Theor Chem Acc (1999) 101:274–281
DOI 10.1007/s00214980m133

Theoretical
Chemistry Accounts

Regular article

Conformational free energy surface of the linear DPDPE peptide: cost of pre-organization for disulfide bond formation

Yan Wang, Krzysztof Kuczera

Department of Chemistry and Department of Molecular Biosciences, University of Kansas, 2010 Malott Hall, Lawrence, KS 66045, USA

Received: 20 April 1998 / Accepted: 9 September 1998 / Published online: 7 December 1998

Abstract. We have calculated the free energy differences between four conformers of the linear form of the opioid pentapeptide DPDPE in aqueous solution. The conformers are Cyc, representing the structure adopted by the linear peptide prior to disulfide bond formation, β_C and β_E, two slightly different β-turns previously identified in unconstrained molecular dynamics simulations, and Ext, an extended structure. Our simulations indicate that β_E is the most stable of the studied conformers of linear DPDPE in aqueous solution, with β_C, Cyc and Ext having free energies higher by 2.3, 6.3, and 28.2 kcal/mol, respectively. The free energy differences of 4.0 kcal/mol between β_C and Cyc, and 6.3 kcal/mol between β_E and Cyc, reflect the cost of pre-organizing the linear peptide into a conformation conducive for disulfide bond formation. Such a conformational change is a pre-requisite for the chemical reaction of S—S bond formation to proceed. The relatively low population of the cyclic-like structure agrees qualitatively with observed lower potency and different receptor specificity of the linear form relative to the cyclic peptide, and with previous unconstrained simulation results. Free energy component analysis indicates that the moderate stability difference of 4.0–6.3 kcal/mol between the β-turns and the cyclic-like structure results from cancellation of two large opposing effects. In accord with intuition, the relaxed β-turns have conformational strain 43–45 kcal/mol lower than the Cyc structure. However, the cyclic-like conformer interacts with water about 39 kcal/mol strongly than the open β-turns. Our simulations are the first application of the recently developed multidimensional conformational free energy thermodynamic integration (CFTI) protocol to a solvated system, with fast convergence of the free energy obtained by fixing all flexible dihedrals. Additionally, the availability of the CFTI multidimensional free energy gradient leads to a new decomposition scheme, giving the contribution of each fixed dihedral to the overall free energy change and providing additional insight into the microscopic mechanisms of the studied processes.

Correspondence to: K. Kuczera
e-mail: kuczera@tedybr.chem.ukans.edu

Key words: Free energy simulations – DPDPE peptide – Peptide conformational equilibria in solution – Conformational free energy – Disulfide bonds

1 Introduction

Disulfide bonds perform a wide range of biological functions, participating in a number of biochemical reactions [1, 2], stabilizing native three-dimensional structures of proteins [3, 4], and inducing conformations important for biological activity in short peptides such as conotoxin [5], endothelin [6], oxytocin, and vasopressin [7]. Disulfide bridges have been utilized in peptide drug design to constrain the peptides to high affinity conformations for their receptors [8]. One such peptide is the opioid DPDPE (Tyr-D-Pen-Gly-Phe-D-Pen), where Pen is penicillamine, or β,β-dimethylcysteine. The cyclic form of DPDPE, designed with topographical and conformational constraints [8], is a highly specific δ opioid. One of the reasons for the intensive research focused on DPDPE and other opioids is that their affinity and selectivity for different receptors are strongly correlated with conformations. DPDPE has been the object of a large number of studies, including low energy conformational search [9, 10], molecular mechanics with NMR constraints [11], quenched molecular dynamics [12], molecular dynamics simulations [13–17], X-ray crystallography [18], and NMR [11, 19, 20].

In what follows we use the terms "linear DPDPE" to specifically denote the reduced, acyclic [DPDPE$(SH)_2$] form of the studied peptide, "cyclic DPDPE" to specifically denote the oxidized, disulfide-bridged form. In a recent study, we have generated 1 ns length molecular dynamics (MD) trajectories of both linear and cyclic forms of the DPDPE peptide in aqueous solution, aimed at delineating the influence of the disulfide bond on peptide properties [21]. In that work we were able to rationalize the observed differences in potency and membrane permeability between linear and cyclic DPDPE in terms of the variation of their physical

properties found in the simulations: structural flexibility, diffusion rates, and dipole moments [21]. The simulation results identified two type IV β-turn conformers as stable, representative structures of linear DPDPE in solution. This was based on the fact that two independent 1 ns simulations of linear DPDPE, starting from a cyclic-like and an extended initial structure, converged to similar conformations of IV β-turn character. Additionally, the simulations indicated a low population of the presumably biologically active cyclic-like structure for linear DPDPE. Because of the limited conformational sampling in the unconstrained trajectories, this population was difficult to quantify.

The primary goal of this work is to use free energy simulations to obtain a quantitative measure of the population of the cyclic-like structure relative to the β-turns for linear DPDPE. This result will provide an estimate of the population of the biologically active form of linear DPDPE in aqueous solution. Further, the calculated free energy difference will also reflect the cost of folding or pre-organizing the linear peptide into a conformation conducive for disulfide bond formation. This will provide valuable information on the pre-organization step, which is a requisite for S—S bond formation in all disulfide-bridged peptide systems. Besides our primary calculation, which evaluates the free energy difference between the cyclic-like structure and the β-turn it converges to after 1 ns simulation, we also obtain free energy differences between the extended conformation and the final β-turn it converges to in unconstrained molecular dynamics, and between the two β-turn conformations.

In our simulations we use the recently developed conformational free energy thermodynamic integration (CFTI) protocol [22, 23]. In this method, MD simulations are performed for a series of intermediate structures between an initial and final conformer, with all flexible dihedrals in the system kept fixed at each point. In the CFTI approach, constraining all the soft degrees of freedom of the system has two goals. First, by restricting the conformational space explored by the flexible solute, fast convergence of the free energy is achieved, and conformational sampling problems are effectively overcome [22, 23]. Second, the method yields not only the derivative of the free energy along the chosen reaction path between the initial and final state, but also the full conformational free energy gradient: derivatives of the conformational free energy with respect to all the fixed coordinates at each path point. This provides more information about the free energy surface of the studied system at no increase in computational cost over standard thermodynamic integration approaches [22–24].

