JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Indirect Readout in Protein-Peptide Recognition: A Different Story from Classical Biomolecular Recognition

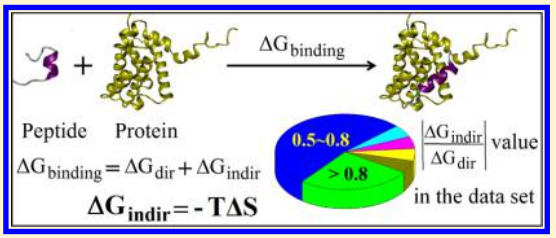
Hua Yu,[†] Peng Zhou,[‡] Maolin Deng,*[,§] and Zhicai Shang*[,†]

[†]Department of Chemistry and [§]Department of Mechanics, Zhejiang University, Hangzhou, Zhejiang 310027, China

[‡]Center of Bioinformatics (COBI), School of Life Science and Technology, University of Electronic Science and Technology of China (UESTC), Chengdu, Sichuan 610054, China

Supporting Information

ABSTRACT: Protein-peptide interactions are prevalent and play essential roles in many living activities. Peptides recognize their protein partners by direct nonbonded interactions and indirect adjustment of conformations. Although processes of protein-peptide recognition have been comprehensively studied in both sequences and structures recently, flexibility of peptides and the configuration entropy penalty in recognition did not get enough attention. In this study, 20 protein-peptide complexes and their corresponding unbound peptides were investigated by molecular dynamics simulations. Energy analysis revealed that configurational entropy penalty introduced by restriction of the degrees of freedom of peptides in indirect readout process of protein-peptide recognition is significant. Configurational entropy penalty has become the main content of the indirect readout energy in protein-peptide recognition instead of deformation energy which is the main source of the indirect readout energy in classical biomolecular recognition phenomena, such as protein−DNA binding. These results provide us a better understanding of protein-peptide recognition and give us some implications in peptide ligand design.



1. INTRODUCTION

Protein-peptide interactions play important roles in many cellular processes, such as signal transduction, apoptotic, immune system, and other pathways.[1−3] Hence, they are also attractive targets for drug design. In recent years, many studies have been done to investigate the sequence and structural basis of the protein-peptide binding process.[4−8] As knowing more about protein-peptide interactions, many researchers have tried to design peptide ligands though rational designs that have shown potent inhibition in vitro or in vivo are still rare, which drives us to do further exploration.[9,10]

Traditionally, the classical biomolecular recognition can be decomposed into two processes in terms of thermodynamics. First, the two partners change their conformations to the one adopted in the complex, which is called indirect readout. Then, the two right conformations bind through direct nonbonded interactions, which is called direct readout. Both processes contribute to the total binding free energy change. A great deal of research has been done by Akinori Sarai's[11−16] group and other researchers[17] to evaluate the direct and indirect readout energies and their relative contributions to the sequence specificity of the protein−DNA and drug−DNA complexes. In their reports, indirect readout energy comes from the conformational energy difference which is called deformation energy between the bound and the unbound DNA conformations. However, situations may be different in protein-peptide indirect readout for the great flexibility of peptides. In 2007, Chang and co-workers reported that ligand configurational entropy loss upon protein−ligand binding is neglected or underestimated for a long time by analyzing the association of amprenavir with HIV protease.[18] In 2009, Killian researched the binding of Tsg101 ubiquitin E2 variant domain with an HIV-derived PTAP nonapeptide and noted that the configurational entropy changes on binding play a role as important as nonbonded interactions.[19] Thus, entropy penalty ($-T\Delta S$) introduced by the restriction of degrees of freedom of the bound peptide may be large and cannot be ignored in protein-peptide recognition. However, almost all present studies on entropy cost focus on one specific complex as mentioned above. We lack a systematic investigation on many different protein-peptide complexes. Hence, in this article we addressed this problem via molecular dynamics simulations on 20 protein-peptide complexes with various peptide structures, various binding modes, and various cellular activities selected from the Protein Data Bank[20] (PDB). Present analysis revealed that peptide configurational entropy penalty ($-T\Delta S$) is significant and becomes the main source of the indirect readout energy in protein-peptide recognition. This suggests to us to pay more attention to the flexibility of peptides whenever predicting binding affinity or designing peptide ligands and so on.

2. MATERIALS AND METHODS

2.1. The Data Set. All of the protein-peptide complexes[21−38] were obtained from the PDB. We selected manually the complexes that have different structures of peptides (Table

Received: January 13, 2014
Published: July 7, 2014

 © 2014 American Chemical Society dx.doi.org/10.1021/ci5000246 | J. Chem. Inf. Model. 2014, 54, 2022−2032

1), different binding modes, and various cellular activities with available dissociation constants (K_d) (Table 2). All peptides consist of natural amino acids only. The statistical numbers of various amino acids present in peptides of the data set were summarized in Figure S1. Proline is overrepresented because of the selection of proline-rich peptides. There were two kinds of objects in our simulations. One was the protein-peptide complex. The other was the unbound peptide obtained from the corresponding complex. Here, for complexes that have multiple molecular models, we uniformly chose the first one. In theory, no matter which model we use, the simulation result will be the same as various NMR models are very likely the actual ensemble of conformations in solution.[39] However, a certain degree of deviation among energy results of different models always exists because of the limited time scale and randomness of the dynamics simulations in practice (1H3H, 1H3H-model-3, 1H3H-model-18, 2KHH, 2KHH-model-12, and 2KHH-model-20 in Table S2). Though the energy variation did not influence the main conclusion of this paper significantly, it did have a little effect on the correlation analysis in Figure 4. A detailed analysis of this kind of influence was shown in Figure S2.

Table 1. Secondary Structure of the Bound Peptide[a]

Num.	PDB code	α-helix	poly-proline type II helix	3_{10}-helix	β-strand	turn	bend	coil
1	1BXL	56%	–	–	–	13%	–	31%
2	1SY9	84%	–	–	–	–	–	16%
3	2K2I	45%	–	–	–	–	15%	40%
4	2PHE	35%	–	12%	–	23%	–	30%
5	1WLP	20%	24%	–	–	–	16%	40%
6	1HAA	–	–	–	31%	15%	15%	39%
7	2ASQ	–	–	–	21%	14%	14%	51%
8	2K00	–	–	–	20%	–	20%	60%
9	1H3H	–	–	27%	–	–	–	73%
10	1JEG	–	22%	22%	–	–	6%	50%
11	1WA7	–	36%	–	–	23%	9%	32%
12	2KXQ	–	20%	–	–	15%	25%	40%
13	1L2Z	–	36%	–	–	–	18%	46%
14	1ZSG	–	–	–	–	36%	32%	32%
15	1KC4	–	–	–	–	11%	16%	73%
16	1JGN	–	–	–	–	27%	14%	59%
17	2KHH	–	–	–	–	–	11%	89%
18	2KFF	–	–	–	–	25%	–	75%
19	2KFG	–	–	–	–	17%	25%	58%
20	2KFH	–	–	–	–	–	50%	50%

[a]Complexes with the similar peptide structure characteristics are colored the same. Structure information comes from the PDB and the corresponding literature.[31,33−35,37]

2.2. Molecular Dynamics Simulations. All of the molecular dynamics (MD) simulations were performed using GROMACS (ver.4.5.4) software package[40] with the Amber03[41] force field and the TIP3P[42] water model. Hydrogen atoms present in the original coordinate files were ignored and then added according to the description in the Amber03 force field using the "pdb2gmx" tool available in GROMACS. Default protonation states were used for all residues. Using the "editconf" tool, each protein-peptide complex\unbound peptide was placed in the center of a dodecahedron box. The distance between the solute and the box was 12 Å. Then the box with the solute inside was immersed into TIP3P water

