ACS chemical biology

ARTICLES

pubs.acs.org/acschemicalbiology

Paramagnetic Relaxation Assisted Docking of a Small Indole Compound in the HIV-1 gp41 Hydrophobic Pocket

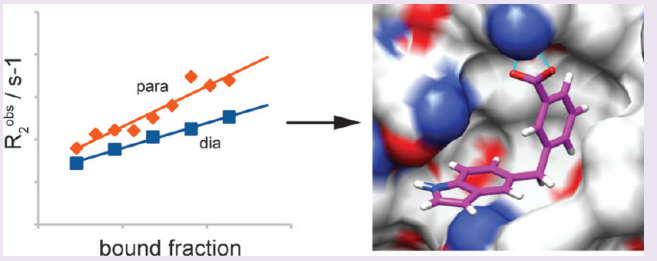
Miriam Gochin,[†,‡,*] Guangyan Zhou,[†] and Aaron H. Phillips[†]

[†]Department of Basic Sciences, Touro University-California, Vallejo, California 94592, United States

[‡]Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, United States

Supporting Information

ABSTRACT: The hydrophobic pocket contained within the gp41 coiled coil is an important target for small molecules designed to inhibit HIV-1 fusion. While various screening experiments have identified molecules purported to bind in this pocket, few have confirmed details of the interaction, instead relying on computational docking to predict the binding mode. This is made more challenging by the fact that residues lining the hydrophobic pocket are highly flexible, as is typical for a protein−protein interaction site, limiting the predictive power of computational tools. In this study, we report on an NMR method to define the binding mode of **1-5i**, a compound in a series of newly developed indole inhibitors. We show that paramagnetic relaxation enhancement of ligand protons due to an MTSL group positioned close to the binding pocket could be applied quantitatively to distinguish between more than 30 different computational poses, selecting a single pose that agreed with the NMR data. In this pose, important hydrophobic and polar contacts occur with pocket lysine, tryptophan, and glutamine residues, including putative hydrogen bonds between the ligand carboxylate and the lysine ε-amino group. A study of the ligand orientation suggests directions for optimization.



There has been a great deal of interest in discovering low molecular weight compounds that bind to the hydrophobic pocket of gp41, as a means to inhibit HIV-1 fusion with a small molecule. The hydrophobic pocket is critical for stability of the six-helix bundle that forms during the fusion reaction,[1] and it is considered a hotspot for inhibiting the protein−protein interaction between the N-heptad and C-heptad repeat regions of gp41. While there have been many studies involving screening of small molecule libraries to identify fusion inhibitors in biological or biochemical assays,[2−7] there has been very little structural information to confirm that the mechanism of inhibition is hydrophobic pocket binding. Furthermore, the plasticity of the protein−protein interaction surface makes it difficult to rely strictly on computational docking studies[8] as evidence of binding or to provide orientational information.

We have recently described a series of indole compounds with confirmed hydrophobic pocket binding activity that is correlated to inhibition of cell−cell fusion.[9] The core scaffold, compound **1**, called **1-6i** in this study, was a small molecule with molecular weight 251 and with good inhibitory potential in the low micromolar range for both hydrophobic pocket binding and cell−cell fusion inhibition. We have continued to investigate the structure−activity relationship properties of this compound and its derivatives, including making an isomer of the compound with a differently substituted indole, **1-5i** (Figure 1). Intriguingly, **1-5i** has 10-fold reduced affinity for the hydrophobic pocket compared with **1-6i**. Modeling predictions do not adequately explain the observed difference and are quite dependent on the receptor structure that

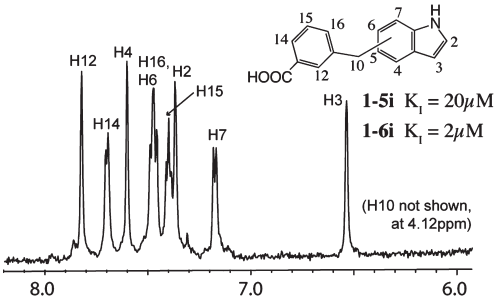


Figure 1. Structure of **1-5i** (5-substituted indole) and **1-6i** (6-substituted indole) and NMR spectrum of **1-5i**. Resonances of protons H6 and H16 overlapped and were not used in the analysis. Dissociation constants measured by fluorescence using a hydrophobic pocket binding assay[5] are indicated for the two isomers.

is used in analysis. There are 170 protein data bank entries for gp41,[10] and they illustrate the variability typical of a protein−protein interface. The base of the pocket, including main chain atoms and residues internal to the coiled coil, is invariant between structures, but side chains of surrounding residues adopt multiple conformations, likely associated with differences in structural resolution, length of the peptides used in structure determination, and the induced fit associated with a bound inhibitor or peptide.

Received: November 12, 2010
Accepted: December 14, 2010
Published: December 15, 2010

It has proved very challenging to determine the structure of small molecules bound in the hydrophobic pocket of gp41, since it has not been possible to crystallize a complex of gp41 bound to a small molecule. Limited information has been obtained from NMR studies in two instances.[11,12] We showed how we could extract paramagnetic relaxation rates for a fast exchanging ligand binding to the hydrophobic pocket, using a spin-labeled peptide MTSL-C29e5.0 that bound adjacent to the pocket of receptor complex Fe(env5.0)$_3$. Using a simple model of a fixed MTSL side chain, we were able to demonstrate the principle of ligand structure determination by second site screening for a weakly binding inhibitor.[11] Here we extend this study to **1-5i**, a ligand of moderate affinity within the indole series. We have used the results to establish a protocol for applying the structural constraints and to determine the resolution and limitations of the method. We found that the paramagnetic relaxation enhancements could be used quantitatively to define ligand orientation, as well as to suggest a receptor structure useful for computational predictions of small molecule binding. We also hoped to understand not only the binding orientation of **1-5i** but also the structure−activity relationship that confers much higher affinity for **1-6i**.