The results presented here are the first estimate of the pre-organization free energy and also the first application of the new CFTI protocol to a solvated system. The CFTI approach to accelerating convergence of thermodynamic averages through severely restricting available configurational space is philosophically different from that of other methods currently in use. For example, the locally enhanced sampling (LES) method introduces multiple copies of a selected subsystem and effectively increases the system flexibility [25], while the weighted histogram analysis method (WHAM) optimizes the statistical treatment of the available data in standard simulations [26, 27]. In applications presented here, the CFTI approach is similar to Monte Carlo simulations with fixed solute geometries [28] or exploration of one-dimensional reaction paths defined by simultaneous variation of several internal coordinates [29, 30]. The additional advantage of the CFTI method is the availability of the multidimensional free energy gradient. This has been used previously to perform free energy optimization and calculate second derivatives of the free energy in model systems [22, 23].In this work we use the free energy gradient to introduce a new type of free energy decomposition: conformational free energy site contributions, which provides additional insight into the molecular mechanism of the simulated processes.

2 Methods

The simulations focus on four conformers of the linear peptide DPDPE$(SH)_2$, denoted by Cyc, β_C, β_E and Ext, described in Table 1 and Fig. 1. These conformers have been generated in two previous unconstrained molecular dynamics simulations of 1 ns length, called cyl-DPDPE and lin-DPDPE in [21]. Simulation cyl-DPDPE started at the Cyc structure and reached the β_C conformer in its final stage. Cyc was obtained from a conformer of cyclic DPDPE (2′ of [11]) with the disulfide bond removed and geometry optimized with harmonic dihedral constraints. It was chosen to represent the structure adopted by the linear peptide prior to disulfide bond formation and is presumably the biologically active conformer. Simulation lin-DPDPE started in the Ext structure and reached β_E in its final stage. Ext was obtained by energy optimization of an arbitrary extended structure with all *trans* dihedrals restrained by with harmonic dihedral constraints. Since the two independent 1 ns simulations converged to similar conformers, β_C and β_E, we concluded that these type IV β-turns were stable, representative conformations of linear DPDPE in aqueous solution [21].

Three free energy simulations were performed, along paths connecting structures Cyc $\rightarrow \beta_C$, Ext $\rightarrow \beta_E$, and $\beta_C \rightarrow \beta_E$. Each simulation consisted of a series of steps; at each step an intermediate structure between the initial and final point was simulated for 60 ps (20 ps equilibration and 40 ps trajectory generation), using the CFTI protocol with all dihedrals needed to determine the conformation of the peptide kept fixed [22, 23].

Table 1. Dihedral values (degrees) for the four conformers of DPDPE$(SH)_2$

Dihedrals	Cyc	β_C	β_E	Ext
ϕ^2	111	128	118	168
ϕ^3	−100	−144	−164	178
ϕ^4	−73	−87	−100	−176
ϕ^5	84	127	108	168
ψ^1	165	90	143	−173
ψ^2	14	−145	−147	−180
ψ^3	−18	−36	84	178
ψ^4	−47	−44	−30	−174
χ_1^1	−163	−155	−179	−174
χ_1^2	177	174	−173	179
χ_1^4	−179	−84	−67	−175
χ_1^5	−74	−54	175	179
χ_2^1	51	60	73	76
χ_2^4	68	150	96	75