Table 2. Selected Complexes and the Main Energy Term

PDB code	peptide sequence (size)	protein	K_d[a] (μM)	ΔG_{dir} (kJ/mol)	ΔG_{indir} (kJ/mol)	$\Delta G_{binding}$ (kJ/mol)	ΔE_{deform} (kJ/mol)
1BXL	GQVGRQLAIIGDDINR (16)	Bcl-xL	0.34 ± 0.03	−409.0905	197.379	−211.7115	41.72
1SY9	GGFRRIARLVGVLREWAYR (19)	Calmodulin	0.004	−620.9627	734.22	113.2573	48.25
2K2I	RADLHHQHSVLHRALQAWVT (20)	Centrin 2	0.23	−254.4187	171.879	−82.5397	91.565
2PHE	YGALDMADFEFEQMFTDALGID EYGG (26)	Transcriptional coactivator positive cofactor 4	15 ± 6	−267.8501	233.289	−34.5611	98.32
1HAA	WRYYESSLEPYPD (13)	α-bungarotoxin	0.009 09	−285.4344	155.58	−129.8544	4.746
1JGN	VVKSNLNPNAKEFVPGVKYGNI (22)	poly(A)-binding protein	0.16	−247.2536	45.105	−202.1486	122.266 44
1KC4	IPGKRTESFYECCKEPYPD (19)	α-bungarotoxin	30	−142.1043	123.207	−18.8973	85.965
2ASQ	KVDVIDLTIESSSD (14)	SUMO-1	6.5 ± 0.75	−185.0797	134.907	−50.1727	90.482
2ASQ-10res[b]	KVDVIDLTIE (10)	SUMO-1	6.5 ± 0.75	−209.4561	31.287	−178.1691	75.8
2K00	GRSKESGWVENEIYY (15)	Talin F3 subdomain	24	−309.5603	190.47	−119.0903	66.12
2KHH	DSGFSFGSK (9)	Mex67 UBA domain	0.276 ± 0.08	−179.0414	104.841	−74.2004	34.874
1JEG	IPPPLPERTPESFIVVEE (18)	Csk SH3 domain	0.8	−305.4695	107.301	−198.1685	81.729
1H3H	APSIDRSTKPA (11)	Gads SH3 domain	0.24 ± 0.045	−201.1084	168.57	−32.5384	17.316
1WA7	WDPGMPTPPLPPRPANLGERQA (22)	Lyn SH3 domain	9.58 ± 0.95	−270.6838	232.428	−38.2558	46.578
1WLP	GPLGSKQPPSNPPPRPPAEARKK PS (25)	p47phox SH3 domain	0.64	−376.3432	277.146	−99.1972	84.447
2KXQ	GPLGSELESPPPPYSRYPMD (20)	Smurf2 WW23 domain	1.7 ± 0.4	−268.4203	162.909	−105.5113	86.183
1ZSG	DATPPPVIAPRPEHTKSVYTRS (22)	β-PIX SH3 domain	7.5 ± 0.3	−220.1056	165.387	−54.7186	145.509
1L2Z	SHRPPPPGHRV (11)	CD2BP2 adaptor GYF domain	190 ± 22	−197.5866	109.344	−88.2426	17.564
2KFH	FNYESTGPFTAK (12)	EHD1 C-terminal EH-domain	2400 ± 100	−140.7908	160.731	19.9402	6.467
2KFG	FNYESTDPFTAK (12)	EHD1 C-terminal EH-domain	1200 ± 100	−121.9551	48.483	−73.4721	39.799
2KFF	FNYESTNPFTAK (12)	EHD1 C-terminal EH-domain	245 ± 35	−172.8848	95.697	−77.1878	19.95

[a]The experimental dissociation constant of the protein-peptide complex obtained from the corresponding literature.[21,23,24,26−31,33−38,52−54] The experimental conditions including T, pH, buffer, concentration, and assay method were summarized in Table S1. [b]Residues 11−14 of the peptide in 2ASQ were cut out. Discussions in the context used 2ASQ if not specified.

dx.doi.org/10.1021/ci5000246 | J. Chem. Inf. Model. 2014, 54, 2022−2032

using the GROMACS utility program "genbox". In each system, proper numbers of Na^+ or Cl^- were added to keep neutralization using "genion". This step was followed by steepest descent energy minimization until the maximum force was smaller than 100 kJ mol^{-1} nm^{-1} on any atom. Afterward, each system was equilibrated in two stages. First, a 1 ns simulation was performed under a constant volume ensemble (NVT) with a harmonic position restraint applied on heavy atoms of the solute. Second, a 1 ns simulation was performed under a constant pressure ensemble (NPT) without any restraint. All production simulations were conducted for 60 ns under NPT ensemble. For all the simulations, the default linear constraint solver (LINCS)[43] algorithm was used to constrain all bonds containing hydrogen atoms. Periodic boundary conditions were applied in all three directions. Berendsen's[44] coupling algorithm with constants of 0.1 and 1.0 ps was used to maintain temperature (300 K) and pressure (1 bar), respectively. Solute, solvent, and the counterions were coupled separately for temperature coupling. The particle-mesh-Ewald (PME)[45,46] method was used to calculate electrostatic interactions. van der Waals interactions were treated using a cutoff of 14 Å. Time step was set to 2 fs. Every 1 ps snapshot was collected for the whole simulation of 60 ns.

2.3. Calculation of the Direct and Indirect Readout Energetic Components Involved in Protein-Peptide Recognition. The same Amber03 force field and the trajectory of the last 20 ns were used to calculate energies for all the simulations. Any necessary analysis tool was available in GROMACS (ver.4.5.4).

2.3.1. Direct Readout Energy. The direct readout energy (ΔG_{dir}) was estimated as the sum of the nonbonded interaction energy ($E_{nonbonded}$) and the solvation free energy (ΔG_{solv}) in the direct readout process

$$\Delta G_{dir} = E_{nonbonded} + \Delta G_{solv} \tag{1}$$

where

$$E_{nonbonded} = \langle E_{ele} \rangle + \langle E_{van} \rangle \tag{2}$$

where $\langle \cdots \rangle$ denotes the average result for a set of structures of a period simulation trajectory. The nonbonded interaction energy contains electrostatic interaction energy (E_{ele}) and van der Waals interaction energy (E_{van}) between peptide and its protein partner. Both can be obtained using "g_energy" tool within GROMACS

$$\Delta G_{solv} = G_{solv}(complex) - G_{solv}(protein) - G_{solv}(peptide) \tag{3}$$

where

$$G_{solv} = \langle G_{polar} \rangle + \langle G_{nonpolar} \rangle \tag{4}$$

We used GBSA[47,48] implemented in GROMACS (ver.4.5.4) to get the value of G_{solv}, where the polar free energy of solvation, G_{polar}, was calculated from generalized Born equation, while the nonpolar term, $G_{non-polar}$, was calculated using a constant of 2.25936 kJ mol^{-1} nm^{-2} for the value of surface tension with SA algorithms. Here, trajectories of protein and peptide needed for calculations of $G_{solv}(protein)$ and $G_{solv}(peptide)$ were extracted from the trajectory of the corresponding complex. Note that the entropic term of solvent was included here in the solvation free energy and hence in the direct readout energy. Therefore, the direct readout energy is a different concept from enthalpy change of protein-peptide binding.

2.3.2. Indirect Readout Energy. The indirect readout energy (ΔG_{indir}) in protein-peptide recognition was estimated as peptide configurational entropy penalty ($-T\Delta S$) which is only one part of the total entropy penalty of protein-peptide recognition

$$\Delta G_{indir} = -T\Delta S = -T[\langle S_{pep}(bound) \rangle - \langle S_{pep}(unbound) \rangle] \tag{5}$$

where S_{pep}, the absolute peptide configurational entropy, was computed based on the Quasiharmonic approach implemented in GROMACS. It describes the entropy associated with fluctuations of the 3N-6 internal coordinates of the molecule and contains both vibrational entropy and conformational entropy. This term is the reflection of flexibility. In this study, we mainly used the entropy difference of the same peptide in bound and unbound states. Though it was reported that the Quasiharmonic approach overestimates the absolute configurational entropy of the flexible systems,[49−51] the entropy difference of the same molecule will be relatively reliable as the errors tend to be canceled.[18,49] $S_{pep}(bound)$ and $S_{pep}(unbound)$ are the absolute peptide configurational entropies in bound and unbound states, respectively. The simulation temperature T = 300 K was used here.

2.3.3. Binding Free Energy. Apparently, the final binding free energy of protein-peptide recognition is the sum of the direct readout energy and the indirect readout energy.

$$\Delta G_{binding} = \Delta G_{dir} + \Delta G_{indir} \tag{6}$$

2.3.4. Deformation Energy. Deformation energy, ΔE_{deform}, was estimated as the total energy change of peptide from the unbound to the bound state

$$\Delta E_{deform} = \langle E_{pep}(bound) \rangle - \langle E_{pep}(unbound) \rangle \tag{7}$$

in which $\langle E_{pep} \rangle$, the average total energy of peptide, was coarsely obtained from the trajectory which only contains peptide with all other molecules and ions removed. Perhaps, this will lead to an overestimation of the potential energy, especially for the bound peptide without considering influences of both protein and water molecules. As kinetic and potential energies of the peptide constitute E_{pep}, as

$$E_{pep} = E_{pep\text{-}kinetic} + E_{pep\text{-}potential} \tag{8}$$

the deformation energy will be overestimated to some degree. The value of E_{pep} was directly obtained via "g_energy" tool in GROMACS.

