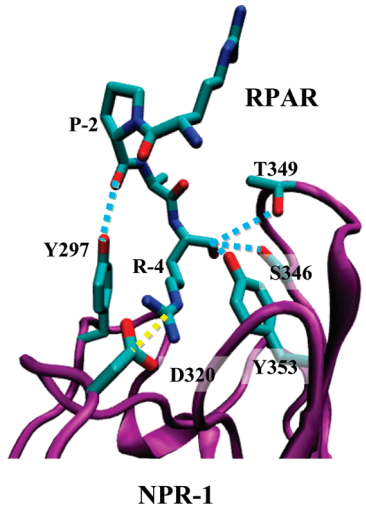

FIGURE 3: Ribbon representation of the NRP-1−RPAR complex showing the most notable interactions found between the peptide and the binding pocket of NRP-1. The ligand and the interacting side chains of the receptor are depicted as solid lines. The NRP-1 backbone is colored purple and the RPAR backbone green. Hydrogen atoms have been omitted for the sake of clarity. Specific interactions are drawn: hydrogen bonds shown as blue dashed lines and salt bridges as yellow dashed lines.

hydrogen bond between the backbone oxygen of the Pro of RPAR and the side chain of Tyr-297 of NRP-1 is well preserved, as reflected in Figure 2d, which shows that the hydrogen bonding distance remains <2 Å over the whole trajectory. (3) A bifurcated hydrogen bond involves the oxygen of the C-terminal Arg of RPAR as a hydrogen bonding acceptor and the side chains of Ser-346 and Thr-349 of NRP-1 as hydrogen bonding donors. The two interactions exist in ~98% of the trajectory. The distances between the donor and acceptor centers of this interaction are illustrated in Figure 2d. As one can see, the donor−acceptor distances are ≤ 2 Å throughout the entire simulation, indicating well-preserved hydrogen bonds.

Figure 3 presents the interaction network, showing that the stronger interactions are near the binding pocket. This allows us to build a molecular template for designing RPAR surrogates based on chemical modifications at specific positions. Table 2 lists the average pairwise interaction energies between the amino acids participating in the interaction discussed above. Averages were calculated over the entire trajectory, and the energy, which was calculated over all the atoms in the participating amino acids and not just in the interacting centers, includes the nonbonded contributions (i.e., van der Waals and electrostatics) with self-energy excluded. The four interactions contribute a highly negative energy indicating strong binding. As expected, the energy of the interaction between the C-terminal Arg of RPAR and Asp-320 is the strongest because of the electrostatic nature of this salt bridge, which is formed between the two charged centers. On the other hand, the energy associated with the interaction between the Pro of RPAR and Tyr-297 corresponds to that typically found in conventional hydrogen bonds. This interaction provides further anchoring of RPAR to the binding site outside the binding pocket. Finally, the energy found for the bifurcated interaction between the C-terminal Arg of RPAR and the side chains of Ser-346 and Thr-349 is intermediate between those of a salt bridge and a conventional hydrogen bond. This is because of the negative charge of the carboxylate of Arg, which enhances the electrostatic nature of the interaction with respect to that of conventional hydrogen bonds.

Almost all the interactions discussed above were also observed in the crystal structure of the NRP-1−tuftsin complex (PDB entry 2ORZ) (*11*), which was used to model the NRP-1−RPAR complex in this study. Thus, these findings suggest that the interactions identified in both NRP-1−tuftsin and NRP-1−RPAR complexes are crucial for the binding of NRP-1 with its ligands. These interactions anchor the C-terminal Arg in the binding pocket and emphasize the importance of the C-terminal Arg in binding to NRP-1 as denoted by the C-end rule (*9*). The additional highly preserved hydrogen bond involving Thr-349 of NRP-1 and the Pro backbone in RPAR (Lys backbone in tuftsin, also in position 2 of the peptide) helps in maintaining the geometry of the ligand and further stabilizing the complex binding. The structural differences between Lys and Pro backbones hinder the adoption of an ideal geometry for such a hydrogen bond when RPAR is docked instead of tuftsin. In spite of this, the distance between the two interacting centers remains very short, as shown in Figure 2d.