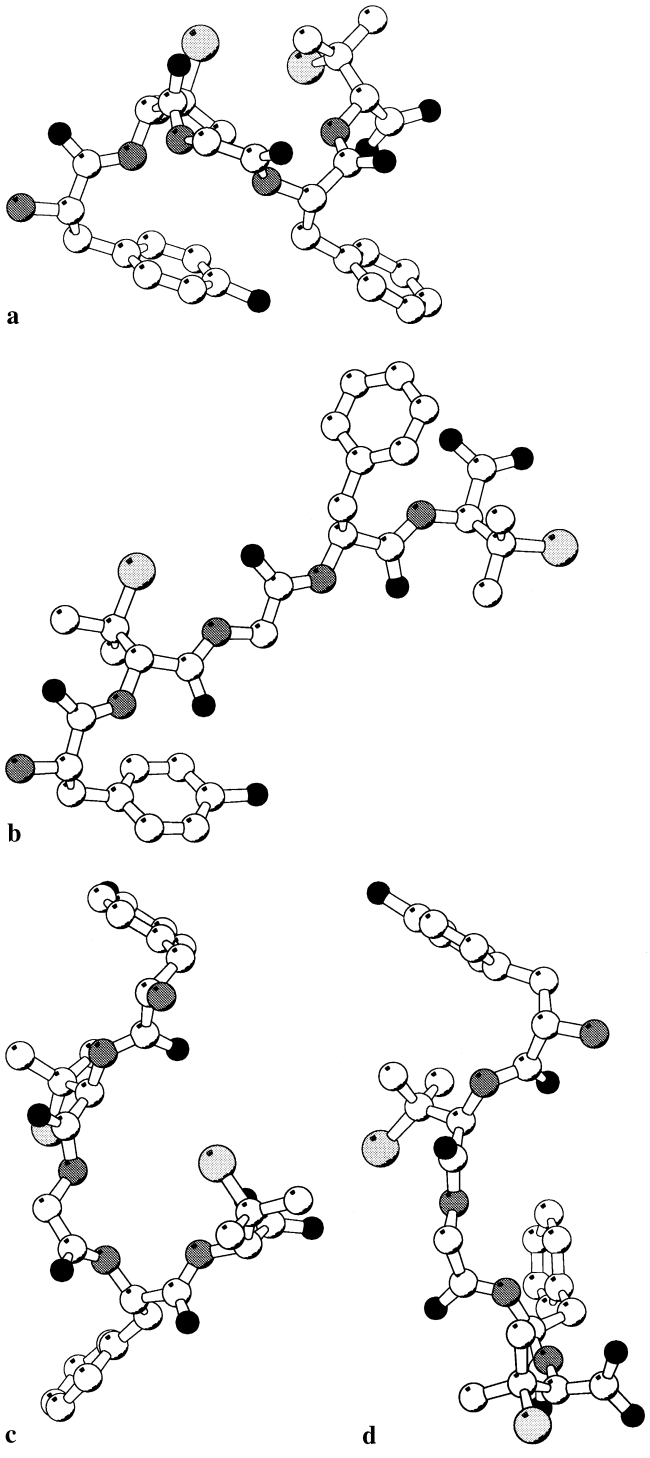

Fig. 1a–d. Studied structures of linear DPDPE in aqueous solution. **a** Cyc conformer. **b** Ext conformer. **c** β_C conformer. **d** β_E conformer. Atom coding: C, *white*; N, *dark gray*; O, *black*; S, *light gray*, shown with enlarged radius; H atoms not shown for clarity

The Cyc $\rightarrow \beta_C$ energy simulation consisted of 29 steps, with steps 1, 12, and 29 taken directly from the cyl-DPDPE trajectory [21]. Step 1 corresponded to the Cyc structure, i.e. the frame used to initiate the cyl-DPDPE trajectory, step 12 to the final frame of the equilibration phase of cyl-DPDPE, and step 29 to the β_C conformer (the final frame of cyl-DPDPE). Structures 2–11 were equally spaced along a straight line connecting 1 and 12 in dihedral angle space. Similarly, structures 13–28 were along a straight line connecting 12 and 29 in dihedral angle space. At steps 1, 12, and 19, trajectory frames from the cyl-DPDPE simulation containing a DPDPE$(SH)_2$ and 875 TIP3P waters box were directly used to perform free energy simulations. For the intermediate structures, the peptide was rebuilt in its new conformation in the solvent cavity of the last frame of the previous free energy step, followed by a brief energy minimization with fixed dihedrals.

The Ext $\rightarrow \beta_E$ free energy simulation involved 13 steps, with step 1 taken as the Ext conformer, i.e. the frame used to initiate the lin-DPDPE trajectory, and step 13 as β_E (the final frame of lin-DPDPE). The $\beta_C \rightarrow \beta_E$ simulation connected the β_C and β_E structures by 13 intermediate steps. In both Ext $\rightarrow \beta_E$ and $\beta_C \rightarrow \beta_E$ the intermediate structures were on a straight line in dihedral angle space connecting the respective initial and final states.

The free energy simulations were performed for a system consisting of linear DPDPE$(SH)_2$ and 875 water molecules in a body-centered cubic cell based on a cube of edge $a = 37.86$ Å, at constant volume and with periodic boundary conditions. At every simulation point the temperature was brought to 300 K by multiple velocity rescaling during the equilibration phase. The trajectory generation phase was then carried out at constant energy. Average temperatures from the different simulation points were between 294 and 305 K; temperature fluctuations were 5 K in each case. The simulation conditions thus approximately correspond to an (NVT) system at 300 K. All 14 dihedrals necessary to determine the conformation of DPDPE were kept fixed during the simulations. As listed in Table 1, these included all backbone ϕ and ψ dihedrals and most sidechain dihedrals; only dihedral angles corresponding to rotations of the peptide bonds, methyl groups, bonds involving hydrogen atoms, and termini were left unconstrained. The CHARMM 22 all-hydrogen parameter set was used in the simulations [31] with the same parameters for Pen as used in [21]. The Verlet algorithm with a time step of 2 fs was used in trajectory generation. SHAKE constraints were applied to all bonds involving hydrogen atoms. A nonbonded cutoff distance of 12.0 Åwas used, with van der Waals terms smoothly eliminated by a switching function between 10 and 12 Å and electrostatic interactions removed by shifting [32]. The truncation distance for the nonbond interaction lists was 14.0 Å. The lists were updated using the heuristic or "as needed" algorithm, which performs an update every time any atom moves by more than half the difference between the list cutoff and the nonbond cutoff, in our case $(14 - 12)/2 = 1$ Å, since previous list was built.

Using the CFTI protocol, we have calculated directly both the derivative of the free energy with respect to the reaction path $\partial A/\partial\lambda$:

$$\frac{\partial A}{\partial \lambda} = \left\langle \frac{\partial U}{\partial \lambda} \right\rangle = \left\langle \sum_{k=1}^{14} \frac{\partial U}{\partial \xi_k} \frac{\partial \xi_k}{\partial \lambda} \right\rangle \tag{1}$$

and the 14 individual derivatives $\partial A/\partial\xi_k$:

$$\frac{\partial A}{\partial \xi_k} = \left\langle \frac{\partial U}{\partial \xi_k} \right\rangle \quad k = 1, \ldots, 14 \tag{2}$$

with respect to all fixed coordinates along the path. The negligibly small generalized force terms involving the Jacobian were ignored [22, 23]. The free energy gradient along the path [Eq. (1)] and the individual free energy derivatives (Eq. 2) were integrated using linear interpolation of the derivative, to give overall free energy changes ΔA and individual dihedral contributions ΔA_k, $k = 1, \ldots, 14$ (Tables 2, 3). Since the volumes contained within the atomic van der Waals spheres of the initial and final peptide structures differed by less than 1 Å^3, we expect that for the studied processes the Helmholtz and Gibbs free energy changes should be similar.