2.3.5. Z-Value and W-Value. In order to measure how strongly peptide configurational entropy penalty influences binding free energy, we defined a Z-value as

$$Z\text{-value} = \left| \frac{\Delta G_{indir}}{\Delta G_{dir}} \right| \tag{9}$$

It is also called entropy effect in the context below. A bigger Z-value means a bigger adverse effect that peptide configurational entropy penalty makes to the binding free energy, while a smaller one means the opposite. Similarly, we introduced a W-value to estimate how much difference deformation energy of peptide can make to the binding process.

$$W\text{-value} = \left| \frac{\Delta E_{deform}}{\Delta G_{dir}} \right| \tag{10}$$

dx.doi.org/10.1021/ci5000246 | J. Chem. Inf. Model. 2014, 54, 2022−2032

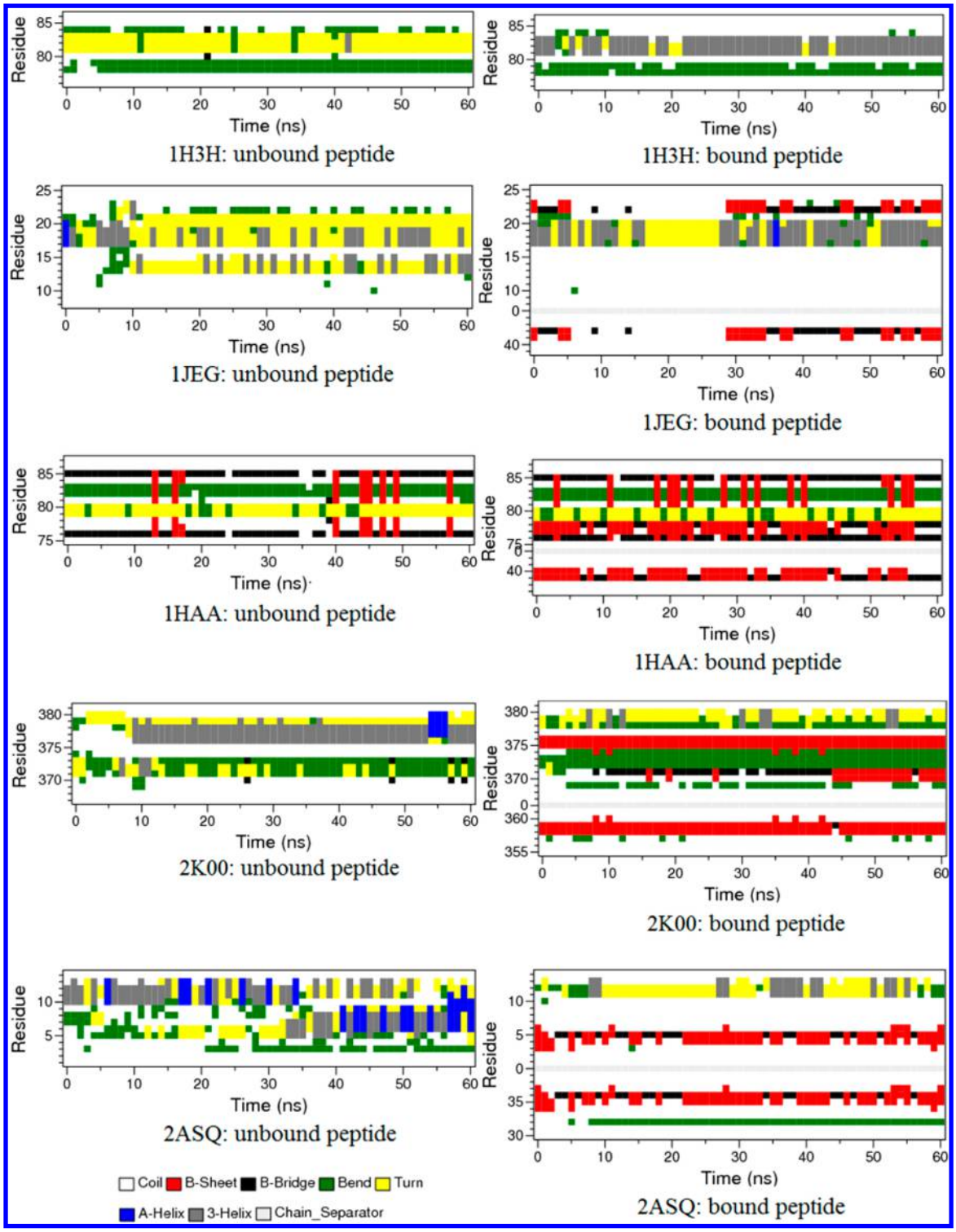


Figure 1. Secondary structure evolution of the unbound and bound peptides during the course of 60 ns production simulations. Figures in the left column show the secondary structure of the unbound peptide; in the right column, the secondary structure of the bound peptide (upper portion above the chain-separator in each figure) and some critical residues of the protein partner (portion below the chain-separator in each figure) that form a β-sheet with a peptide. Residues are numbered as to what they are when we downloaded them from the PDB. Zero in the ordinate of the right column figures represents the chain-separator. These figures were analyzed using "do_dssp" available in GROMACS with DSSP[55] installed.

3. RESULTS AND DISCUSSION

3.1. Description of the Simulation Results. We have simulated 20 complexes in which the bound peptides involve α-helix, right-handed 3_{10}-helix, left-handed poly proline type II helix, β-stand, turn, bend, and coil structures (Table 1 and Table 2). Another 2ASQ-10res obtained from a truncated 2ASQ, where residues 11−14 of the peptide in 2ASQ were cut out, was also simulated for comparison. Analysis of the root-mean-square deviation (RMSD) applied on trajectories of all 25 complexes and 25 unbound peptides (including another four models of 1H3H and 2KHH) (Figure S3) shows equilibrium of

dx.doi.org/10.1021/ci5000246 | J. Chem. Inf. Model. 2014, 54, 2022−2032

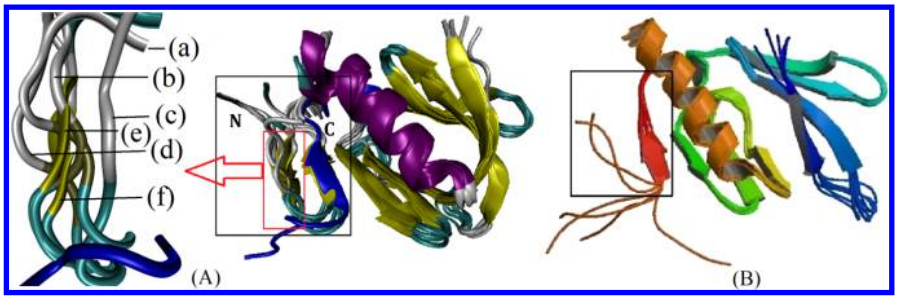


Figure 2. Structures of the 2K00 complex. (A) The right figure shows the superposed complex structures including the starting structure (peptide of which is colored blue) and conformations of six different frames during production simulation of the 2K00 complex; the left figure is the amplification of structures in the red square of the right figure: (a), frame of 10 ns; (b), 20 ns; (c), 30 ns; (d), 40 ns; (e), 50 ns; (f), 60 ns. Yellow represents the extended β-strand. (B) The ensemble of NMR structures of the 2K00 complex downloaded from the PDB. Structures in the black square in (A) and (B) are the similar structures of the peptide.

nearly all simulations in 40 ns−60 ns. Most of the unbound peptides fold into a more compact structure in the last 20 ns (Figure S4). Detailed results of all the energetic components were shown in Table S2.

Unbound peptides of 1BXL, 2K2I, and 2PHE can still keep some stable α-helixes during the 60 ns simulations (Figure S5). The unusual and relatively unstable 3_{10}-helix is observed to live stable in 40−60 ns simulations of complexes 1H3H (residues 81−83) and 1JEG (residues 17−20) (Figure 1) just like their NMR structures; but the 3_{10}-helix disappears in the unbound peptide simulation of 1H3H (Figure 1), which agrees well with the experimental conclusion that 3_{10}-helix in the peptide of 1H3H is induced upon binding rather than being preformed in the unbound state.[32] However, the 3_{10}-helix still exists in several frames in the unbound peptide simulation of 1JEG (Figure 1). Actually, it was reported that the 3_{10}-helix of this peptide in the bound state has only a structural role and forms no direct stabilizing interactions with its protein partner.[31] This indicates that the stability of this 3_{10}-helix mainly comes from the internal factors, and this is probably why it still forms in the unbound peptide of 1JEG.