METHODS

Materials. Chemicals were obtained from Sigma, Inc. or Synthonix, Inc. NHR (env5.0) and CHR (C29-e5.0) peptides for the study were obtained by solid state synthesis from Biosynthesis Inc. or Bio-Peptides. env5.0 has the sequence bpy-GQAVSGIVQQQNNLLRAIEAQQH-LLQLTVWGIKQLQARILAVEKK-NH_2, where bpy is 5-carboxy-2,2′-bipyridine. Three bipyridine units coordinate one ferrous ion, creating a trimeric complex representing the gp41 NHR coiled coil.[13] C29-e5.0 has the sequence Ac-CYTSLIESLIRESQEQQEKNEQELRELDK- NH_2, modified from the wild-type CHR to contain several salt-bridges and an N-terminal cysteine residue. **1-5i** was synthesized according to the procedure described in the Supporting Information. MTSL (*S*-(2,2,5,5-tetramethyl-2,5-dihydro-1*H*-pyrrol-3-yl)methyl methanesulfonothioate) was purchased from Toronto Research Chemicals and attached using standard procedures at the N-terminal cysteine of C29e5.0.

NMR Measurements. Samples of **1-5i** for NMR measurements contained 10% d_6-DMSO, 25 mM d_6-tris, 25 mM d_3-acetate, and 0.1 mM d_4-TSP in 100% D_2O at pD 6.8. In each experiment, the concentration of **1-5i** was 200 μM, and the concentrations of Fe(env5.0)$_3$ and C29-e5.0 (either spin-labeled or unlabeled) were varied between 2 and 10 μM. NMR measurements were recorded on a Bruker Avance DMX 600 spectrometer equipped with a cryogenic probe. R_2 relaxation parameters were determined from a CPMG experiment with presaturation to suppress the residual water peak.[14−16] The delay between 180° pulses was 100 μs in order to suppress relaxation due to the *J*-coupling between geminal protons. Relaxation parameters were fit to two points, with CPMG pulse trains of duration 0.57 and 710 ms with a third point collected at 355 ms to fit fast relaxing resonances.[17] Chemical shift assignments of **1-5i** were confirmed with COSY spectrum.

The transverse relaxation rate of the diamagnetic samples was fit to a straight line as a function of the fraction of the bound ligand, f_b, which varies as follows for a ligand in fast exchange:[11]

$$R_{2obs}{}^{dia} = f_b(R_{2b}{}^{dia} - R_{2f}) + R_{2ex} + R_{2f} \quad (1)$$

R_{2f} and $R_{2b}{}^{dia}$ are the relaxation rates of free and bound ligand, respectively, and R_{2ex} is the exchange contribution to the relaxation. In fast exchange, R_{2ex} is proportional to f_b. f_b was calculated from the **1-5i** inhibition constant of 20 μM, which was determined by competitive inhibition.[5] Similarly, the observed relaxation rate of the ligand in the paramagnetic samples depends linearly on the product of f_b with the fractional occupancy of the spin-labeled probe peptide, f_b'.

$$R_{2obs}{}^{para} = f_b f_b{}' \Gamma_2 + R_{2obs}{}^{dia} \quad (2)$$

f_b' was determined from the inhibition constant of peptide binding, which was 3 μM for unlabeled peptide.[11] Γ_2 is the transverse relaxation rate of the bound ligand in the paramagnetic complex. Subtraction of the diamagnetic best-fit line from the paramagnetic data and plotting against $f_b \cdot f_b'$ yields Γ_2, which is related to the distance between the unpaired electron on the MTSL spin-label and the proton in question according to the Solomon−Bloembergen equation:[18]

$$\Gamma_2 = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_H^2 g^2 \mu_B^2 S(S+1)[4J(0)] + 3J(\omega_H) \quad (3)$$

$$J(\omega) = r^{-6}\frac{\tau_c}{1+(\omega_H\tau_c)^2} \quad (4)$$

where μ_0 is the permeability of free space, γ_H is the gyromagnetic ratio for the hydrogen nucleus, g is the electronic g-factor, μ_B is the Bohr magneton, and S is the number of unpaired electrons at the paramagnetic center.

Computational Procedures. A low energy solution conformation of **1-5i** was obtained from a SMILES string using Omega2 and Szybki (OpenEye Inc.) and docked into three different models of the hydrophobic pocket using AutoDock4.2[19] or AutoDock-Vina.[20] Poses selected for analysis with the NMR data included the 9 lowest energy poses from AutoDock-Vina and 3−5 poses from the lowest energy clusters representing at least 90 of the 100 poses obtained by AutoDock4.2. Structure calculations were carried out using XPLOR-NIH.[21] Paramagnetic relaxation enhancement restraints (PRE) Γ_2 obtained for the ligand were converted into distance bounds using eqs 3 and 4 and handled in one of two ways: (a) a single flexible MTSL was used to define the spin label position, or (b) an ensemble of 3 or 10 randomized MTSL side chains was created to represent the flexible paramagnetic group. In the latter case, the distance in eq 4 is replaced by the ensemble average $\langle r^{-6}\rangle$ and the correlation time by an apparent correlation time, $\tau_c{}^{app}$. This approach corresponds to the initial minimization step outlined by Iwahara *et al.*[22] Annealing consisted of a simple minimization step for the ligand−receptor−peptide complex, holding ligand and main chain atoms fixed, followed by a 3000-step minimization to determine ligand position using the distance bounds as NOE constraints with a force constant scaling factor of 50. A soft-square well NOE energy function was used with a well width equal to the distance range corresponding to the observed PRE ± observed error. Calculations were repeated for molecular correlation times in the range 2−20 ns.