On the other hand, an additional hydrogen bond was detected in the NRP-1−tuftsin complex (*11*) but not in the NRP-1−RPAR complex. This corresponds to the hydrogen bond between the backbone C-terminal Arg of the peptide and the side chain of Tyr-353. This interaction did not exist in our simulation, the average distance between the donor and acceptor atoms being approximately 3.8 Å. The absence of this additional hydrogen bond has been attributed to the different geometry of the RPAR peptide with respect to tuftsin, which is a consequence of the rigidity of the Pro residue in RPAR. Because the binding of RPAR and NRP-1 remains strong, as indicated by the rmsd and the other well-preserved interactions, we conclude that one missing interaction is not critical for the binding, especially if the two other interactions exist between the backbone of the C-terminal

Table 2: Averaged Interaction Energies for the Specific Interactions Observed in the NRP-1−RPAR Complex[a]

NRP-1−RPAR interaction	average energy (kcal/mol)
Ser-346−C-Arg	−25.2 ± 0.6
Thr-349−C-Arg	−23.4 ± 0.7
Asp-320−C-Arg	−78.7 ± 0.9
Tyr-297−Pro-2	−6.0 ± 0.2

[a]The energies, which were averaged over the whole trajectories, were calculated considering the nonbonding contributions (van de Waals and electrostatics). The first amino acid refers to NRP-1 and the second to RPAR. C-Arg corresponds to the C-terminal Arg in RPAR. Interaction energies were evaluated considering all the atoms in the participating amino acids and not just the interacting centers.

Arg of RPAR and the side chains of NRP-1. Moreover, even though the distance between Tyr-353 and the C-terminal Arg is too large for the bond to be considered as a strong interaction, both polar groups form a weak charge−dipole interaction that evidently contributes to the global stability of the complex.

Most recently, we began expanding the search for other molecular organizations that could act as NRP-1 ligands. Our preliminary results are quite unexpected. Simplified models that mimic protein loops and that were based on an oligopeptide rich in glycine residues suggested potential binding between the NRP-1 receptor and species that do not follow the C-end rule. Thus, three different heptapeptides (G_7, G_3RG_3, and G_4RG_2), all of them featuring a type I β-turn, were probed to see if they could dock in the receptor binding site. Not only was G_7 able to dock, causing a minimal deformation of the receptor topology, but also peptides that presented arginine residues (such as G_3RG_3 or G_4RG_2) could be inserted into the binding site with only a small deformation of the receptor. The average rmsd between the topology of the binding site when bound to RPAR, compared to the heptapeptides, was less than 0.30 Å for G_7 and 0.45 Å for the other glycine-based peptides. These results may suggest alternative therapeutic approaches beyond the mutant analogues that are presented below.

New RPAR Analogues Resistant to Proteolysis. (*i*) *Rational Design of Synthetic Residues*. To explore possible chemical modifications that would enhance the affinity of RPAR for NRP-1 while hindering its proteolytic degradation, we focused on two main approaches. In the first, we attempted to optimize the docking of the homing peptide to the receptor by introducing chemical modifications into the side chain of the most active residue, the C-terminal Arg. Specifically, we modified the aliphatic segment of the Arg side group, which should not dramatically affect the interaction between the charged guanidinium group and the carboxylate side chain of Asp-320. The investigated analogues are schematically represented in Figure 4a and labeled II−VI. The modifications inserted in the Arg template were either based on the incorporation of an unsaturated bond into the aliphatic segment (Figure 4a, analogues II−IV) or based on the variation of the number of methylene units (Figure 4a, analogues V and VI). Two isomers, which differ in the relative position of the non-hydrogen atoms attached to the doubly bounded carbon atoms, i.e., cis and trans isomers, are obtained by incorporation of an unsaturated aliphatic segment. The two isomers have been investigated in all cases.

In the second approach, we tried to make the peptide more resistant to proteases by following two strategies. In the first one, we considered chirality. If the chirality of one or more residues of the RPAR peptide is to be inverted, cleavage could be prevented