The bound peptide of the 1HAA complex adopts a β-hairpin structure. It forms an intermolecular antiparallel β-sheet with its protein partner (residues 39−40 in protein chain), meanwhile, nearly all residues of the unbound peptide keep the same structure characteristics with its bound state during the simulation (Figure 1). Little change of the peptide conformations before and after binding is presumably the main reason for small deformation energy in the 1HAA complex forming process (Table 2). The β-strand also exists in the bound peptides of complexes 2K00 (residues 375−376) and 2ASQ (residues 4−5), which respectively forms an antiparallel and a parallel β-sheet with their corresponding protein partner (protein residues 358−359 in 2K00; protein residues 35−36 in 2ASQ) (Figure 1); but both β-strands disappear in the unbound peptides because of instability of singular β-strand. In addition, it is worth noting that N-terminal residues 370−371 of the bound peptide of 2K00 form another relatively stable β-strand after 40 ns simulation (Figure 1), although they do not adopt this kind of structure in our starting structure (Figure 2(A)). Trajectory analysis shows that this N-terminal β-strand (residues 370−371) is probably induced by the nearby pre-existing C-terminal β-strand (residues 375−376). These two β-strands form an antiparallel β-sheet. This phenomenon may not be that strange because there are also similar conformations in the ensemble of NMR structures of the 2K00 complex present in the PDB (Figure 2(B)).

Theoretically, the antiparallel β-sheet is more stable than the parallel one. Our calculated direct readout energies of 2K00 (−309.5603 kJ/mol) and 2ASQ (−185.0797 kJ/mol) also confirm a much stronger intermolecular interaction of 2K00 than 2ASQ, and the calculated binding free energy of 2K00 (−119.0903 kJ/mol) is smaller than that of 2ASQ (−50.1727 kJ/mol). However, the experimental dissociation constant of 2K00 (K_d = 24 μM) is much bigger than that of 2ASQ (K_d = 6.5 ± 0.75 μM). In fact, according to the literature,[28] the dissociation constant of 2ASQ was measured using a truncated peptide involving only residues 1−10, while residues 11−14 of the peptide were ignored because NMR experiment shows that these residues are highly flexible and not restricted by forming complex and that these residues do not form a defined structure. In our simulation results, we found that residues 11−14 of the bound peptide are indeed of high flexibility according to Figure 3 and structured like those in the unbound peptide

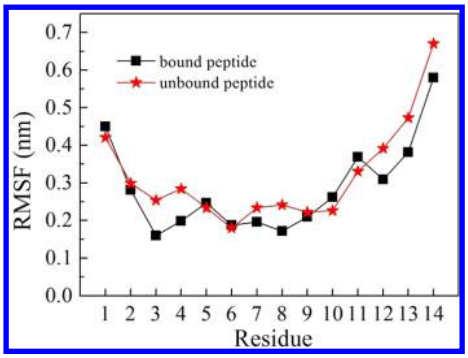


Figure 3. Residue-based root-mean-square fluctuation (RMSF) of all atoms averaged over the last 20 ns simulations of the bound and unbound peptides of the 2ASQ complex.

according to Figure 1. Even so, Figure 3 also shows that flexibility of these four residues in the bound peptide is indeed lower than those in the unbound state on the whole, which suggests that these four residues may provide a nonignorable entropy penalty in binding process. The larger entropy penalty of 2ASQ than 2ASQ-10res shown in Table 2 also reconfirms this. Binding free energy of 2ASQ-10res (−178.1691 kJ/mol) is smaller than 2K00 (−119.0903 kJ/mol). This agrees well with the relative size of their experimental dissociation constants (2K00, 24 μM; 2ASQ-10res, 6.5 ± 0.75 μM). At last, another interesting thing is that complexes 2KFF, 2KFG, and 2KFH which differ from each other only in one peptide residue differ a

dx.doi.org/10.1021/ci5000246 | J. Chem. Inf. Model. 2014, 54, 2022−2032

lot in binding affinities, and our calculated binding free energies agree well with their relative stability.

3.2. Correlation Analysis of the Calculated Binding Free Energy with the Experimental Dissociation Constant. The calculated binding free energy of the 1SY9 complex is a big positive number, while its experimental dissociation constant is very small. The methodology of solvation free energy calculation is unsuitable for this particular system according to our analysis. Therefore, we discarded 1SY9 in the data analysis below. Correlation analysis applied on the remaining 19 complexes shows that there exists a moderate negative correlation (r = −0.527) between the calculated binding free energy and the experimental pK_d (Figure 4). In

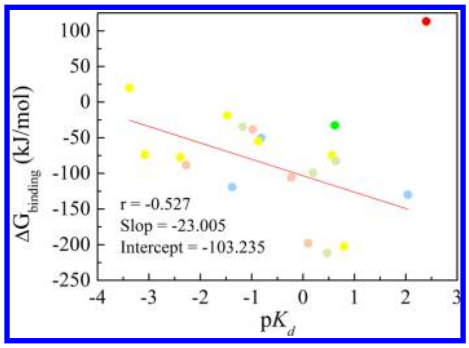


Figure 4. Correlation analysis of the calculated binding free energy with the experimental pK_d. The red dot represents the discarded 1SY9 complex. The other 19 complexes are colored according to the secondary structure of the bound peptide classified in Table 1. The colors are consistent with Table 1.

general, complexes with the similar peptide structure characteristics, colored the same in Figure 4, all distribute uniformly in both sides of the line. This indicates no obvious bias of the energy estimation methodology to a certain kind of peptide structure. On the whole, the calculation results can provide a qualitative analysis though a precise quantitative research is unreachable.

3.3. *Z*-Value and *W*-Value in Protein-Peptide Recognition. Figure 5 intuitively shows that the energy value of the direct readout is usually larger than that of the indirect readout

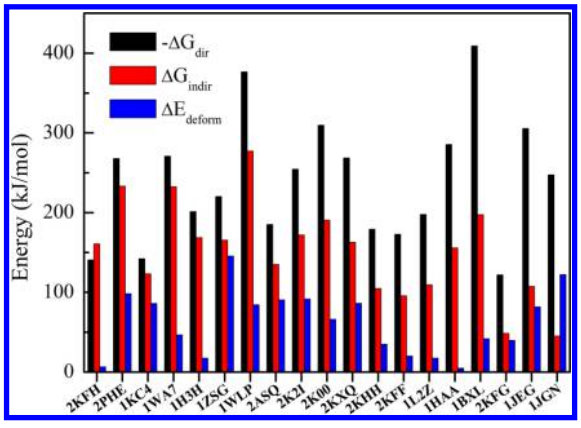


Figure 5. Direct and indirect readout energies and deformation energies of 19 complexes. In order to have a visualized comparison of different energies from the graph, we used the opposite number of the direct readout energy in the figure. Complexes rank as the *Z*-value decreases from left to right.

and the latter is larger than that of the deformation in each complex excluding 2KFH and 1JGN plotted in graph. Indirect peptide configurational entropy penalty is large and not allowed to be ignored. In Table 3, complexes rank as the *Z*-value

Table 3. *Z*-Values and *W*-Values of 19 Complexes

no.	PDB code	*Z*-value	*W*-value
1	2KFH	1.1416	0.0459
2	2PHE	0.8710	0.3671
3	1KC4	0.8670	0.6049
4	1WA7	0.8587	0.1721
5	1H3H	0.8382	0.0861
6	1ZSG	0.7514	0.6611
7	1WLP	0.7364	0.2244
8	2ASQ	0.7289	0.4889
9	2K2I	0.6756	0.3599
10	2K00	0.6153	0.2136
11	2KXQ	0.6069	0.3211
12	2KHH	0.5856	0.1948
13	2KFF	0.5535	0.1154
14	1L2Z	0.5534	0.0889
15	1HAA	0.5451	0.0166
16	1BXL	0.4825	0.1020
17	2KFG	0.3975	0.3263
18	1JEG	0.3513	0.2676
19	1JGN	0.1824	0.4945

decreases. *Z*-values of the first 15 complexes are more than 0.5. Value of the indirect readout energy even exceeds four-fifths of the direct readout energy for the first five complexes. Then, *Z*-values of the next six complexes vary from 0.8 to 0.6 followed by another four complexes whose *Z*-values decrease within the scope of 0.6−0.5. Among the remaining four complexes, *Z*-values range from 0.5 to 0.1. However, there are 17 complexes whose *W*-values are less than 0.5 with the largest and smallest values equal to 0.4945 and 0.0166, respectively. Though the remaining two complexes, 1KC4 and 1ZSG, have larger *W*-values, they are still smaller than their respective *Z*-values. The large *Z*-values in Table 3 suggest a large entropy effect in protein-peptide recognition, and this further indicates important roles of the flexibility of peptides in recognition.