RESULTS AND DISCUSSION

Computational Docking Studies. Figure 2 shows examples of **1-5i** docked into the PDB structures 3P7K, 2R5D, and 2KP8 using AutoDock-Vina. 3P7K is a crystal structure of 45 residue NHR;[23] 2R5D is a crystal structure of the IQN17:PIE7 D-peptide complex,[24] and 2KP8 is an NMR-based structure of a truncated 6-helix bundle with a small molecule.[12] Most docked poses conformed to the experimental requirement for a free carboxylate group on the ligand, since putative hydrogen bonds or salt bridges could be formed with either Arg579′ (group 1 poses) or Lys574 (group 2 poses). Within each group, different poses were classified as A or B for group 1 and C, D, or E for group 2. Poses A and B and poses D and E have the orientation of the indole ring flipped relative to one another. Pose C is a group of poses in which the carboxylate is buried more deeply in the

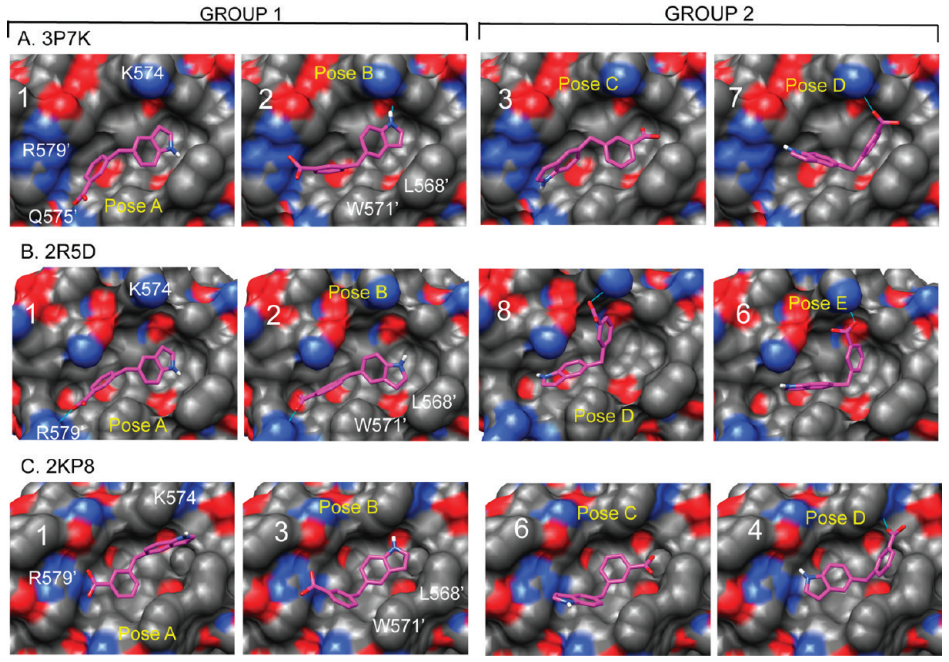


Figure 2. Principle poses observed by Vina docking of **1-5i** into the hydrophobic pocket of receptor structures (A) 3P7K, (B) 2R5D, and (C) 2KP8. The carboxylate group was unprotonated. The poses, categorized as A−E, are divided into two groups depending on the orientation of the ligand. In group 1, the ligand carboxylate points toward R579′ and in group 2 toward K574. The rank of the poses depicted is indicated in each panel, and the binding energies are given in Table 1. Calculated hydrogen bonds are shown in blue lines and include the following: in 3P7K, a hydrogen bond to Q575′-γNH_2 in rank 1, Val570-O in rank 2, and K574-εNH in rank 7; in 2R5D, hydrogen bonds to R579′-γNH_2 in ranks 1 and 2 and K574-εNH in ranks 6 and 8; in 2KP8, a hydrogen bond to Lys574′ in rank 4. Poses not shown are similar to those shown or have a ligand that is not in the main part of the pocket. Note that pose C exists for 2R5D (rank 3) but is not shown in this figure.

Table 1. Calculated Binding Data for Matching Poses of 1-5i and 1-6i in AutoDock-Vina Calculations on Three Receptor Structures

		1-5i, obsd K_I[a] = 18.1 ± 3.8					**1-6i**, obsd K_I[a] = 1.7 ± 0.2				
		group 1[b] poses		group 2[b] poses			group 1[b] poses		group 2[b] poses		
		A	B	C	D	E	A	B	C	D	E
3P7K	calcd ΔG[c]	−6.7	−6.7	−6.5	−6.1	−	−7.1	−6.9	−6.8	−6.4	−6.1
	calcd K_I[a]	**12**	**12**	**17**	**34**		**6**	**9**	**10**	**20**	34
	rank	1	2	3	7		1	2	3	7	9
2R5D	calcd ΔG[c]	−7.6	−7.1	−7.0	−6.4	−6.6	−7.6	−7.1	−7.3	−7.2	−
	calcd K_I[a]	**3**	**6**	**7**	**20**	**14**	**3**	**6**	**4**	**5**	
	rank	1	2	3	8	6	1	4	2	3	
2KP8	calcd ΔG[c]	−6.6	−6.5	−6.2	−6.4	−	−6.5	−6.2	−6.2	−6.5	−
	calcd K_I[a]	**14**	**17**	**28**	**20**		**17**	**28**	**28**	**17**	
	rank	1	3	6	4		1	5	4	2	

[a] In μM. [b] Groups and poses as defined in Figure 2 and Supplementary Figure 1S. Equivalent poses of **1-5i** and **1-6i** have the indole NH pointing in the opposite direction. [c] In kcal/mol.

pocket and cannot form a hydrogen bond to the lysine side chain. These poses are representative of the 9 poses calculated altogether for each structure. All 9 poses were subjected to NMR analysis. In 2R5D, the lysine side chain was rotated (a 143° rotation about the Cδ−Cε bond) so that the ε-NH_2 was accessible to a pocket binding ligand. In all of the poses, an approximately perpendicular orientation of the two ring systems was maintained in the ligand, in keeping with the lowest energy solution structure, where the angle between the planes is 85.2°. Figure 2 shows the variation in distribution of charges around the pocket in different receptor structures, due to the flexibility of the surrounding residues. The result is subtle differences in the orientation of the ligand between different structures.