Using the linear relationship between the potential energy and free energy change characteristic for all thermodynamic integration methods, we decompose the overall ΔA into contributions from peptide internal strain, nonbonded interactions within the peptide (solute-solute term), and interactions of the peptide with the solvent water (solute-solvent term). The internal strain component consists of bond stretching, Urey-Bradley, angle, proper,

Table 2. Conformational free energy simulation of linear DPDPE; changes in free energy and its components (kcal/mol)

Component	Cyc → β_C	Ext → β_E	$\beta_C \rightarrow \beta_E$
Free energy, ΔA	−4.0 ± 0.7	−28.2 ± 0.9	−2.3 ± 0.8
Entropy, $-T\Delta S$	−66.1 ± 31.2	10.4 ± 36.4	39.2 ± 33.6
Potential energy, ΔU	62.1 ± 31.2	−38.6 ± 36.4	−41.5 ± 33.7
	Free energy components		
Internal strain of solute	3.5 ± 0.1	−3.3 ± 0.1	−1.3 ± 0.1
Solute-solute interaction	−46.8 ± 0.1	−8.9 ± 0.1	−0.6 ± 0.1
Solute-solvent interaction	39.3 ± 0.9	−16.0 ± 1.2	−0.4 ± 1.1

Table 3. Individual dihedral site contributions to the overall free energy change

Dihedral	Cyc → β		Ext → β_E		$\beta_C \rightarrow \beta_E$	
	$\Delta\xi_k$ (°)	ΔA_k (kcal/mol)	$\Delta\xi_k$ (°)	ΔA_k (kcal/mol)	$\Delta\xi_k$ (°)	ΔA_k (kcal/mol)
ϕ^2	17	−0.4 ± 0.1	−50	−7.3 ± 0.2	−10	−0.4 ± 0.1
ϕ^3	−44	0.3 ± 0.1	18	−0.1 ± 0.1	−20	−0.1 ± 0.1
ϕ^4	−14	−0.4 ± 0.3	76	−3.4 ± 0.2	−13	−0.5 ± 0.1
ϕ^5	43	1.0 ± 0.2	−60	−6.8 ± 0.2	−19	−0.7 ± 0.1
ψ^1	−75	0.7 ± 0.2	−44	−2.8 ± 0.2	53	−0.8 ± 0.2
ψ^2	−159	−7.3 ± 0.6	33	−2.4 ± 0.1	−2	0.0 ± 0.1
ψ^3	−18	2.0 ± 0.4	−94	1.3 ± 0.2	120	1.3 ± 0.3
ψ^4	3	−0.5 ± 0.1	144	−1.3 ± 0.5	14	0.2 ± 0.1
χ_1^1	8	1.0 ± 0.1	−5	0.3 ± 0.1	−24	−0.5 ± 0.1
χ_1^2	−3	0.3 ± 0.1	8	0.7 ± 0.1	13	0.5 ± 0.1
χ_1^4	95	−1.9 ± 0.2	108	−5.7 ± 0.3	17	−2.1 ± 0.1
χ_1^5	20	−0.5 ± 0.1	−4	0.1 ± 0.1	−131	2.1 ± 0.3
χ_2^1	9	−0.2 ± 0.1	−3	−0.2 ± 0.1	13	0.1 ± 0.1
χ_2^4	82	1.9 ± 0.2	21	0.5 ± 0.1	−54	−1.5 ± 0.1

and improper dihedral deformation terms. The nonbonded interactions are a sum of van der Waals and electrostatic terms. The sum of the internal strain and solute-solute interaction terms encompasses all the energy components involving atoms of the solute only, and may be called the conformational strain contribution. The decompositions described above are nonrigorous, and do not yield path-independent thermodynamic functions. However, they are a useful tool in analyzing the molecular details of the overall effects.

Within each simulation step, statistical errors of the averages in Eqs. (1, 2) were obtained as standard deviations of the mean of sub-averages obtained by dividing the data into 20 contiguous blocks. The statistical uncertainties of the free energies were then obtained by applying standard error propagation to the formula for the numerical integral, which is effectively a trapezoidal rule (see above).

Our simulations are the first application of the recently developed CFTI algorithm to a solvated system [22, 23]. In the CFTI method, simulations with several coordinates kept fixed yield the derivatives of the free energy with respect to all coordinates in the set. This approach was shown to converge very quickly in vacuum simulations when all "soft" degrees of freedom were kept fixed [22, 23]. In the case of the peptide molecule in solution, the averaging is done both over the "hard" degrees of freedom of the peptide and over the solvent distribution around the constrained solute. The simulations performed here show that restricting the sampled configuration space by fixing the peptide conformation leads to quickly convergent averages also in the presence of solvent. From relatively short simulations (20 ps equilibration and 40 ps trajectory generation per step), we obtain reliable estimates of the free energy derivatives. For both $\partial A/\partial\lambda$ and $\partial A/\partial\xi_k$, absolute values were in the 0.1–20.0 kcal/(mol rad) range, while statistical errors were 0.3–1.5 kcal/(mol rad), with a positive correlation between the two quantities. This led to uncertainties of overall free energy change of below 1 kcal/mol (Table 2). In this initial application we made no attempt to lower uncertainties by optimizing the selection of path points or simulation lengths. Since the thermodynamic integration approach calculates the local values of the free energy derivatives along the reaction path, the path points (windows) may be quite widely spaced, and overlap between distributions sampled in consecutive windows is not necessary, unlike the case in umbrella sampling or thermodynamic perturbation calculations.

The following example further illustrates the power and usefulness of the CFTI approach in treating flexible systems in solution. To determine the starting and final structures of the free energy path we used individual trajectory frames, which are somewhat arbitrary but have reasonable internal geometry and solvation. To make sure that this did not introduce significant bias into our simulations, we used the free energy gradient calculated at the end points of the reaction path to estimate the free energy change involved in a transition from the structure in the chosen frame to that of the corresponding 20 ps MD average structure containing that frame: $\Delta A = \sum_k (\partial A/\partial\xi_k)\Delta\xi_k$, where $\Delta\xi_k$ is the change in dihedral k, $k = 1, \ldots, 14$. The calculated free energy changes were small, of the order of 0.1 kcal/mol, indicating that no large bias was introduced.