3.4. Value of the Peptide Configurational Entropy Penalty. Both a stronger intermolecular interaction and a more flexible unbound peptide can result in a larger reduction of degrees of freedom and thus a larger decrease of the entropy (ΔS) for peptide in recognition. Hence, a larger peptide configurational entropy penalty appears. For different complexes, if they have similar interactions, value of the peptide configurational entropy penalty will depend on the flexibility of their peptides. 1WA7, 2KXQ, and 2PHE for example, the absolute configurational entropies of their unbound peptides are 6892.16 J/(mol K), 5556.32 J/(mol K), and 7901.85 J/(mol K) (Table S2), respectively, while the differences among their direct readout energies are small (1WA7, −270.6838 kJ/mol; 2KXQ, −268.4203 kJ/mol; 2PHE, −267.8501 kJ/mol) (Table 2). Peptide configurational entropy penalty of 2PHE is larger than that of 1WA7, while the latter is larger than that of 2KXQ as expected (2PHE, 233.289 kJ/mol; 1WA7, 232.428 kJ/mol; 2KXQ, 162.909 kJ/mol) (Table 2). In another example, 2KFF and 1L2Z, though peptide of 2KFF is a little more flexible than that of 1L2Z (2KFF, 3642.34 J/(mol K); 1L2Z, 3605.57 J/(mol K), its indirect readout energy is less than the latter (2KFF,

dx.doi.org/10.1021/ci5000246 | J. Chem. Inf. Model. 2014, 54, 2022−2032

95.697 kJ/mol; 1L2Z, 109.344 kJ/mol) because of a stronger intermolecular interaction in 1L2Z complex (2KFF, −172.8848 kJ/mol; 1L2Z, −197.5866 kJ/mol). This situation occurs between 2KHH and 1H3H as well.

3.5. Does a Larger Peptide Configurational Entropy Penalty Mean a Bigger Entropy Effect, i.e. a Bigger *Z*-Value? To answer this question, we studied the relationship between the direct and indirect readout energies. Figure 6

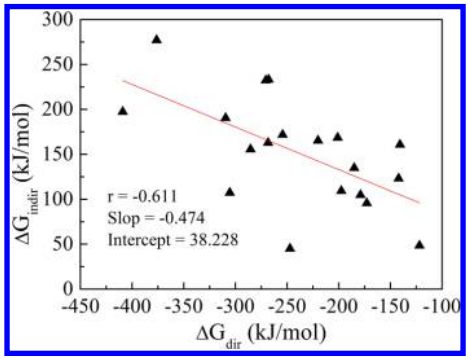


Figure 6. Correlation analysis between the direct and indirect readout energies.

shows that a moderate negative Pearson correlation exists (r = −0.611). This situation is very similar to enthalpy−entropy compensation. A binding process with a larger indirect readout energy usually has a corresponding smaller direct readout energy, and this may lead to a smaller *Z*-value. Complex 1WLP for example, its indirect readout energy is the largest of all (Figure 5) while its *Z*-value is not. However, a smaller indirect readout energy is usually accompanied by a larger direct readout energy, and this may lead to a larger *Z*-value. Complex 1KC4 for example, its indirect readout energy is much smaller than 1WLP (Figure 5), but its *Z*-value is larger. Therefore, size of the *Z*-value cannot be directly judged by the size of the indirect readout energy without exact calculations. A larger peptide configurational entropy penalty does not necessarily mean a bigger entropy effect.

However, a large entropy effect is obviously unfavorable in protein-peptide recognition. In order to get a tight-binding protein-peptide complex, we need to ensure a small total binding free energy. We can try to increase nonbonded interactions and decrease the unbound peptide flexibility at the same time. In fact, it has long been recognized that a less flexible ligand is likely to induce fewer entropy penalty upon binding and will be a good selection of candidate ligands.[18] In addition, it is a common observation that "known drug peptides all have low entropy".[56] However, more rigid peptide is weaker in adjusting to fit the binding pocket, which results in low nonbonded interactions even nonrecognition. Thus, how to balance higher nonbonded interactions and more rigid peptide with suitable conformation is still a big challenge.

3.6. Structure and Entropy Penalty. Table 1 clearly shows the secondary structure of the bound peptide. We classified these 20 complexes according to the similarities of peptide structure. Different classes were distinguished by different colors. In Figure 7, we calculated the average entropy penalty of residue ($-T\Delta S$ divided by the number of residues) in each peptide. The colors are consistent with Table 1. Complexes colored yellow are observed an obvious aggregation in the left part of the figure. All their bound peptides adopt flexible turn, bend, and coil structures (Table 1). This indicates

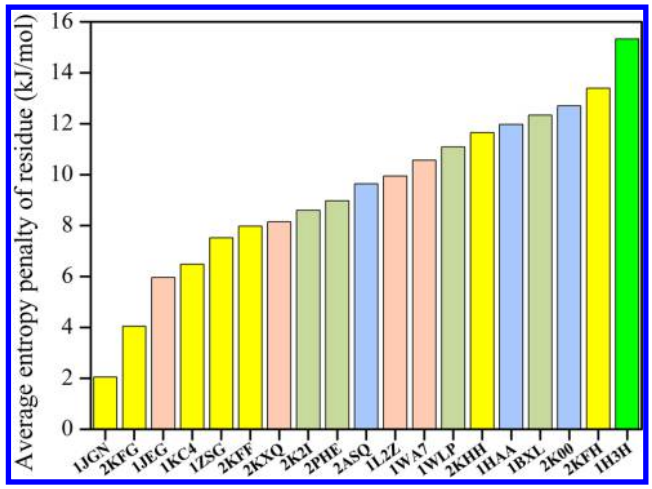


Figure 7. Average entropy penalty of residue in each peptide of the 19 complexes.

a tendency of less average entropy penalty of residue, if the peptide adopts only flexible structures after binding. Complexes colored light red are distributed in the left and middle part of the figure. Their bound peptides adopt poly proline type II helix which is more rigid than turn, bend, and coil but more flexible than other helixes (α-helix, 3_{10}-helix and so on) for the lack of internal hydrogen bonding. The remaining complexes which their bound peptides adopt part of the rigid α-helix, 3_{10}-helix, or β-sheet tend to scatter in the middle and right part. Average entropy penalty of residue in these peptides tend to be larger. The largest average entropy penalty occurs in complex 1H3H. Structures of its unbound peptide are all turn, bend, and coil, while its bound peptide adopts a stable existed 3_{10}-helix (Figure 1). Maybe this big difference in the flexibility of structure is a reason for the big entropy penalty. On the whole, the appearance of the rigid secondary structure in the bound peptide might cause an increase of the average entropy penalty of residue.

3.7. Flexibility Analysis of the Peptide Residues of the 2KFF Complex. During the last 20 ns simulations of the 2KFF complex and its unbound peptide, four terminal residues in each peptide (residues 143, 144, 153, and 154) keep flexible coil structure while those in the central position form relatively rigid turn, bend, and helix structures (Figure 8). This structural difference leads to higher RMSF values of the terminal residues compared with the central ones in both peptides except one end of the unbound peptide (Figure 9). Figure 10(A) and 10(B) show that residue Lys154 of the unbound peptide can form two long-lived hydrogen bonds with Ser147 adjacent in space. A moderate (N−H···O: 2.804 Å) main chain-main chain hydrogen bond exists between one carboxyl oxygen atom of Lys154 and amino hydrogen of Ser147, at the same time, another strong (O−H···O: 2.609 Å) main chain-side chain hydrogen bond forms between the same oxygen atom of Lys154 and hydroxyl hydrogen of Ser147 (Figure 11(A)). Both these two hydrogen bonds nearly live in the whole 60 ns simulation (Figure 10(A) or 10(B)), which severely restrains the degrees of freedom of Lys154 and Ser147. In addition, Figure 10(A) shows that there also exists a shorter-lived and weaker (N−H···O: 2.958 Å) hydrogen bond between another oxygen atom of Lys154 and amino hydrogen of Glu146 (Figure 11(A)). These three hydrogen bonds are presumably the main reason for low flexibility of the terminal that Lys154 locates.

dx.doi.org/10.1021/ci5000246 | J. Chem. Inf. Model. 2014, 54, 2022−2032

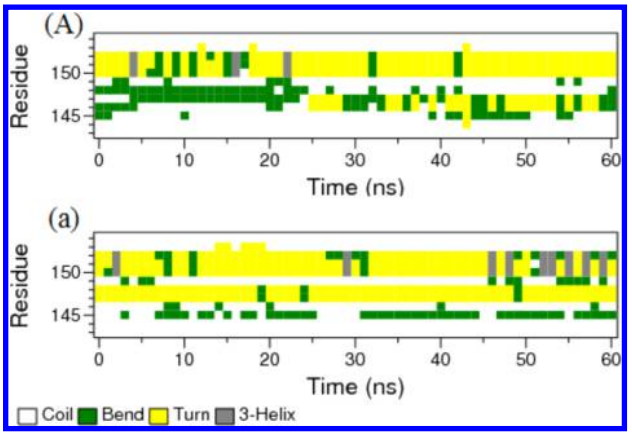


Figure 8. Secondary structure evolution of the unbound and bound peptides of the 2KFF complex during the course of 60 ns production simulations: (A) the unbound peptide of the 2KFF complex and (a) the bound peptide in the 2KFF complex.