Table 1 lists the calculated energies for these structures, together with calculated energies for the isomer **1-6i** in matched poses for each receptor. The poses of **1-6i** are shown in Figure 1 in Supporting Information. In most cases, a match was found for the docked poses of **1-5i** and **1-6i** in each receptor, with the indole NH and C3 positions switched. One exception was that pose E was not found for **1-5i** docked into 3P7K, although it was

dx.doi.org/10.1021/cb100368d |*ACS Chem. Biol.* 2011, 6, 267–274

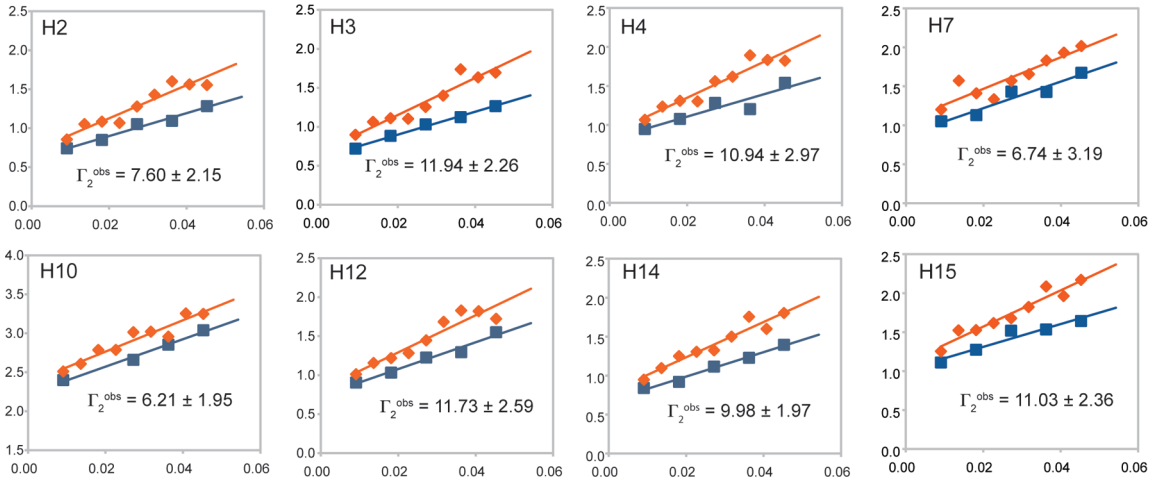


Figure 3. Diamagnetic relaxation (■) plotted against f_b and paramagnetic relaxation (◆) plotted against $f_b \cdot f_b'$ for 8 protons of **1-5i**. Best-fit straight lines are shown through the data.

found for **1-6i** in 3P7K. Calculated affinity was consistently lower for **1-5i** compared to **1-6i** with receptor structure 3P7K and for group 2 poses in 2R5D, by a factor of ∼2−4. Calculated binding affinities appeared lower overall in 2KP8, possibly due to the widened groove in that structure and therefore reduced hydrophobic contacts with a small ligand.

The selection and rank ordering of poses calculated using the program AutoDock4.2 was different. In 3P7K, 6 of the 9 Vina-docked poses of **1-5i**, including the top ranked poses, fell into group 1 as defined above, whereas AutoDock4.2 predicted the lowest energy cluster, comprising 42% of all structures, to be in group 2, poses D or E. The default clustering parameters did not distinguish between these two poses, although the lowest energy poses in the cluster were in pose D. Pose E was not detected by Vina in multiple repeat runs. A similar result occurred for **1-6i**, with the majority, including the top two Vina-docked structures, belonging to group 1, whereas with AutoDock4.2, 95% of the structures were in group 2, pose E. Differences also occurred with the other two receptors. The principal observed poses calculated for **1-5i** by AutoDock4.2 were poses B and D/E in 3P7K, poses B and D in 2R5D, and poses D and E in 2KP8. In 2KP8, a wider groove permitted the indole rings to lie flat against the pocket in these structures, while it is angled relative to the base of the pocket in the other two receptors (*e.g.*, Figure 2, poses D and E). Autodock4.2 results are shown in Figure 2 in Supporting Information.

The disparity between different docking protocols and different structures, especially in the case of a small ligand targeted to a flexible protein−protein interaction surface, makes it difficult to rely on computational calculations alone to predict binding orientation. We therefore turned to paramagnetic second site screening as a way to distinguish between the poses. This method relies on the transfer of distance-dependent paramagnetic effects from bound ligand structure to free ligand present in excess in solution. It was applicable to **1-5i**, which is in the fast exchange regime, and results can only be inferred for **1-6i**.

Addition of NMR Restraints. The 1D NMR spectrum of **1-5i**, shown in Figure 1, has 8 well-resolved protons, for which transverse relaxation rates R_2 were obtained in the presence of increasing concentrations of receptor complex Fe(env5.0)$_3$ and either spin-labeled or unlabeled peptide MTSL-C29e5.0. C29e5.0 binds to the NHR coiled coil with its N-terminus adjacent to the hydrophobic pocket.[11,25,26] In the spin-labeled system, the paramagnetic relaxation enhancement (PRE) due to the unpaired electron provides information on the distance from the spin label to the observed proton (eq 3). It is obtained as the difference between the relaxation in the paramagnetic and diamagnetic systems. Since the system is trimeric, up to three MTSL groups could in principle contribute to relaxation of ligand protons; however, the spin label of peptides binding on the second and third face of the coiled coil are twice as far away and can be neglected.

By making multiple measurements, the PRE could be obtained more precisely than in a single point study.[11] Nine data points were accumulated for the paramagnetic system and five for the diamagnetic reference. Figure 3 shows the data obtained for each resonance. The slope of the plot of R_{2obs}^{dia} against f_b was similar for all resonances, with a value of 16.01 ± 1.44 s^{-1}; there is no spatial variation in R_{2b}^{dia} across the ligand. Subtracting the diamagnetic best-fit line from the paramagnetic data gave a straight line with slope Γ_2, when plotted against $f_b \cdot f_b'$ (eq 2). Experimental errors were obtained from the sum of the errors in the slopes. There is a relatively low dispersion of PRE across the ligand, doubtless because of its small size as well as its orientation with respect to the spin label.