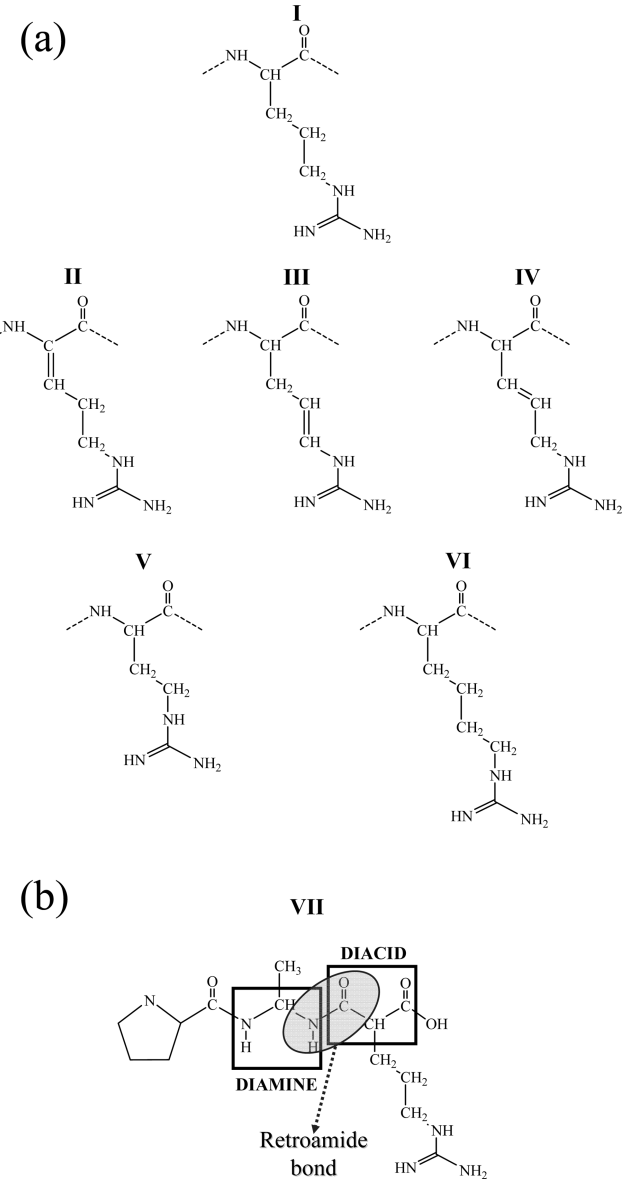

FIGURE 4: Schematic representation of the most relevant chemical modifications introduced on the Arg amino acid. (a) Modifications on the side chain: (I) coded Arg, (II) α-dehydro Arg, (III) β-dehydro Arg, (IV) γ-dehydro Arg, (V) Arg in which the side chain has been reduced by one methylenic unit (also denoted nor-arginine), and (VI) Arg in which the side chain has been enlarged by one methylenic unit (also denoted homo-arginine). (b) Retromodification of the main chain to obtain a retroamide bond (retro-Arg) (VII).

because of the loss of the protease recognition pattern. However, in a previous study, it was shown that a D-peptide, in which the chirality of all the residues was simultaneously inverted, did not interact with NRP-1 (*9*). This made us reduce the modification range to single mutations (i.e., one-by-one residue). Another potential tool for obstructing the proteolytic action is the modification of the amide bond. Accordingly, we examined the possibility of changing the directionality of the amide bond by introducing a retromodification, which consists of the substitution of two consecutive amino acids with the *gem*-diamine and the acid, respectively, as indicated in Figure 4b. The net effect over the peptide scaffold is a partial inversion of the local directionality of the peptide main chain, the resulting compound being known as a retropeptide (*15*).

Table 3: Most Relevant Structural and Chemical Properties Featured by Each RPAR Analogue Interacting with NRP-1 Averaged over the 5 ns MD Simulation[a]

RPAR analogue	average rmsd (Å)	% preserved interactions[b]	total interaction energy (kcal/mol)[c]
(D)RPAR	1.25 ± 0.9	97	−125.3 ± 3.4
R(D)PAR	2.01 ± 1.0	85	−94.0 ± 7.7
RP(D)AR	2.21 ± 1.1	79	−88.4 ± 5.6
RPA(D)R	1.31 ± 0.5	96	−129.9 ± 3.2
I	1.24 ± 0.2[d]	99[e]	−133.6 ± 2.5[f]
II			
cis	1.44 ± 0.7	78	−98.5 ± 3.3
trans	1.29 ± 0.6	98	−129.3 ± 3.9
III			
cis	1.85 ± 0.4	96	−111.1 ± 5.1
trans	1.92 ± 0.5	93	−109.4 ± 4.7
IV			
cis	1.99 ± 0.6	91	−101.0 ± 2.3
trans	1.34 ± 0.3	97	−122.6 ± 3.5
V	1.81 ± 0.91	72	−77.2 ± 5.7
VI	2.18 ± 0.94	73	−89.2 ± 6.1
VII	1.25 ± 0.3	99	−131.3 ± 2.9