The calculations were performed using the program CHARMM version 22 [31, 32], modified to enable CFTI simulations as described in [22], on several IBM RS/6000 workstations at the Departments of Chemistry and Biochemistry of the University of Kansas, Lawrence, the Molecular Graphics and Modeling Laboratory and the Kansas Institute for Theoretical and Computational Science. One step of the free energy simulation protocol took about 3 days CPU time on an RS/6000-550.

3 Results and discussion

The linear DPDPE conformers considered here are described in Fig. 1 and Table 1. The free energy simulation results are presented in Fig. 2 and Tables 2 and 3.

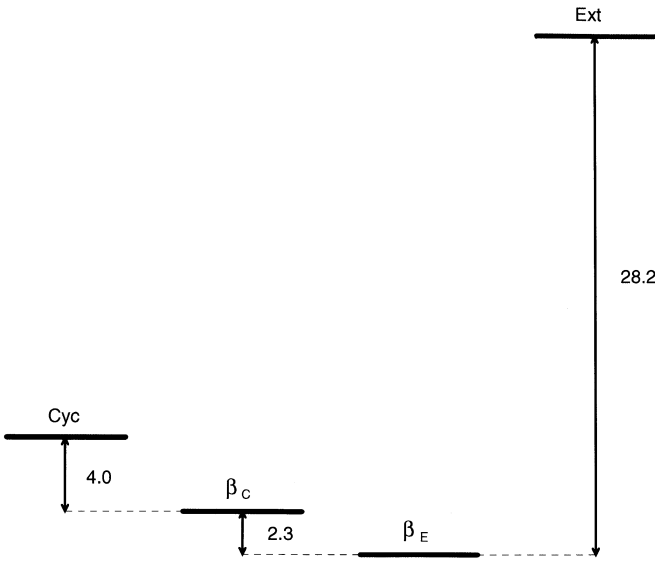

Fig. 2. Free energy diagram of the four studied conformers of linear DPDPE. Helmholtz free energy differences in kcal/mol

3.1 Overall results

The calculated overall free energy changes ΔA were -4.0 ± 0.8 kcal/mol for the Cyc $\rightarrow$ β_C "unfolding" transition, -28.2 ± 0.9 kcal/mol for the Ext $\rightarrow$ β_E transition, and -2.3 ± 0.8 kcal/mol for the $\beta_C \rightarrow \beta_E$ transition (Table 2). The most stable of the examined conformers is thus β_E, with β_C, Cyc, and Ext having free energies 2.3, 6.3, and 28.2 kcal/mol higher, respectively. The predicted conformer population ratios are thus 42 000:860:1 for β_E:β_C:Cyc; the population of the extended conformer should be negligible. The corresponding free energy level diagram is shown in Fig. 2.

The two β-turn structures, β_C and β_E, are the most stable among those considered. This is in accord with the unconstrained nanosecond simulations of linear DPDPE$(SH)_2$, which converged to these conformers [21]. The relatively low population of the cyclic-like structure also agrees qualitatively with the unconstrained simulation results [21]. The Cyc conformer represents the structure adopted by the linear peptide prior to disulfide bond formation, while the two β-turns are representative stable structures of linear DPDPE. The free energy differences of 4.0 kcal/mol between β_C and Cyc, and 6.3 kcal/mol between β_E and Cyc, reflect the cost of pre-organizing the linear peptide into a conformation conducive for disulfide bond formation. Such a conformational change is a pre-requisite for the chemical reaction of S—S bond formation to proceed.

The two β-turn conformers, β_C and β_E, are similar in Cartesian space, with a backbone rms deviation of 1.6 Å, but exhibit some significant differences in several dihedral angles (Table 1). These differences lead to a greater stability of the β_E structure. The Cyc $\rightarrow$ β_C and Ext $\rightarrow$ β_E free energy simulations connect the initial and final structures of our two previous unconstrained 1 ns linear DPDPE trajectories, cyl-DPDPE and lin-DPDPE, respectively [21]. Since both the Cyc $\rightarrow$ β_C and Ext $\rightarrow$ β_E free energy simulations result in a negative free energy change, this suggests that the underlying unconstrained trajectories describe nonequilibrium relaxation processes, with the peptides evolving from initial higher free energy (lower probability) states to final lower free energy (higher probability) structures.

It is not surprising that the extended conformation, which was just a guess used to generate a starting structure for long-time MD simulations [21], is highly unfavorable.

3.2 Energy-entropy decompositions

The energy-entropy decomposition results are presented in Table 2. Owing to significantly higher statistical errors, these results are much less reliable than the overall ΔA. It appears that the CFTI approach, while leading to quickly converging free energy values, does not provide improved energy-entropy results compared to other methods [33, 34]. In the Ext $\rightarrow$ β_E and $\beta_C \rightarrow \beta_E$ simulations, the energetic and entropic contributions are lower in magnitude or comparable to their respective errors, and no reliable conclusions can be drawn from these results.

In the Cyc $\rightarrow$ β_C simulation, entropic and energetic terms were $-T\Delta S = -66 \pm 31$ kcal/mol and $\Delta U = 62 \pm 31$ kcal/mol, respectively, suggesting that the β_C conformer is entropically favored over Cyc. The large entropy change may be rationalized in terms of the potential energy components: the decrease of intramolecular strain leads to more disorder in the peptide, while the less favorable interactions with the solvent lead to disorder in the water. The free energy components are discussed in more detail below.

3.3 Component analysis: solute and solvent

Using the linear relationship between the potential energy and free energy change characteristic for all thermodynamic integration methods, we decompose the overall ΔA into contributions from peptide internal deformations, nonbonded interactions within the peptide (solute-solute term), and interactions of the peptide with the solvent water (solute-solvent term). The results are given in Table 2.