PRE constraints have received limited application in quantitative structure calculations because of the low accuracy with which relaxation times can be measured and the uncertainty in spin-label position due to flexibility of the attached paramagnetic group. Applications to bound ligand structure have been limited to qualitative assessment of binding, in the second site screening approach.[27,28] In our case, making multiple measurements of R_2 at different concentrations of peptide−receptor complex enabled us to average the error associated with measurement of an individual R_2 and to obtain a relatively narrowly defined constraint for each proton. Nevertheless, the low number of constraints for a small molecule ligand limited the structure refinement to a simple minimization protocol, in order not to drive the annealing by other nonexperimental constraints. We therefore used the PRE to examine different ligand poses obtained by computational docking and allowed the ligands to relax into a local minimum with respect to the PRE constraints. Both single and multiple MTSL conformer representations were used. The multiple conformer representation containing a minimum of 3 randomized spin label

Table 2. Fit of PRE Data to Calculated Poses of 1-5i in Receptor 3P7K[a]

pose definition[b]							
rank	group	pose	rmsd to Γ_2	NOE violations >1.0 Å	$\langle R^{-6}\rangle^{-1/6}$ rmsd (Å)	τ_c (ns)[c]	affinity (kcal/mol)
V-1	1	A	5.2	2	2.16	10	−6.7
V-2	1	B	5.1	3	2.28	4	−6.7
V-3	2	C	3.5	3	1.16	10	−6.5
V-4	2	C	5.4	3	2.51	10	−6.4
V-5	1	B	5.2	4	1.94	2	−6.3
V-6	out of pocket		5.9	3	2.22	6	−6.1
V-7	2	D	3.7	0	0.39	18	−6.1
V-8	out of pocket		3.7	1	1.01	20	−6.1
V-9	1	A	11.1	4	3.12	10	−6.0
AD-1	2	D	4.7	3	1.80	6	−7.0
AD-2	1	B	5.1	3	2.34	6	−6.7

[a] Three conformer representation of MTSL. [b] Poses defined in Figure 2, Table 1 or Supplementary Figure 2S (V = Vina, AD = AutoDock). [c] Correlation time at which minimum number of distance violations were observed.

side chains has been used successfully to represent conformational flexibility.[22,29−31]

Evaluation of Calculated Poses. Existing poses determined by AutoDock-Vina or AutoDock4.2 were first evaluated to see if any of them directly fit the data. This method would help to identify a receptor applicable to small molecule docking and to some degree to evaluate the computational algorithm, albeit specific to the current system. It also depends on the placement of the spin label, which varied slightly in the three receptors. A C-peptide was docked into 3P7K and 2R5D using homology modeling from 1IF3.[26] 2KP8 already contained the C-peptide, which required elongation by two residues in an α-helical conformation.[12] The resulting position of the MTSL was slightly different in the three structures, an inherent uncertainty in the method, but in keeping with multiple possible positions of the spin label due to the flexible anchor. A three-conformer model for the MTSL was applied, first randomizing the position of the MTSL side chain and then allowing the side chains to move during minimization while holding the ligand and protein backbones fixed. The calculated PRE for each ligand proton was determined from the ensemble-averaged distance for multiple correlation times (τ_c) varying from 2 to 20 ns.

The only pose in agreement with experimental data was the Vina-docked pose D in group 2 in the receptor structure 3P7K (Figure 2). No PRE-based distance restraints were violated by greater than 1.0 Å over a τ_c range of 12−20 ns. It was possible to fit the data with multiple correlation times, since the flexible MTSL group could move away from the ligand binding site as τ_c was increased. No Vina-docked poses in 2R5D or 2KP8 matched the data directly, nor did any of the AutoDock4.2 poses in all three structures. Table 2 shows the results for **1-5i** in 3P7K, listing the correlation time corresponding to the minimum observed PRE-based distance restraint violation. The data are consistent with an association between the carboxylate group of **1-5i** and Lys574-εNH_2. They also suggest that 3P7K is the best representative model for small molecule docking among the three structures tested, although it is not a perfect model.

The data fitting was sensitive to the starting orientation of **1-5i** and to receptor structure. Convergence was not obtained for ligands in quite similar orientations, *e.g.*, pose D in 3P7K (0 distance violations >1 Å) *versus* 2R5D (3 violations) or 2KP8 (5 violations). The minimization protocol samples a restricted space, so that a global minimum is not assured. It is rather a probe of the accuracy of the docking algorithm and permits evaluation of the receptor model.

The multiconformer representation of MTSL did not improve the analysis. A similar agreement between observed and calculated data was obtained using a single MTSL side chain to represent the spin label position and using a 10-conformer model (Tables 1 and 2 in Supporting Information).

Minimization of Ligand Position. The protocol was subsequently modified to see if the agreement with experimental constraints could be improved by adjusting ligand position. After a short restraints-free minimization during which ligand and peptide backbone atoms were held fixed, NOE restraints corresponding to the 8 PRE values were added, and **1-5i**, MTSL, and receptor side chains were allowed to move during 3000 steps of minimization. The calculation was repeated for molecular correlation times τ_c ranging from 2 to 16 ns. As before, τ_c acted like a scaling factor with distances becoming longer with increasing τ_c, accommodated by movement of the MTSL side chain. Final structures were therefore selected on the basis of the best agreement with experimental data, since in this case variations in τ_c can be considered to be a reflection of the ambiguity in spin label position. Ligands were examined by eye to ensure that they were not displaced from the pocket. Structures in which the indole or phenyl ring extruded from the receptor were excluded.

This analysis is considered to provide a more accurate reflection of ligand position, since fixing the ligand in the original docked position resulted in distortion or tight clustering of MTSL conformers in some structures. Furthermore, a significant improvement in the agreement between observed and calculated data was obtained (Table 3). A multiple conformer model of the MTSL side chain could not be used with this minimization protocol; in the general case, the presence of so many degrees of freedom and so few constraints yielded a flat energy surface with poor discrimination between ligand poses.

Table 3 lists the results obtained starting with AutoDock-Vina poses shown in part in Figure 2 and listed in Table 1 and Autodock4.2 clusters shown in Figure 2 in Supporting Information. Structures that do not appear in Table 3 yielded a ligand that was clearly forced out of the pocket, with one of the ring systems making little hydrophobic contact with surrounding pocket residues. This included structures resulting from starting pose

Table 3. Results of Minimization Using 8 PRE Restraints on AutoDock4.2- or AutoDock-Vina-Derived Poses[a]

starting pose definition[b]							
receptor	rank[c]	group	pose	rmsd to Γ_2	NOE violations >0.5 Å	$\langle R^{-6}\rangle^{-1/6}$ rmsd (Å)	τ_c (ns)[d]
3P7K	V-1	1	A	5.07	2	1.84	8
2R5D	V-1	1	A	4.77	2	1.73	12
3P7K	V-2	1	B	4.97	2	1.99	8
2R5D	V-2	1	B	3.97	1	0.83	8
3P7K	AD-2	1	B	5.00	2	2.02	8
2R5D	AD-1	1	B	5.23	3	2.55	8
3P7K	V-7	2	D	2.53	0	0.094	12
2R5D	V-8	2	D	2.62	0	0.112	12
3P7K	AD-1	2	D	2.44	0	0.076	12
2R5D	AD-2	2	D	2.47	0	0.067	12
2R5D	V-6	2	E	4.60	3	2.10	12

[a] Single conformer representation of MTSL. [b] Poses defined in Figure 2, Table 1, and Supporting Information Figure 1 (see text). [c] Poses ranked by Vina (V) or AutoDock4.2 (AD), for Autodock poses, the lowest ranked structure in the cluster was used. [d] Correlation time at which minimum experimental violation was observed.