[a]Averaged values for wild-type RPAR (system I) were obtained using the last 5 ns of the 40 ns simulation. [b]Computed as the percentage of preserved specific interactions analyzed in Table 2 over the 5 ns simulation. [c]The total interaction energy was determined by summing the individual interaction energies, which were obtained for each specific interaction as indicated in Table 2. [d]Values extracted for those residues that constitute the receptor model. [e]Values extracted combining information from panels c and d of Figure 1. [f]Values of the individual interactions energies extracted from Table 2.

All these analogues were systematically tested by running MD simulations, using our simplified receptor model. A complete list of the simulations is provided in Table 3. For each analogue family, its potential usefulness was assessed on the basis of the requirements noted above: a good candidate should not distort the receptor organization around the binding site and should preserve the interactions that characterized the association of the RPAR peptide with the receptor.

(*ii*) *Simulations of RPAR Analogues*. Table 3 lists the averaged rmsd, the percentage of preserved interactions, and the total interaction energy values, which were calculated by summing the interaction energies determined for all the specific interactions (i.e., adding the pairwise interaction energies between the amino acids participating in all salt bridges and hydrogen bonds), for all the investigated systems. As expected, results for D-amino acid replacements depended on the D-residue position: the potential affinity for the receptor decreased when the mutated residues were in the central segment of the peptide because they affected the orientation of the RPAR analogue in the binding pocket. Even though in all cases the interaction pattern was acceptably preserved, with percentages of >80% of the total amount for 5 ns, only (D)RPAR and RPA(D)R satisfied all the screening criteria. Replacements in the central site distorted the receptor organization, as reflected by the higher rmsd values, reducing the interaction strength as the computed higher energy suggests (system I in Table 3). On the other hand, replacement of any terminal Arg residue does not perturb the formation of the receptor−peptide complex, giving very similar results in both cases. A detailed view of the organization obtained for the RPA(D)R system is shown in Figure 5a.

The modeling of all the remaining analogues was based on the introduction of several chemical modifications at the C-terminal Arg. Thus, all present the RPAX motif, where X corresponds to a new noncoded amino acid that replaces a natural Arg in the RPAR peptide. Within this set of molecules, when the length of the methylene segment in the C-terminal Arg was modified, most of the necessary conditions for successful interaction were not satisfied (V and VI in Figure 4a and Table 3). Under such conditions, the relocation of the guanidinium group in the binding pocket, which is guided by the interaction between such a group and the Asp-320 side chain, induces the loss of some interactions between RPAX and the side chains of Thr-349 and Ser-346. Table 2 shows that the percentage of preserved interactions drops below 75% in both studied analogues for the 5 ns simulation. This loss results in important changes in the receptor topology indicating that both analogues, with an average rmsd that almost doubles the structural fluctuations observed in the wild-type segment for the same amount of time, are unlikely.

On the other hand, insertion of unsaturated bonds along the aliphatic segment of Arg leads to binding modes that depend strongly on both the particular topology of the double bond (i.e., cis or trans) and its position in the alkyl segment. The only analogues that were able to fulfill the selection criteria were those in which the geometrical restrictions of the new constitution drove the side chain toward an organization that resembles the conformation adopted by the side chain of the coded Arg in RPAR. Among six possible analogues (II, III, and IV in Figure 4a, each with cis and trans isomers), only three formed the proper interactions with the binding site residues: trans-II (Figure 5b) and cis-III and trans-IV (Figure 5c). However, cis-III should be discarded because its docking produced rearrangements in the receptor organization (with rmsd fluctuations 45% higher than in the wild-type peptide) and diminished the strength of the interaction between the ligand and the receptor. However, both trans-II and trans-IV satisfy the requirements and are good candidates for an RPAR analogue, reaching interaction energies close to the values computed for the wild-type peptide.