In the Cyc $\rightarrow$ β_C simulation the contributions are: 3.5 kcal/mol from internal deformations, −46.8 kcal/mol from solute-solute, and 39.3 kcal/mol from solute-solvent interactions. Thus our results indicate that the preference of the linear peptide for the β-turn is due to significantly more favorable solute-solute interactions in that conformer. The sum of the solute-solute and internal strain terms, −43.3 kcal/mol, represents the conformational strain energy released after the peptide is allowed to relax from the cyclic-like to the β_C structure. This provides a quantitative measure for the intuitive concept of release of strain energy after removal of the S—S conformational constraint. Interactions with the

solvent water provide a compensating effect, preferentially stabilizing the cyclic-like structure. Strong solute-solvent interactions have been seen previously in standard MD simulations of cyclic DPDPE [13, 21], which exhibits parallel orientation of carbonyl groups, aggregation of hydrophobic groups, and an exceptionally high dipole moment. The two large effects of internal strain and solvation have opposite signs and mainly cancel, leading to a moderate conformational free energy difference between the cyclic-like and β_C structures. This appears to be a general property of solution thermodynamics of biological systems, where large contributions hidden in the overall free energy change are revealed by theoretical analysis [33].

In the $\beta_C \rightarrow \beta_E$ simulation, all free energy components were small and negative, with contributions of $-1.3, -0.6$, and -0.4 kcal/mol from internal strain, solute-solute, and solute-solvent interactions, respectively. This indicates that the main reason for the higher stability of the β_E structure is its lower conformational strain relative to β_C. The solvation of the two β conformers is the same within the statistical errors (see Table 2). Because of the small magnitude of the $\beta_C \rightarrow \beta_E$ free energy components, the sources of the stability difference between β_E and Cyc are essentially the same as between β_C and Cyc. Adding up the Cyc $\rightarrow \beta_C$ and $\beta_C \rightarrow \beta_E$ components shows that the the β_E structure has less favorable internal strain and solvation by 2.2 kcal/mol and 38.9 kcal/mol, respectively, while having more favorable intramolecular nonbonded interactions by 47.4 kcal/mol.

In the Ext $\rightarrow \beta_E$ simulation, all the free energy components were negative, with internal strain contributing -3.2 kcal/mol, solute-solute interactions -8.8 kcal/mol, and solute-solvent interactions -15.9 kcal/mol. Thus, the β_E conformer is favored over the arbitrary extended structure in all energetic aspects, having both lower conformational strain and better interactions with solvent.

3.4 Component analysis: individual dihedrals

Our conformational free energy simulation method, CFTI, yields the 14-dimensional free energy gradient along the reaction path, as in Eq. (2). Integrating each of the components $(\partial A/\partial \xi_k)$, $k = 1, \ldots, 14$, along the path enables the determination of the contribution of conformational change in each of the dihedrals to the overall free energy change. This leads to a novel free energy decomposition scheme, which may be called conformational site contribution analysis. As with all free energy decomposition schemes, this is not thermodynamically rigorous, but leads to useful insights into peptide structure, function, and possibly design of peptide and peptidomimetic drugs. The site contribution analysis results are presented in Table 3 .

In the case of the Cyc $\rightarrow \beta_C$ transition in linear DPDPE, the largest contribution to the overall free energy change, -7.3 kcal/mol, comes from ψ^2. This dihedral also undergoes the largest conformational shift. A conclusion from this decomposition result is that in order to stabilize the peptide in the Cyc conformation, a constraint at ψ^2 needs to be added. Such a constraint is indeed generated by the disulfide bond in cyclic DPDPE, in which our previous 1 ns simulations show ψ^2 remaining in the g^+ conformer [21]. The free energy decompositions suggest that other types of constraints, e.g. chemical modifications, might also be effective. Several other dihedral angles have site contributions of absolute magnitude of 1 kcal/mol or more: ψ^3, 2.0 kcal/mol; ϕ^5, 1.0 kcal/mol; χ_1^1, 1.0 kcal/mol; χ_1^4, -1.9 kcal/mol; and χ_2^4, 1.9 kcal/mol. In the case of the ϕ^5, χ_1^4, and χ_2^4 dihedrals the free energy contributions may be rationalized by their large shifts in the Cyc $\rightarrow \beta_C$ conformational transition. However, for ψ^3 and χ_1^1 the structural shifts are small, and the large contributions unexpected. As found in other free energy decomposition schemes, a number of contributions of opposing sign contribute to the overall observed effect [33].

All the backbone dihedrals except for ϕ^3 contribute significantly to the free energy change in the Ext $\rightarrow \beta_E$ simulation, with ϕ^2 and ϕ^5 having largest terms. Another major contributor is χ_1^4. In this transition, involving a large negative overall ΔA, most of the major components are also negative, with the 1.3 kcal/mol contribution from ψ^3 being the only exception. This analysis agrees with the solute-solute and solute-solvent decompositions described above, indicating that the β_E conformer is favorable relative to the extended one in all respects. The dihedrals with the largest site contributions (ϕ^2, ϕ^5, and χ_1^4) undergo large structural shifts in the Ext $\rightarrow \beta_E$ transition; however, two dihedral undergoing shifts close to 100°, ψ^3 and ψ^4, have only moderate contributions.

Most of the individual dihedral site contributions are small in the $\beta_C \rightarrow \beta_E$ simulation. The largest terms were for ψ^3, 1.3 kcal/mol; χ_1^4 -2.1 kcal/mol; and χ_2^4 -1.5 kcal/mol. These values largely reflect the structural shift in the $\beta_C \rightarrow \beta_E$ transition.