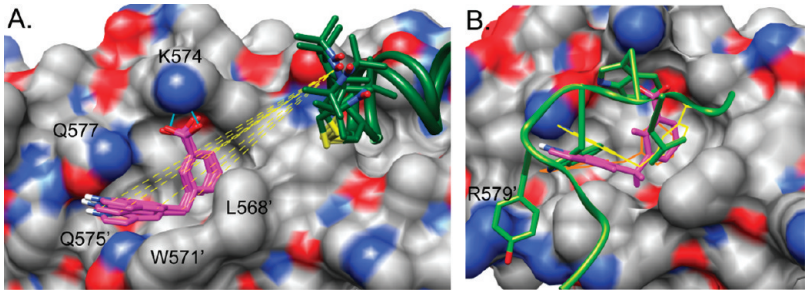


Figure 4. (A) Lowest energy minimized structure of **1-5i** (magenta) bound in the 3P7K hydrophobic pocket resulting from PRE-constrained minimization of the four group 2D poses listed in Table 3, using τ_c = 12 ns. 3P7K and 2R5D structures were superimposed by matching the backbone atoms of the N-peptides. C-Peptide helices and MTSL groups for the four 2R5D and 3P7K structures are shown in green, with distances defining the ligand position indicated in dashed yellow lines. Putative hydrogen bonds are indicated in cyan. (B) Superposition of the final pose of **1-5i** with hydrophobic residues dTyr7, dTrp10, dTrp12, dLeu13 of the D-peptide PIE7. Also shown as wire structures are the starting Vina (yellow) and AutodDock4.2 (orange) poses in the 3P7K structure.

2C and structures using receptor 2KP8. The results indicate that only structures derived from starting pose 2D in receptors 3P7K and 2R5D were able to fit the data, and agreement could not be obtained for structures derived from 1A, 1B, or 2E. This further confirms ligand orientation with the carboxylate oxygen atoms making hydrogen bond contact to Lys574-εNH_2. In addition, the data clearly discern the orientation of the indole ring, distinguishing between the indole NH pointing toward (pose 2D) or away (pose 2E) from the pocket. Since minimization alone is not able to surmount the energy barrier required to flip the indole ring, poses starting from orientation 2E do not converge to a final structure with low experimental rmsd's. It is interesting that despite the limitations of a flexible MTSL group, a clear and fairly highly resolved final structure of the ligand was obtained. The consensus structure is shown in Figure 4 as well as a superposition of the small molecule onto the crystal structure of the D-peptide PIE7 in 2R5D.[24] Figure 5 shows the data fitting for the resulting structures, with observed and calculated PRE and distances for a sample pose. An excellent agreement with experimental data was obtained, with an rmsd in Γ_2 approaching the observed errors. The four structures obtained with two different receptors converged well, with an all-atom positional rmsd in the range of 0.69−0.87 Å between structures. The indole ring forms hydrophobic interactions with Trp571′ on the receptor and possibly a polar interaction with Gln577. The aromatic groups in **1-5i** appear to overlap with hydrophobic dTrp10 and dLeu13 residues on PIE7 in 2R5D, occupying the same region of the pocket. Not shown in Figure 4B are polar interactions of PIE7 in the pocket, but these do not include hydrogen bonding to the pocket Lys574, as is seen here for **1-5i**. The similarity in pocket occupancy may explain in part the relatively high affinity of **1-5i** and **1-6i** for the hydrophobic pocket, despite their small size (MW 251).

Analysis of PRE-Guided Ligand Docking. The results demonstrated that NMR-derived PRE constraints could validate and improve upon computationally docked structures of a small ligand, **1-5i**, bound to the gp41 hydrophobic pocket. The PREs were able to distinguish between several docked poses of **1-5i** in the hydrophobic pocket that were close in energy, an especially helpful feature for binding pockets that display such structural plasticity. The simulations suggested one of two approximately iso-energetic orientations of the ligand, with the carboxylate group facing either toward Lys574 or toward Arg579′, which are at opposite ends of the pocket. The accuracy with which we were able to obtain PRE values enabled us to discriminate between these orientations, in favor of the Lys574-carboxylate interaction, and furthermore to obtain definitive information about the orientation of the indole ring.

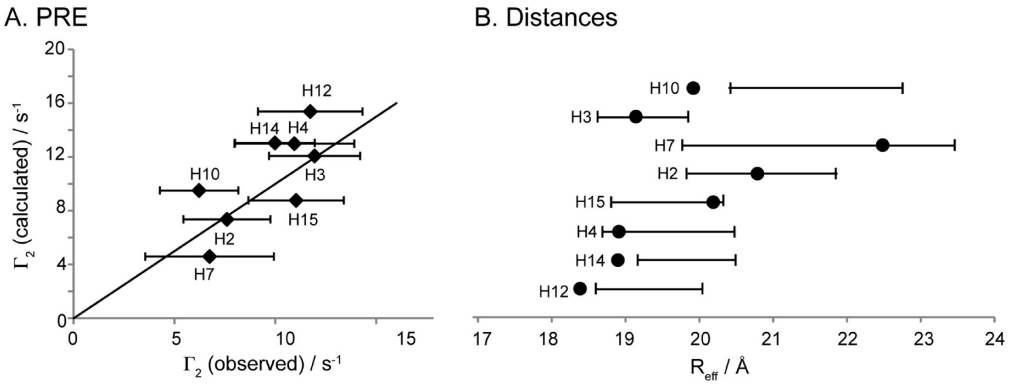


Figure 5. (A) Observed versus calculated PRE and (B) corresponding distances for **1-5i** protons after minimization using τ_c = 12 ns. The starting structure was receptor 2R5D, single MTSL conformer, AutoDock4.2 lowest energy docked ligand from cluster 2, pose 2D. In A, horizontal error bars indicate the experimental errors, and a diagonal line is drawn through $y = x$. In B, the horizontal bars indicate the experimental distance range, and the symbols show the calculated distances for each specified proton.