Thus, peptide VII, which consists of a RPAX analogue in which X is retro-Arg (Figure 4b), becomes the most promising candidate. This is because the majority of the interactions that stabilize the NRP-1−RPAR complex were supplied either by the Arg side chain or by the terminal carboxylate, and therefore, alteration of the main chain at the N-terminus does not preclude binding of NRP-1. Indeed, the results obtained with retro-Arg proved that this analogue perfectly suits all the previously imposed selection conditions (Figure 5d), showing structural fluctuations

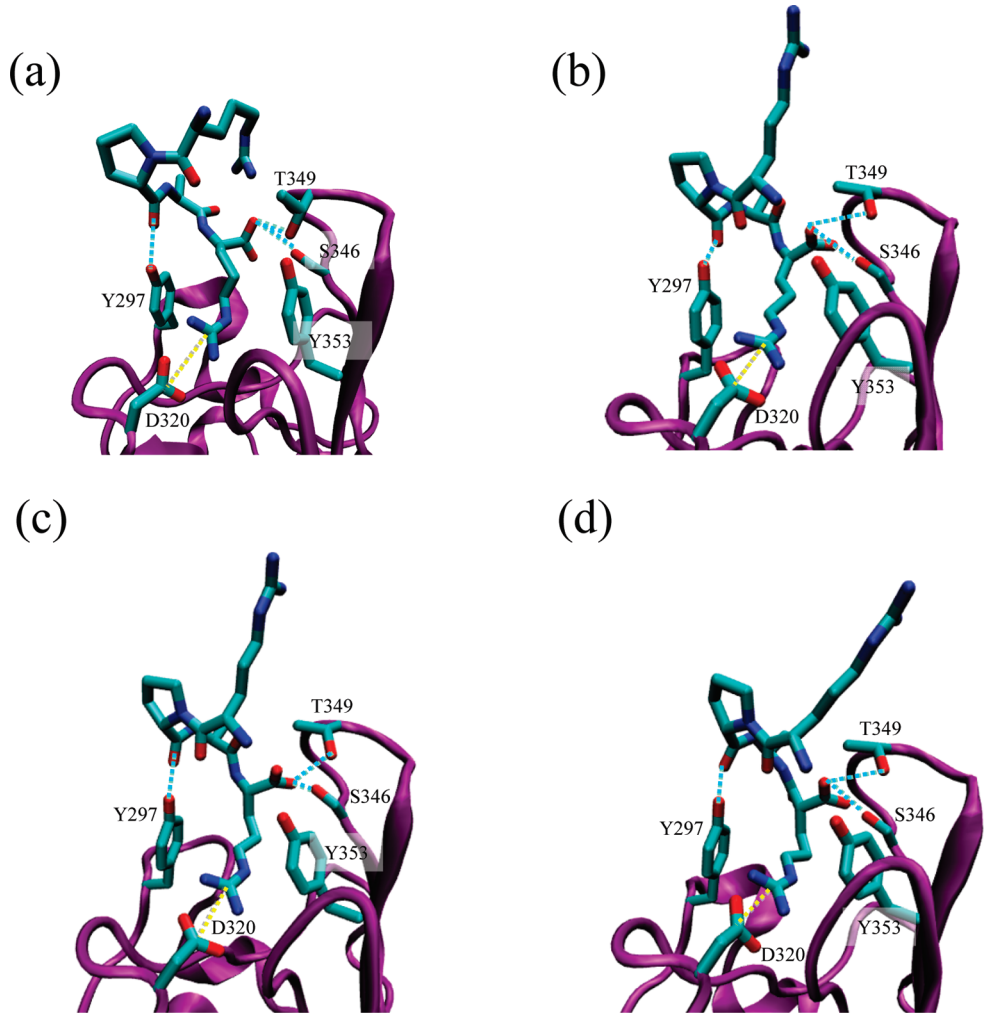

FIGURE 5: Ribbon representation of the NRP-1 binding site with RPAR analogues docking in NRP-1−RPAX complexes, where X is (a) (D)R, (b) trans-II, (c) trans-IV, and (d) retro-Arg. The ligand and the interacting side chains of the receptor are depicted as solid lines. The NRP-1 backbone is colored purple and the RPAX backbone green. Hydrogen atoms have been omitted for the sake of clarity. Specific interactions are drawn: hydrogen bonds blue dashed lines and salt bridges as yellow dashed lines.

and interaction energies almost identical to those observed with the wild-type peptide (Table 3).