3.5 Correction for thermodynamic state

The free energy differences obtained from our constrained simulations refer to strictly specified states, defined by single points in the 14-dimensional dihedral space. Standard concepts of a molecular conformation include some region, or volume in that space, explored by thermal fluctuations around a transient equilibrium structure. To obtain the free energy differences between conformers of the unconstrained peptide, a correction for the thermodynamic state is needed. To obtain an estimate of the correction, one starts with the standard chemical definition of the free energy difference between conformers I and J: $\Delta A_{IJ} \approx \Delta G_{IJ} = -kT \ln P_J/P_I$, where P_I and P_J are populations, or probabilities of the two states. These probabilities may be expressed as integrals of the probability density ρ over volumes of conformational space assigned to the respective conformers. In terms of the quantities defined in the Methods section, the probability density is simply related to the free energy

surface: $\rho(\xi) = Z(\xi)/Z = \exp(-A(\xi)/kT)/Z$ where ξ is used to denote the whole collection of conformational coordinates $(\xi_1, \xi_2, \ldots, \xi_m)$. If one then assumes that (1) A is a multidimensional parabolic function in the vicinity of each conformer, e.g. for conformer I:

$$A(\xi) \approx A(\xi_0^{\mathrm{I}}) + \frac{1}{2}\sum_{k,l=1}^{m} H_{kl}^{\mathrm{I}}(\xi_0^{\mathrm{I}})(\xi_k - \xi_{0,k}^{\mathrm{I}})(\xi_l - \xi_{0,l}^{\mathrm{I}}) \qquad (3)$$

and (2) contributions close to the bottom of the parabola dominate in each of the volume integrals, then each volume integral may be approximated by the integral of the appropriate local parabolic expansion over all space. The population of a given conformer may then be approximated by

$$P_{\mathrm{I}} = \frac{\exp(-A(\xi_0^{\mathrm{I}})/kT)}{Z}\frac{(2\pi kT)^{m/2}}{|H_{\mathrm{I}}|^{1/2}} \qquad (4)$$

where $A(\xi_0^{\mathrm{I}})$ is the free energy at the bottom of the parabola describing conformer I and $|H_{\mathrm{I}}|$ is the determinant of the matrix of free energy second derivatives with respect to the m selected conformational coordinates, evaluated at this position. The formula for the free energy difference between conformers I and J then becomes

$$\begin{aligned}\Delta A_{\mathrm{IJ}} &= -kT\ln\frac{P_{\mathrm{J}}}{P_{\mathrm{I}}} \\ &= A(\xi_0^{\mathrm{J}}) - A(\xi_0^{\mathrm{I}}) + \frac{1}{2}kT\ln\frac{|H_{\mathrm{J}}|}{|H_{\mathrm{I}}|}\end{aligned} \qquad (5)$$

where the first two terms in the last equation represent the quantities calculated in our work (Table 2), i.e. free energy difference between two points in 14-dimensional space, and the final term is the desired correction due to thermal fluctuations.

The determinant $|H_{\mathrm{I}}|$ may be calculated from the well-known connection between the free energy second derivative matrix $\mathbf{H}$ and the covariance matrix of the selected degrees of freedom $\mathbf{C}{:}\mathbf{H} = kT\mathbf{C}^{-1}$, where $C_{ij} = \langle(\xi_i - \langle\xi_i\rangle)(\xi_j - \langle\xi_j\rangle)\rangle$ [35]. For each of the four selected conformers, three 20 ps simulations were performed with each of the dihedrals, which were kept fixed in the free energy simulation constrained in its initial value with a harmonic restraint potential $\frac{1}{2}k(\phi - \phi_0)^2$, with decreasing force constants $k = 3$, 2, and 1 kcal/(mol rad^2). The logarithms of the determinants $|C_{\mathrm{I}}|$ of the covariance matrices were calculated and extrapolated to $k = 0$ for each conformer I. The correction to the free energy difference between conformers I and J was then calculated as $(-1/2)kT\ln(|C_{\mathrm{J}}|/|C_{\mathrm{I}}|)$ at $k = 0$ [24]. The corrections obtained were -0.2 kcal/mol for Cyc $\rightarrow \beta_C$, 0.8 kcal/mol for Ext $\rightarrow \beta_E$, and 0.2 kcal/mol for $\beta_C \rightarrow \beta_E$, respectively. These corrections are comparable to the errors of the calculated free energy differences and do not affect the conclusions [24]. Thus it appears that the multidimensional free energy gradient obtained by fixing all flexible dihedrals in a pentapeptide may be integrated to obtain free energy differences between conformers without introducing significant bias.

4 Conclusions

We have used computer simulations to explore the free energy surface of the linear form of the opioid pentapeptide DPDPE in aqueous solution by calculating free energy differences between four conformers, denoted by Cyc, β_C, β_E, and Ext. The β-turn conformations, β_C and β_E, had been previously identified as stable, representative solution structures for linear DPDPE in unconstrained molecular dynamics simulations. The Cyc, or cyclic-like, conformer was similar to a proposed structure of the cyclic form of DPDPE, chosen to represent the structure adopted by the linear peptide prior to disulfide bond formation; the Ext conformer was an extended structure, with all *trans* dihedrals.

Our simulations indicate that β_E is the most stable of the studied conformers of linear DPDPE in aqueous solution, with β_C having 2.3 kcal/mol higher free energy. The cyclic-like Cyc conformer was less stable than both β-turns, with free energy 4.0 kcal/mol over β_C and 6.3 kcal/mol over β_E. Finally, the least stable extended structure had a free energy of 28.2 kcal/mol over β_E.

Our primary result is the determination of the free energy differences between the representative stable structures β_C and β_E and the cyclic-like conformer of linear DPDPE in aqueous solution. These free energy differences, 4.0 kcal/mol between β_C and Cyc, and 6.3 kcal/mol between β_E and Cyc, reflect the cost of preorganizing the linear peptide into a conformation conducive for disulfide bond formation. Such a conformational change is a pre-requisite for the chemical reaction of S—S bond formation to proceed. The predicted population ratio is 42 000:860:1 for β_E:β_C:Cyc. The relatively low population of the cyclic-like structure, which is presumably the biologically active conformer, agrees qualitatively with the observed lower potency and different receptor specificity of the linear form relative to the cyclic peptide. This low population is also in accord with previous unconstrained simulation results, from which such a population could not reliably be extracted owing to limited sampling.

Component analysis of the free energy difference between the β-turns and the cyclic-like structure indicate that the moderate stability difference of 4.0–6.3 kcal/mol is a result of the cancellation of two large opposing effects. In accord with intuition, the relaxed β-turns have conformational strain 43–45 kcal/mol lower than the Cyc structure. However, the cyclic-like conformer is able to interact much more strongly with water than the open β-turns, by ca. 39 kcal/mol. Such strong peptide-water interactions are in accord with previous simulations of the cyclic form of DPDPE.