We consistently found pose D (Figures 2, 4) to match the data, including in evaluating existing computational docked structures as well as allowing them to relax into lower experimental energy structures. This is in agreement with previous peptide studies that suggested that a salt bridge between Lys574 and Asp632 on the gp41 C-peptide plays an important role in the energetics of binding.[32] We identified 3P7K as the only receptor for which a computational pose was found that agreed with the experimental data, and then only for the docking software AutoDock-Vina. However, it was not the lowest energy structure selected by Vina, instead being ranked seventh or eighth in different runs. AutoDock4.2 selected a pose in category 2D for **1-5i** as the lowest energy structure among 100 docked poses in 3P7K and as the second ranked cluster in 2R5D, although some adjustment was required for these poses to fit the NMR data (Figure 4B). This is likely not a reflection of the adequacy of the computations but rather includes the uncertainty in receptor structure. The protein−protein interface is highly fluid, as evidenced by the multiple different entries in the PDB, including the three receptor structures tested here. The agreement with 3P7K may reflect the fact that **1-5i** is a very small ligand and 3P7K is an apo-structure, whereas the other structures used contained either a bound peptide (2R5D) or a rather large organic molecule (2KP8), which likely mold the pocket in different ways. 2KP8 had a clearly increased groove width, and no final structures could be obtained with this receptor that matched the experimental data.

Refinement of the docked ligand indicated a specific indole-ring orientation, which may provide a clue as to the difference in affinities between **1-5i** and **1-6i**. Examination of Figure 4A indicates that the indole ring is not buried deeply into the pocket. It forms hydrophobic interactions with Trp571′. The indole NH points toward the hydrophobic receptor, possibly forming a polar interaction with Gln577. It may preclude deeper insertion into the pocket because of its polarity. This might explain the higher affinity of **1-6i**, which could be buried further into the pocket if the NH group is replaced with a CH; at the same time, an NH group on the outer edge of the ring might hydrogen bond with Gln577 or Gln575′. This suggestion assumes that the two isomers bind in a similar orientation, given their small size and similar structures, although it is not impossible that they could adopt entirely different orientations. It is also possible that a ligand pose may exist that was not detected, because the analysis was restricted to computationally derived poses. With such a small ligand, however, it is likely that sample space has been covered.

Attempts to use a multiconformer analysis strategy for the flexible MTSL group were not very successful, since they did not result in improved agreement with experimental restraints or in better definition of the ligand. In fact, the multiconformer model appeared to be an over-representation of the system in this case where there are so few restraints. On the other hand, the single MTSL conformer model enabled us to arrive at a consistent minimized pose with low experimental energy starting from different structures. Previous studies have shown that the ensemble representation of the spin-label probe may reduce structural accuracy due to an unrealistic sampling of probe conformational space.[33] The parameter that seems to be most critical is τ_c, the molecular correlation time, which essentially acts like a nonlinear scaling factor for the distances. The final structure varied slightly (mainly in the exact position of the indole ring) depending on the choice of τ_c. The high precision with which we could determine the PRE enabled us to clearly select τ_c from the minimum in the rmsd between calculated *versus* experimental PRE and distances. This enabled higher structural definition. Although the value that was obtained, 12 ns, agrees well with the expected correlation time for this complex, it could be affected by inaccuracies in the system including the exact placement of the C-peptide and errors in the experimental binding constants for ligand and C-peptide.

Despite the limitations of this method, namely, limited precision due to flexibility of the MTSL side chain and restricted sampling of chemical space, the analysis resulted in an improved understanding of the docked ligand interactions in the hydrophobic pocket. The orientation is in agreement with the finding that adding hydrophobic groups onto the indole NH of **1-6i** improves potency,[9] as these are likely to interact with Trp571 and Gln575′ in the channel occupied by dTyr7 of PIE7. Knowledge of the bound structure and comparison to PIE7 suggest a direction for optimization of the ligand as an HIV fusion inhibitor. Extension of the method to ligands in slow exchange should permit rapid evaluation of more tightly bound inhibitors.

■ ASSOCIATED CONTENT

Supporting Information. This material is available free of charge *via* the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*Phone: (707) 638-5463. Fax: (707) 638-5255. Email: miriam.gochin@tu.edu.

ACKNOWLEDGMENT

This work was supported by NIH grant NS066469 to M.G. Molecular graphics images were prepared using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR001081). The authors thank C. Huang for conversion software enabling AutoDock output files to be read using ViewDock. The authors also gratefully acknowledge use of the UC Berkeley Biomolecular NMR facility. The authors thank E. Springman at Locus Pharmaceuticals (currently Ansaris) for providing the coordinates of 3P7K prior to publication.

REFERENCES

(1) Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998) Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. *Proc. Natl. Acad. Sci. U.S.A. 95*, 15613–15617.

(2) Jiang, S., Lu, H., Liu, S., Zhao, Q., He, Y., and Debnath, A. K. (2004) N-Substituted pyrrole derivatives as novel human immunodeficiency virus type 1 entry inhibitors. *Antimicrob. Agents Chemother. 48*, 4349–4359.

(3) Liu, S., Wu, S., and Jiang, S. (2007) HIV entry inhibitors targeting gp41: from polypeptides to small-molecule compounds. *Curr. Pharm. Des. 13*, 143–162.

(4) Frey, G., Rits-Volloch, S., Zhang, X. Q., Schooley, R. T., Chen, B., and Harrison, S. C. (2006) Small molecules that bind the inner core of gp41 and inhibit HIV envelope-mediated fusion. *Proc. Natl. Acad. Sci. U.S.A. 103*, 13938–13943.

(5) Cai, L., and Gochin, M. (2007) A novel fluorescence intensity screening assay identifies new low molecular weight inhibitors of the gp41 coiled coil domain of HIV-1. *Antimicrob. Agents Chemother. 51*, 2388–2395.