CONCLUSIONS

In this study, we used computer simulation methods to model the binding of several peptides to the NRP-1 receptor, which plays an important role in angiogenesis and vascular permeability. Previous work showed that peptides containing the consensus K/RXXK/R sequences promote binding and internalization into tumor cells, while blocking the C-terminal Arg prevented its internalization. This property was coined the C-end rule. The molecular details that allow binding of RPAR to NRP-1 have been characterized in this work. The structural features of the complex have been modeled using as a molecular template the crystal structure of NRP-1 bound to tuftsin, a tetrapeptide that mimics the C-terminal sequence of the tumor-secreted vascular permeability factor (VEGF) (*11*). Furthermore, the specific interactions found in the RPAR−NRP-1 complex have been used to suggest RPAR analogues with improved functional features.

Our simulations have demonstrated that the RPAR−NRP-1 complex is supported by a network of stabilizing hydrogen bonds and salt bridges, similar to that previously observed in the NRP-1−tuftsin cocrystal. The most relevant interaction is a salt bridge between the C-terminal Arg of RPAR and Asp-230 of NRP-1. The stabilization of the C-terminus of RPAR is further enhanced with a bifurcated hydrogen bond between the carboxylate group of the terminal Arg and the side chains of Ser-346 and Thr-349. Finally, RPAR features an extra anchoring point in the NRP-1 pocket through a hydrogen bond between the backbone of the Pro residue of RPAR and the side chain of Tyr-297 of NRP-1. However, this interaction, which is also present in the NRP-1−tuftsin crystal template, is weaker than the other three because it does not involve any charged center, which suggests a secondary role.

The potential therapeutic use of the RPAR peptide as an internalization trigger in anticancer nanodevices requires increasing its resistance to protease cleavage. The use of a reduced model for the NRP-1 receptor combined with short molecular simulations has allowed a quick screening of potential RPAR analogues. The analogues that exhibit a rapid destabilization can be immediately discarded. Those that formed a stable molecular complex should be further studied as potential therapeutic approaches. In this context, we have explored potential RPAR analogues that are the outcome of chemical modifications aimed to make them less sensitive to proteolytic action. The suitability of each modification has been evaluated using molecular simulations by comparing the molecular features of each studied analogue when binding NRP-1 with those characterized for the original RPAR segment. Thus,

modifications that disrupted the formation of the previously described interactions with NRP-1 have been discarded. Among all potential chemical modifications that have been introduced into the RPAR peptide, only those modifications in the C-terminal Arg that either did not alter the shape of the C-terminus (D-Arg and retro-Arg) or did not restrict the orientation of the guanidinium group upon addition of a double bond to the side chain of the C-terminal Arg (trans-II and trans-IV) have been categorized as suitable candidates. When any of those features were absent in a tested analogue, the resulting pattern of interaction between the NRP-1 receptor and the modified peptide did not present all the necessary interacting groups and the internalization activity may be diminished. Therefore, chemical changes in RPAR for increasing the proteolytic resistance need to maintain the carboxylate moiety and should not excessively restrict the conformational freedom of the C-terminal Arg side chain.

ACKNOWLEDGMENT

Computer resources were generously provided by the Centre de Supercomputació de Catalunya (CESCA). We thank the National Cancer Institute for partial allocation of computing time and staff support at the Advanced Biomedical Computing Center of the Frederick Cancer Research and Development Center and the high-performance computational capabilities of the Biowulf PC/Linux cluster at the National Institutes of Health (http://biowulf.nih.gov).