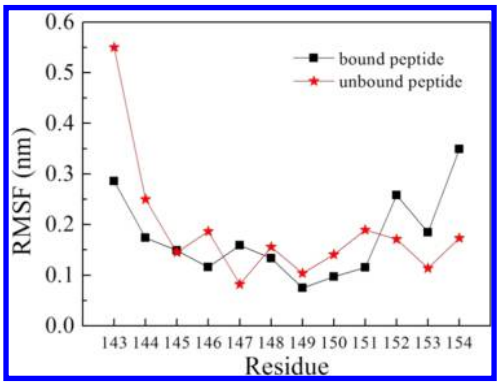


Figure 9. Residue-based root-mean-square fluctuation (RMSF) of all atoms averaged over the last 20 ns simulations of the bound and unbound peptides of the 2KFF complex.

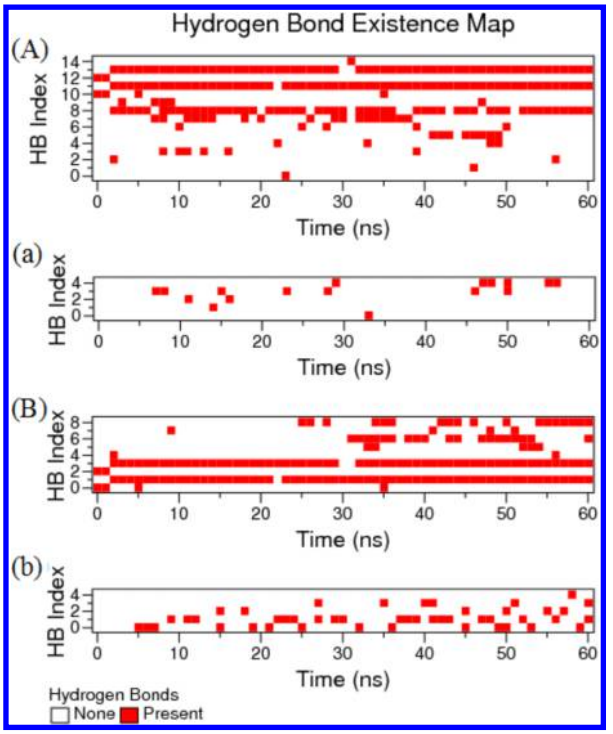


Figure 10. Hydrogen bond evolution of residues Lys154 and Ser147 during the whole 60 ns production simulations of the unbound peptide and the complex. The HB Index refers to the hydrogen bond index. (A) and (B), hydrogen bonds of residues Lys154 and Ser147 in the unbound peptide formed with all the remaining unbound peptide residues, respectively. (a) and (b), hydrogen bonds of residues Lys154 and Ser147 in the bound peptide formed with all the remaining residues of the 2KFF complex, respectively. HB Indexes 11 and 13 in (A) and 1 and 3 in (B) represent the two hydrogen bonds between residue Lys154 and Ser147. The HB Index 8 in (A) represents the weaker hydrogen bond between residue Lys154 and Glu146.

However, after peptide binds to its protein partner, the two hydrogen bonds between Lys154 and Ser147 disappear because of large distances between corresponding atoms (H···O: 14.1 and 14.8 Å in Figure 11(B)); and owing to exposure outside of the binding pocket (Figure 11(B)), these two residues almost cannot form strong interactions with protein, neither. Figure 10(a) and 10(b) show that hydrogen bonds of the two residues in the bound state are few and transient in the last 20 ns simulation. Under these circumstances, flexibility of Ser147 and Lys154 in bound peptide is higher than those unbound. With the removal of the main restraints imposed by hydrogen bonds, terminal residues in the bound peptide exhibit their original high-flexibility of coil structure.

The less flexible end of the unbound peptide indicates that not all the terminal coil structures are more flexible than the central relatively rigid ones. Meanwhile, the increased flexibility of this end and residue Ser147 in the bound peptide suggests that not every residue is certain to become less flexible after binding though the total peptide configurational entropy decreases.

3.8. Amino Acid Propensities of the Unbound Peptide Flexibility. Flexibility of a residue is not permanent. It changes with its own secondary structure and the influence of the surrounding space environment. In order to compare the flexibility of different amino acids, we computed the average RMSF value of each residue of the unbound peptides. A statistical RMSF figure involving all influences of structures and surrounding environments, Figure 12, was done using all of the unbound peptide data we had now. This figure shows that the RMSF value of the backbone is usually more than half of the value of all atoms. This indicates that the backbone flexibility contributes more to the entire peptide flexibility than the rest part of the residue especially for Gly and Ala because of their fewer atoms of side chains. Excluding the residues of less than 10 statistical samples (His, Met, Trp, Gln, and Cys in Figure S1), Asp and Lys are more flexible than other residues in both backbone flexibility only and flexibility of all atoms; but owing to small samples here, this figure perhaps cannot show the real scale of flexibility of each residue. For more accurate results, we need more samples.

4. CONCLUSIONS

In classical biomolecular recognition, protein−DNA recognition for example, protein and DNA bind directly by amino acid−base interactions and indirectly by deformation of DNA. The main source of the indirect readout energy is the DNA deformation energy. However, in this paper, we made a systematic quantitative research on 20 protein-peptide complexes via molecular dynamics simulations and found that the peptide configurational entropy penalty in the indirect readout process of protein-peptide recognition is large and not allowed to be ignored, while deformation energy of peptide is

dx.doi.org/10.1021/ci5000246 | J. Chem. Inf. Model. 2014, 54, 2022−2032

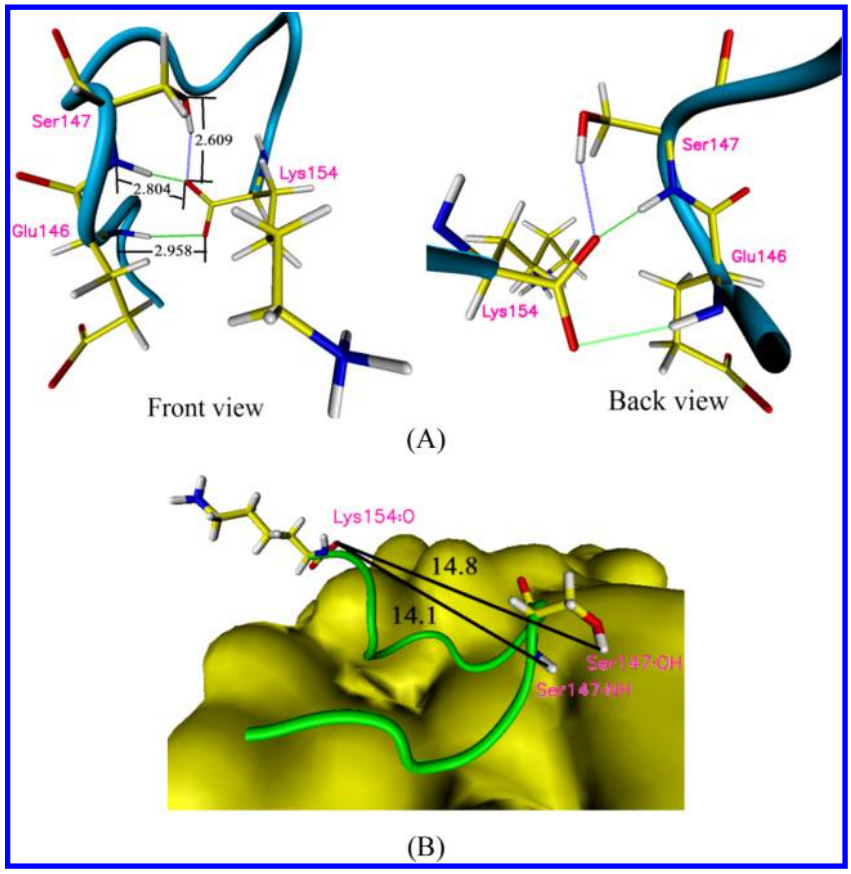


Figure 11. (A) Hydrogen bonds among Lys154, Ser147, and Glu146 of the unbound peptide (cyan) of the 2KFF complex in front and back views. Main chain-main chain hydrogen bonds are colored green; main chain-side chain hydrogen bonds, blue. Hydrogen bond length (O−H···O and N−H···O) is the average distance between donor (O/N) and acceptor (O) of the last 20 ns. (B) 2KFF complex. Distance between atoms (H···O) is the average value over the last 20 ns. Hydrogen bond length in (A) and distance between atoms in (B) are all in Å. Lys154, Ser147, and Glu146 are shown in stick style, which yellow represents carbon; red, oxygen; blue, nitrogen; white, hydrogen.