Our simulations are the first application of the recently developed CFTI protocol to a solvated system [22,23]. In this method, MD simulations are performed for a series of intermediate structures between an initial and final conformer, with all flexible dihedrals in the system kept fixed at each point. This approach has two important advantages. First, because all "soft" degrees of freedom of the solute are fixed, the simulations do not suffer from conformational sampling problems common in other free energy simulation protocols, and all averages converge

very quickly. Second, the method yields not only the derivative of the free energy along the chosen reaction path between the initial and final state, but also the free energy gradients with respect to all the fixed coordinates at each path point. This provides more information about the free energy surface of the studied system at no increase in computational cost over standard thermodynamic integration approaches. The CFTI thermodynamic averages converged quite quickly, over tens of picoseconds. This is contrast to standard one-dimensional reaction path simulations, for example using the peptide end-to-end distance as the reaction coordinate, which generally require much longer sampling at each window. By averaging over the hard degrees of freedom of the solute and over the solvent distribution around the constrained solute, well defined thermodynamic states are generated at the ends and at the intermediate steps of the path, and the internal strain energy term is included in the free energy evaluation. Our results indicate that the CFTI method is a powerful, useful tool for simulating flexible molecules in solution. In our future work we plan to use the multidimensional free energy gradient to perform free energy optimization for such systems.

Acknowledgements. This work was supported in part by the Petroleum Research Fund of the American Chemical Society (grant 29566-G4), by a K*STAR EPSCoR FIRST Award, and by the University of Kansas General Research Fund. The authors wish to thank the Kansas Institute for Theoretical and Computational Science and Molecular Graphics and Modeling Laboratory at the University of Kansas for use of computer resources. Figure 1 was created using the program MolScript [36], and Fig. 2 using program ACE/gr [37].

References

1. Holmgren A (1985) Annu Rev Biochem 54:237–271
2. Creighton TE (1988) Bioessays 8:57–63
3. Thornton JM (1981) J Mol Biol 151:261–287
4. Creighton TE (1985) J Phys Chem 89:2452–2459
5. Nishiuchi Y, Sakakibara S (1982) FEBS Lett 148:260–262
6. Yanagisawa M, Kurihara H, Kimura S, Tomobe Y, Kobayashi M, Mitsui Y, Yazaki Y, Goto K, Masaki T (1988) Nature 332:441–415
7. Hruby VJ, Smith CW (1987) In: Udenfriend S, Meierhofer J (eds) The peptides: analysis, synthesis, biology. Academic Press, Orlando, FL, pp 77–197
8. Mosberg HI, Hurst R, Hruby VJ, Gee K, Yamamura HI, Galligan JJ Burks TF (1983) Proc Natl Acad Sci USA 80:5871–5874.
9. Froimowitz M, Hruby VJ (1989) Int J Pept Protein Res 34:86–96
10. Nikiforovich GV, Hruby VJ, Prakash O, Gehrig CA (1991) Biopolymers 31:941–955
11. Hruby VJ Kao L-F, Pettitt BM, Karplus M (1988) J Am Chem Soc 110:3351–3359
12. Pettitt BM, Matsunaga T, Al-Obeidi F, Gehrig C, Karplus M (1991) Biophys J 60:1540–1544
13. Smith PE, Dang LX, Pettitt BM, (1991) J Am Chem Soc 113:67–73
14. Smith PE, Pettitt BM (1992) Biopolymers 32:1623–1629
15. Wilkes BC, Schiller PW (1991) J Comput Aided Mol Des 5:293–302
16. Chew C, Villar HO, Loew GH (1991) Mol Pharmacol 39:502–510
17. Smith PE, Marlow GE, Pettitt BM (1993) J Am Chem Soc 115:7493–7498
18. Flippen-Anderson JL, Hruby VJ, Collins N, George C, Cudney B (1994) J Am Chem Soc 116:7523–7531
19. Mosberg HI, Sobczyk-Kojiro K, Subramanian P, Crippen GM, Ramalingam K, Woodard RW (1990) J Am Chem Soc 112: 822–829
20. Chuang L-C, Chen S-T, Ya C (1993) Biochim Biophys Acta 1158: 209–216
21. Wang Y, Kuczera K (1996) J Phys Chem 100:2555–2563
22. Kuczera K (1996) J Comput Chem 17:1726–1749
23. Wang Y, Kuczera K (1997) J Phys Chem B 101:5205–5213
24. Wang Y (1997) Computational approach to the influence of the disulfide bond on peptide properties. PhD thesis, University of Kansas
25. Verkhivker G, Elber R, Nowak W (1992) J Phys Chem 97:7838–7841
26. Kumar S, Bouzida D, Swendsen RH, Kollman PA Rosenberg JM (1992) J Comput Chem 13:1011–1021
27. Roux B (1995) Comput Phys Commun 91:275–282
28. Tirado-Rives J, Maxwell DS, Jorgensen WL (1993) J Am Chem Soc 115:11590–11593
29. Tobias DJ, Brooks C.L.B., III (1992) J Phys Chem 96:3864–3870
30. Elber R (1990) J Chem Phys 93:4312–4321
31. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-NcCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Karplus M (1998) J Phys Chem B 102:3386–3616
32. Brooks BR, Bruccoleri R, Olafson B, States D, Swaminathan S, Karplus M (1983) J Comput Chem 4:187–217
33. Gao J, Kuczera K, Tidor B, Karplus M (1989) Science 244:1069–1072
34. Kuczera K, Gao J, Tidor B, Karplus M (1990) Proc Natl Acad Sci USA 87:8481–8485
35. Brooks CL III, Karplus M, Petitt BM (1988) Proteins: a theoretical perspective of dynamics, structure, and thermodynamics. Wiley, New York
36. Kraulis PJ (1991) J Appl Crystallogr 24:946–950
37. Turner PJ ACE/gr program, available from ftp.ccalmr. ogi.edu