REFERENCES

1. Rabbani, M. L., and Rogers, P. A. (2001) Role of vascular endothelial growth factor in endometrial vascular events before implantation in rats. *Reproduction 122*, 85–90.
2. Maeda, H., Fang, J., Inutsuka, T., and Kitamoto, Y. (2003) Vascular permeability enhancement in solid tumor: Various factors, mechanisms involved and its implications. *Int. Immunopharmacol. 3*, 319–328.
3. Senger, D. R., Galli, S. J., Dvorak, A. M., Perruzzi, C. A., Harvey, V. S., and Dvorak, H. F. (1983) Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid. *Science 219*, 983–985.
4. Baban, D. F., and Seymour, L. W. (1998) Control of tumour vascular permeability. *Adv. Drug Delivery Rev. 34*, 109–119.
5. Ferrara, N., and Henzel, W. J. (1989) Pituitary follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells. *Biochem. Biophys. Res. Commun. 161*, 851–858.
6. Ferrara, N., Gerber, H. P., and LeCouter, J. (2003) The biology of VEGF and its receptors. *Nat. Med. 9*, 669–676.
7. Takashima, S., Kitakaze, M., Asakura, M., Asanuma, H., Sanada, S., Tashiro, F., Niwa, H., Miyazaki, J., Hirota, S., Kitamura, Y., Kitsukawa, T., Fujisawa, H., Klagsbrun, M., and Hori, M. (2002) Targeting of both mouse neuropilin-1 and neuropilin-2 genes severely impairs developmental yolk sac and embryonic angiogenesis. *Proc. Natl. Acad. Sci. U.S.A. 99*, 3657–3662.
8. Staton, C. A., Kumar, I., Reed, M. W. R., and Brown, N. J. (2007) Neuropilins in physiological and pathological angiogenesis. *J. Pathol. 212*, 237–248.
9. Teesalu, T., Sugahara, K. N., Kotamraju, V. R., and Ruoslahti, E. (2009) C-end rule peptides mediate neuropilin-1 dependent cell, vascular and tissue penetration. *Proc. Natl. Acad. Sci. U.S.A. 106*, 16157–16162.
10. Sugahara, K. N., Teesalu, T., Karmali, P. P., Kotamraju, V. R., Agemy, L., Greenwald, D. R., and Ruoslahti, E. (2010) Coadministration of a Tumor-Penetrating Peptide Enhances the Efficacy of Cancer Drugs. *Science 328*, 1031–1035.
11. Vander Kooi, C. W., Jusino, M. A., Perman, B., Neau, D. B., Bellamy, H. D., and Leahy, D. J. (2007) Structural basis for ligand and heparin binding to neuropilin B domains. *Proc. Natl. Acad. Sci. U.S.A. 104*, 6152–6157.
12. von Wronski, M. A., Raju, N., Pillai, R., Bogdan, N. J., Marinelli, E. R., Nanjappan, P., Ramalingam, K., Arunachalam, T., Eaton, S., Linder, K. E., Yan, F., Pochon, S., Tweedle, M. F., and Nunn, A. D. (2006) Tuftsin binds neuropilin-1 through a sequence similar to that encoded by exon 8 of vascular endothelial growth factor. *J. Biol. Chem. 281*, 5702–5710.
13. Agemy, L., Sugaharai, K. N., Kotamraju, V. R., Gujraty, K., Girard, O. M., Kono, Y., Mattrey, R. F., Park, J. H., Sailor, M. J., Jiménez, A. I., Cativiela, C., Zanuy, D., Sayago, F. J., Alemán, C., Nussinov, R., and Ruoslahti, E. (2010) Nanoparticles-induced vascular blockade in human prostate cancer. *Blood 116*, 2847–2856.
14. Simberg, D., Duza, T., Park, J. H., Essler, M., Pilch, J., Zhang, L., Derfus, A. M., Yang, M., Hoffman, R. M., Bathia, S., Sailor, M. J., and Ruoslahti, E. (2007) Biomimetic amplification of nanoparticle homing to tumors. *Proc. Natl. Acad. Sci. U.S.A. 104*, 932–936.
15. Alemán, C., and Puiggali, J. (1995) Retromodified Residues: Small Peptides and Polymers. Interactions, Force-Field Parametrization and Conformational Analyses. *J. Org. Chem. 60*, 910–924.
16. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781–1802.
17. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force-Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc. 117*, 5179–5197.
18. 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., and Kollman, P. A. (2003) Point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem. 24*, 1999–2012.
19. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926–935.
20. Darden, T., Perer, L., Li, L., and Pedersen, L. (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure 7*, R55–R60.
21. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual Molecular Dynamics. *J. Mol. Graphics 14*, 33–38.
22. Zanuy, D., Flores-Ortega, A., Casanovas, J., Curco, D., Nussinov, R., and Aleman, C. (2008) The Energy Landscape of a Selective Tumor-Homing Pentapeptide. *J. Phys. Chem. B 112*, 8692–8700.
23. Case, D. A., Darden, T. A., Cheatham, T. E., III, 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., Matthews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) AMBER 9, University of California, San Francisco.