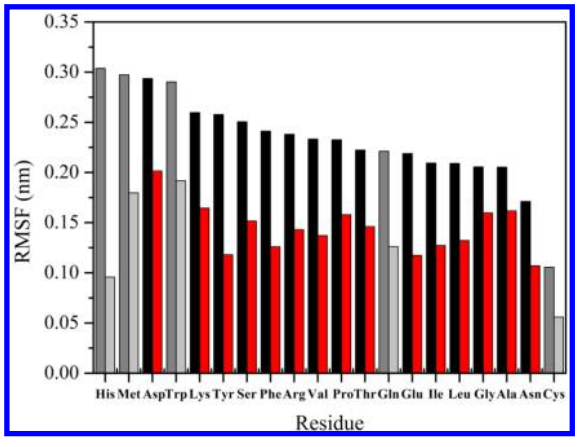


Figure 12. Statistical root-mean-square fluctuation (RMSF) of 20 amino acid residues. Black represents RMSF of all atoms; red, RMSF of the backbone. For His, Met, Trp, Gln and Cys residues, heavy gray and light gray are used to show all-atoms RMSF and the backbone RMSF, instead.

not obvious in comparison. Peptide configurational entropy penalty becomes the main source of the indirect readout energy of protein-peptide recognition. We further discussed the scale of the entropy penalty, the size of the entropy effect, relationship between structure and entropy penalty, and the flexibility of peptide and residues in detail. We hope that our discussions can be helpful for a better understanding of protein-peptide recognition and the design of peptide ligands.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Structures of 20 protein-peptide complexes and their experimental conditions and assay methods of dissociation constants (K_d); detailed results of all the energetic components; statistical numbers of various amino acids in peptides of the data set; influence of different models on the correlation analysis; root-mean-square deviation (RMSD) of all 25 complexes and 25 unbound peptides; the aligned unbound peptide structures; secondary structure evolution of the unbound and bound peptides during the course of 60 ns production simulations of 1BXL, 2K2I, and 2PHE. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: mldeng@zju.edu.cn (M.D.).
*E-mail: shangzc@zju.edu.cn (Z.S.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank National Natural Science Foundation of China (No. 11172259) for financial support.

■ REFERENCES

(1) Neduva, V.; Linding, R.; Su-Angrand, I.; Stark, A.; Masi, F. D.; Gibson, T. J.; Lewis, J.; Serrano, L.; Russell, R. B. Systematic Discovery of New Recognition Peptides Mediating Protein Interaction Networks. *PLoS Biol.* **2005**, *3*, e405.

(2) Neduva, V.; Russell, R. B. Peptides Mediating Interaction Networks: New Leads at Last. *Curr. Opin. Biotechnol.* **2006**, *17*, 465−471.

(3) Petsalaki, E.; Russell, R. B. Peptide-Mediated Interactions in Biological Systems: New Discoveries and Applications. *Curr. Opin. Biotechnol.* **2008**, *19*, 344−350.

(4) Puntervoll, P.; Linding, R.; Gemünd, C.; Chabanis-Davidson, S.; Mattingsdal, M.; Cameron, S.; Martin, D. M. A.; Ausiello, G.; Brannetti, B.; Costantini, A.; Ferrè, F.; Maselli, V.; Via, A.; Cesareni, G.; Diella, F.; Superti-Furga, G.; Wyrwicz, L.; Ramu, C.; Mcguigan, C.; Gudavalli, R.; Letunic, I.; Bork, P.; Rychlewski, L.; Küster, B.; Helmer-Citterich, M.; Hunter, W. N.; Aasland, R.; Gibson, T. J. ELM Server: A New Resource for Investigating Short Functional Sites in Modular Eukaryotic Proteins. *Nucleic Acids Res.* **2003**, *31*, 3625−3630.

(5) Vanhee, P.; Stricher, F.; Baeten, L.; Verschueren, E.; Lenaerts, T.; Serrano, L.; Rousseau, F.; Schymkowitz, J. Protein-Peptide Interactions Adopt the Same Structural Motifs as Monomeric Protein Folds. *Structure* **2009**, *17*, 1128−1136.

(6) London, N.; Movshovitz-Attias, D.; Schueler-Furman, O. The Structural Basis of Peptide-Protein Binding Strategies. *Structure* **2010**, *18*, 188−199.

(7) Stein, A.; Aloy, P. Novel Peptide-Mediated Interactions Derived from High-Resolution 3-Dimensional Structures. *PLoS Comput. Biol.* **2010**, *6*, e1000789.

(8) Vanhee, P.; Reumers, J.; Stricher, F.; Baeten, L.; Serrano, L.; Schymkowitz, J.; Rousseau, F. PepX: A Structural Database of Non-Redundant Protein−Peptide Complexes. *Nucleic Acids Res.* **2010**, *38*, D545−D551.

(9) Grigoryan, G.; Reinke, A. W.; Keating, A. E. Design of Protein-Interaction Specificity Gives Selective bZIP-Binding Peptides. *Nature* **2009**, *458*, 859−864.

(10) Vanhee, P.; Van Der Sloot, A. M.; Verschueren, E.; Serrano, L.; Rousseau, F.; Schymkowitz, J. Computational Design of Peptide Ligands. *Trends Biotechnol.* **2011**, *29*, 231−239.

(11) Michael Gromiha, M.; Siebers, J. G.; Selvaraj, S.; Kono, H.; Sarai, A. Intermolecular and Intramolecular Readout Mechanisms in Protein−DNA Recognition. *J. Mol. Biol.* **2004**, *337*, 285−294.

(12) Ahmad, S.; Kono, H.; Araúzo-Bravo, M. J.; Sarai, A. ReadOut: Structure-Based Calculation of Direct and Indirect Readout Energies and Specificities for Protein−DNA Recognition. *Nucleic Acids Res.* **2006**, *34*, W124−W127.

(13) Fujii, S.; Kono, H.; Takenaka, S.; Go, N.; Sarai, A. Sequence-Dependent DNA Deformability Studied Using Molecular Dynamics Simulations. *Nucleic Acids Res.* **2007**, *35*, 6063−6074.

(14) Araúzo-Bravo, M. J.; Sarai, A. Indirect Readout in Drug-DNA Recognition: Role of Sequence-Dependent DNA Conformation. *Nucleic Acids Res.* **2008**, *36*, 376−386.

(15) Yamasaki, S.; Terada, T.; Shimizu, K.; Kono, H.; Sarai, A. A Generalized Conformational Energy Function of DNA Derived from Molecular Dynamics Simulations. *Nucleic Acids Res.* **2009**, *37*, e135.

(16) Yamasaki, S.; Terada, T.; Kono, H.; Shimizu, K.; Sarai, A. A New Method for Evaluating the Specificity of Indirect Readout in Protein−DNA Recognition. *Nucleic Acids Res.* **2012**, *40*, e129.

(17) Steffen, N. R.; Murphy, S. D.; Tolleri, L.; Hatfield, G. W.; Lathrop, R. H. DNA Sequence and Structure: Direct and Indirect Recognition in Protein-DNA Binding. *Bioinformatics* **2002**, *18*, S22−S30.

(18) Chang, C.-E. A.; Chen, W.; Gilson, M. K. Ligand Configurational Entropy and Protein Binding. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1534−1539.

(19) Killian, B. J.; Kravitz, J. Y.; Somani, S.; Dasgupta, P.; Pang, Y.-P.; Gilson, M. K. Configurational Entropy in Protein−Peptide Binding: Computational Study of Tsg101 Ubiquitin E2 Variant Domain with an HIV-Derived PTAP Nonapeptide. *J. Mol. Biol.* **2009**, *389*, 315−335.

(20) Rose, P. W.; Bi, C.; Bluhm, W. F.; Christie, C. H.; Dimitropoulos, D.; Dutta, S.; Green, R. K.; Goodsell, D. S.; Prlić, A.; Quesada, M.; Quinn, G. B.; Ramos, A. G.; Westbrook, J. D.; Young, J.; Zardecki, C.; Berman, H. M.; Bourne, P. E. The RCSB Protein Data Bank: New Resources for Research and Education. *Nucleic Acids Res.* **2013**, *41*, D475−D482.