(6) Wang, Y., Lu, H., Zhu, Q., Jiang, S., and Liao, Y. (2010) Structure-based design, synthesis and biological evaluation of new N-carboxyphenylpyrrole derivatives as HIV fusion inhibitors targeting gp41. *Bioorg. Med. Chem. Lett. 20*, 189–192.

(7) Katritzky, A. R., Tala, S. R., Lu, H., Vakulenko, A. V., Chen, Q. Y., Sivapackiam, J., Pandya, K., Jiang, S., and Debnath, A. K. (2009) Design, synthesis, and structure-activity relationship of a novel series of 2-aryl 5-(4-oxo-3-phenethyl-2-thioxothiazolidinylidenemethyl)furans as HIV-1 entry inhibitors. *J. Med. Chem. 52*, 7631–7639.

(8) Teixeira, C., Barbault, F., Rebehmed, J., Liu, K., Xie, L., Lu, H., Jiang, S., Fan, B., and Maurel, F. (2008) Molecular modeling studies of N-substituted pyrrole derivatives—potential HIV-1 gp41 inhibitors. *Bioorg. Med. Chem. 16*, 3039–3048.

(9) Zhou, G., Wu, D., Hermel, E., Balogh, E., and Gochin, M. (2010) Design, synthesis, and evaluation of indole compounds as novel inhibitors targeting Gp41. *Bioorg. Med. Chem. Lett. 20*, 1500–1503.

(10) Berman, H., Henrick, K., and Nakamura, H. (2003) Announcing the worldwide Protein Data Bank (www.pdb.orgwww.pdb.org). *Nat. Struct. Biol.10*, 980.

(11) Balogh, E., Wu, D., Zhou, G., and Gochin, M. (2009) NMR second site screening for structure determination of ligands bound in the hydrophobic pocket of HIV-1 gp41. *J. Am. Chem. Soc. 131*, 2821–2823.

(12) Stewart, K. D., Huth, J. R., Ng, T. I., McDaniel, K., Hutchinson, R. N., Stoll, V. S., Mendoza, R. R., Matayoshi, E. D., Carrick, R., Mo, H., Severin, J., Walter, K., Richardson, P. L., Barrett, L. W., Meadows, R., Anderson, S., Kohlbrenner, W., Maring, C., Kempf, D. J., Molla, A., and Olejniczak, E. T. (2010) Non-peptide entry inhibitors of HIV-1 that target the gp41 coiled coil pocket. *Bioorg. Med. Chem. Lett. 20*, 612–617.

(13) Gochin, M., Guy, R. K., and Case, M. A. (2003) A metallopeptide assembly of the HIV-1 gp41 coiled coil is an ideal receptor in fluorescence detection of ligand binding. *Angew. Chem., Int. Ed. 42*, 5325–5328.

(14) Carr, H. Y., and Purcell, E. M. (1954) Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev. 94*, 630–63894, 630−638.

(15) Hoult, D. (1976) Solvent peak saturation with single phase and quadrature Fourier transformation. *J. Magn. Reson. 21*, 337–347.

(16) Meiboom, S., and Gill, D. (1958) Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum. 29*, 688–691.

(17) Mulder, F. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Measurement of slow (micros-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc. 123*, 967–975.

(18) Solomon, I., and Bloembergen, N. (1956) Nuclear magnetic interactions in the HF Molecule. *J. Chem. Phys. 25*, 261–266.

(19) Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem. 30*, 2785–2791.

(20) Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem. 31*, 455–461.

(21) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson. 160*, 65–73.

(22) Iwahara, J., Schwieters, C. D., and Clore, G. M. (2004) Ensemble approach for NMR structure refinement against (1)H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J. Am. Chem. Soc. 126*, 5879–5896.

(23) Springman, E. B., Celtaxis Inc, personal communication.

(24) Welch, B. D., VanDemark, A. P., Heroux, A., Hill, C. P., and Kay, M. S. (2007) Potent D-peptide inhibitors of HIV-1 entry. *Proc. Natl. Acad. Sci. U.S.A. 104*, 16828–16833.

(25) Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., and Clore, G. M. (1998) Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J. 17*, 4572–4584.

(26) Caffrey, M. (2001) Model for the structure of the HIV-1 gp41 ectodomain: insight into the intermolecular interactions of the gp41 loop. *Biochim. Biophys. Acta 1536*, 116–122.

(27) Jahnke, W., Perez, L. B., Paris, C. G., Strauss, A., Fendrich, G., and Nalin, C. M. (2000) Second-site NMR screening with a spin-labeled first ligand. *J. Am. Chem. Soc. 122*, 7394–7395.

(28) Vazquez, J., De, S. K., Chen, L. H., Riel-Mehan, M., Emdadi, A., Cellitti, J., Stebbins, J. L., Rega, M. F., and Pellecchia, M. (2008) Development of paramagnetic probes for molecular recognition studies in protein kinases. *J. Med. Chem. 51*, 3460–3465.

(29) Yu, D., Volkov, A. N., and Tang, C. (2009) Characterizing dynamic protein-protein interactions using differentially scaled paramagnetic relaxation enhancement. *J. Am. Chem. Soc. 131*, 17291–17297.

(30) Liu, Y., Kahn, R. A., and Prestegard, J. H. (2010) Dynamic structure of membrane-anchored Arf*GTP. *Nat. Struct. Mol. Biol. 17*, 876–881.

(31) Volkov, A. N., Worrall, J. A., Holtzmann, E., and Ubbink, M. (2006) Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. *Proc. Natl. Acad. Sci. U.S.A. 103*, 18945–18950.

(32) He, Y., Liu, S., Jing, W., Lu, H., Cai, D., Chin, D. J., Debnath, A. K., Kirchhoff, F., and Jiang, S. (2007) Conserved residue Lys574 in the cavity of HIV-1 Gp41 coiled-coil domain is critical for six-helix bundle stability and virus entry. *J. Biol. Chem. 282*, 25631–25639.

(33) Bermejo, G. A., Strub, M. P., Ho, C., and Tjandra, N. (2009) Determination of the solution-bound conformation of an amino acid binding protein by NMR paramagnetic relaxation enhancement: use of a single flexible paramagnetic probe with improved estimation of its sampling space. *J. Am. Chem. Soc. 131*, 9532–9537.