(21) Sattler, M.; Liang, H.; Nettesheim, D.; Meadows, R. P.; Harlan, J. E.; Eberstadt, M.; Yoon, H. S.; Shuker, S. B.; Chang, B. S.; Minn, A. J.; Thompson, C. B.; Fesik, S. W. Structure of Bcl-x_L-Bak Peptide Complex: Recognition Between Regulators of Apoptosis. *Science* **1997**, *275*, 983−986.

(22) Contessa, G.; Orsale, M.; Melino, S.; Torre, V.; Paci, M.; Desideri, A.; Cicero, D. Structure of Calmodulin Complexed with an Olfactory CNG Channelfragment and Role of the Central Linker: Residual Dipolar Couplingsto Evaluate Calmodulin Binding Modes Outside the Kinase Family. *J. Biomol. NMR* **2005**, *31*, 185−199.

(23) Martinez-Sanz, J.; Kateb, F.; Assairi, L.; Blouquit, Y.; Bodenhausen, G.; Abergel, D.; Mouawad, L.; Craescu, C. T. Structure, Dynamics and Thermodynamics of the Human Centrin 2/hSfi1 Complex. *J. Mol. Biol.* **2010**, *395*, 191−204.

(24) Jonker, H. R. A.; Wechselberger, R. W.; Boelens, R.; Folkers, G. E.; Kaptein, R. Structural Properties of the Promiscuous VP16 Activation Domain. *Biochemistry* **2004**, *44*, 827−839.

(25) Scherf, T.; Kasher, R.; Balass, M.; Fridkin, M.; Fuchs, S.; Katchalski-Katzir, E. A β-Hairpin Structure in a 13-Mer Peptide That Binds α-Bungarotoxin with High Affinity and Neutralizes Its Toxicity. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 6629−6634.

(26) Kozlov, G.; De Crescenzo, G.; Lim, N. S.; Siddiqui, N.; Fantus, D.; Kahvejian, A.; Trempe, J.-F.; Elias, D.; Ekiel, I.; Sonenberg, N.; O'Connor-McCourt, M.; Gehring, K. Structural Basis of Ligand Recognition by PABC, a Highly Specific Peptide-Binding Domain Found in Poly(A)-Binding Protein and a HECT Ubiquitin Ligase. *EMBO J.* **2004**, *23*, 272−281.

(27) Moise, L.; Piserchio, A.; Basus, V. J.; Hawrot, E. NMR Structural Analysis of α-Bungarotoxin and Its Complex with the Principal α-Neurotoxin-Binding Sequence on the α7 Subunit of a Neuronal Nicotinic Acetylcholine Receptor. *J. Biol. Chem.* **2002**, *277*, 12406−12417.

(28) Song, J.; Zhang, Z.; Hu, W.; Chen, Y. Small Ubiquitin-Like Modifier (SUMO) Recognition of a SUMO Binding Motif: A Reversal of the Bound Orientation. *J. Biol. Chem.* **2005**, *280*, 40122−40129.

(29) Wegener, K. L.; Basran, J.; Bagshaw, C. R.; Campbell, I. D.; Roberts, G. C. K.; Critchley, D. R.; Barsukov, I. L. Structural Basis for the Interaction Between the Cytoplasmic Domain of the Hyaluronate Receptor Layilin and the Talin F3 Subdomain. *J. Mol. Biol.* **2008**, *382*, 112−126.

(30) Hobeika, M.; Brockmann, C.; Gruessing, F.; Neuhaus, D.; Divita, G.; Stewart, M.; Dargemont, C. Structural Requirements for the Ubiquitin-Associated Domain of the mRNA Export Factor Mex67 to Bind Its Specific Targets, the Transcription Elongation THO Complex Component Hpr1 and Nucleoporin FXFG Repeats. *J. Biol. Chem.* **2009**, *284*, 17575−17583.

(31) Ghose, R.; Shekhtman, A.; Goger, M. J.; Ji, H.; Cowburn, D. A Novel, Specific Interaction Involving the Csk SH3 Domain and Its Natural Ligand. *Nat. Struct. Mol. Biol.* **2001**, *8*, 998−1004.

(32) Liu, Q.; Berry, D.; Nash, P.; Pawson, T.; McGlade, C. J.; Li, S. S.-C. Structural Basis for Specific Binding of the Gads SH3 Domain to an RxxK Motif-Containing SLP-76 Peptide: A Novel Mode of Peptide Recognition. *Mol. Cell* **2003**, *11*, 471−481.

(33) Schweimer, K.; Hoffmann, S.; Bauer, F.; Friedrich, U.; Kardinal, C.; Feller, S. M.; Biesinger, B.; Sticht, H. Structural Investigation of the

Binding of a Herpesviral Protein to the SH3 Domain of Tyrosine Kinase Lck. *Biochemistry* **2002**, *41*, 5120−5130.
(34) Ogura, K.; Nobuhisa, I.; Yuzawa, S.; Takeya, R.; Torikai, S.; Saikawa, K.; Sumimoto, H.; Inagaki, F. NMR Solution Structure of the Tandem Src Homology 3 Domains of P47phox Complexed with a P22phox-Derived Proline-Rich Peptide. *J. Biol. Chem.* **2006**, *281*, 3660−3668.
(35) Chong, P. A.; Lin, H.; Wrana, J. L.; Forman-Kay, J. D. Coupling of Tandem Smad Ubiquitination Regulatory Factor (Smurf) WW Domains Modulates Target Specificity. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 18404−18409.
(36) Mott, H. R.; Nietlispach, D.; Evetts, K. A.; Owen, D. Structural Analysis of the SH3 Domain of β-PIX and Its Interaction with α-P21 Activated Kinase (PAK). *Biochemistry* **2005**, *44*, 10977−10983.
(37) Freund, C.; Kühne, R.; Yang, H.; Park, S.; Reinherz, E. L.; Wagner, G. Dynamic Interaction of CD2 with the GYF and the SH3 Domain of Compartmentalized Effector Molecules. *EMBO J.* **2002**, *21*, 5985−5995.
(38) Kieken, F.; Jović, M.; Tonelli, M.; Naslavsky, N.; Caplan, S.; Sorgen, P. L. Structural Insight into the Interaction of Proteins Containing NPF, DPF, and GPF Motifs with the C-Terminal EH-Domain of EHD1. *Protein Sci.* **2009**, *18*, 2471−2479.
(39) Yang, L. W.; Eyal, E.; Chennubhotla, C.; Jee, J.; Gronenborn, A. M.; Bahar, I. Insights into Equilibrium Dynamics of Proteins from Comparison of NMR and X-Ray Data with Computational Predictions. *Structure* **2007**, *15*, 741−749.
(40) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(41) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(43) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(44) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(45) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N [Center-Dot] Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(46) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(47) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins* **2004**, *55*, 383−394.
(48) Schaefer, M.; Bartels, C.; Karplus, M. Solution Conformations and Thermodynamics of Structured Peptides: Molecular Dynamics Simulation with an Implicit Solvation Model. *J. Mol. Biol.* **1998**, *284*, 835−848.
(49) Schlitter, J. Estimation of Absolute and Relative Entropies of Macromolecules Using the Covariance Matrix. *Chem. Phys. Lett.* **1993**, *215*, 617−621.
(50) Andricioaei, L.; Karplus, M. On the Calculation of Entropy from Covariance Matrices of the Atomic Fluctuations. *J. Chem. Phys.* **2001**, *115*, 6289−6292.
(51) Chang, C.-E.; Chen, W.; Gilson, M. K. Evaluating the Accuracy of the Quasiharmonic Approximation. *J. Chem. Theory. Comput.* **2005**, *1*, 1017−1028.
(52) Liu, M.; Chen, T. Y.; Ahamed, B.; Li, J.; Yau, K. W. Calcium-Calmodulin Modulation of the Olfactory Cyclic Nucleotide-Gated Cation Channel. *Science* **1994**, *266*, 1348−1354.
(53) Spiga, O.; Bernini, A.; Scarselli, M.; Ciutti, A.; Bracci, L.; Lozzi, L.; Lelli, B.; Di Maro, D.; Calamandrei, D.; Niccolai, N. Peptide−Protein Interactions Studied by Surface Plasmon and Nuclear Magnetic Resonances. *FEBS Lett.* **2002**, *511*, 33−35.
(54) Berry, D. M.; Nash, P.; Liu, S. K. W.; Pawson, T.; McGlade, C. J. A High-Affinity Arg-X-X-Lys SH3 Binding Motif Confers Specificity for the Interaction Between Gads and SLP-76 in T Cell Signaling. *Curr. Biol.* **2002**, *12*, 1336−1341.
(55) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22*, 2577−2637.
(56) Unal, E. B.; Gursoy, A.; Erman, B. Conformational Energies and Entropies of Peptides, and the Peptide−Protein Binding Problem. *Phys. Biol.* **2009**, *6*, 036014.