Quantifying the Intrinsic Effects of Two Point Mutation Models of Proline−Proline Diamino Acid Diamide: A First-Principle Computational Study

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · OCTOBER 2007

Impact Factor: 3.3 · DOI: 10.1021/jp073471h · Source: PubMed

Quantifying the Intrinsic Effects of Two Point Mutation Models of Proline−Proline Diamino Acid Diamide: A First-Principle Computational Study

Michelle A. Sahai,*,†,‡ Bela Viskolcz,§ Emil F. Pai,†,‡,⊥ and Imre G. Csizmadia*,§,#

Ontario Cancer Institute, Division of Cancer Genomics and Proteomics, MaRS Center, Toronto Medical Discovery Tower, 101 College Street, Room 5-359, Toronto, Ontario, M5G 1L7, Canada, Department of Medical Biophysics, University of Toronto, Toronto, Ontario, M5G 2M9, Canada, Department of Chemistry and Chemical Informatics, Faculty of Education, University of Szeged, Boldogasszony sgt. 6, Szeged, H-6725, Hungary, Department of Biochemistry and Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, M5S 1A8, Canada, and Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, Ontario, M5S 3H6, Canada

Received: May 7, 2007; In Final Form: July 16, 2007

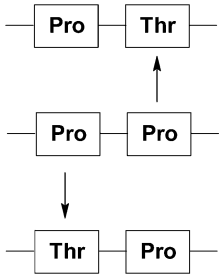
Two sites of a Pro-Pro diamide were subjected to individual Pro → Thr point mutations. The parent diamide Pro-Pro as well as selected conformers of the Pro-Thr and Thr-Pro mutant models were subjected to molecular computations at the B3LYP/6-31G(d) level of theory. At the optimized geometries, thermodynamic functions (*S*, *H*, and *G*) were computed. In order to assess relative stabilities of the mutant models, isodesmic reactions were constructed to calculate ΔS, ΔH, and ΔG, relative to the initial Pro-Pro state. The importance of intramolecular hydrogen bonds, involving the −OH group of the Thr side chain, which emerged after the point mutations were also examined. Our findings suggest a novel approach to analyzing the stability of point mutants in peptide models through the analysis of thermodynamic functions.

1. Introduction

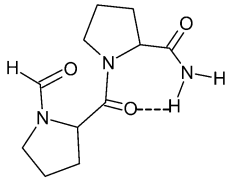
Many hereditary diseases are caused by mutations in the primary structure of some proteins. The textbook example, which can be traced back to the middle of the 20th century and can now be described by structural information, is sickle cell anemia in which glutamic acid, at position 6, is replaced by valine.[1] A more recent example is the observation in *congenital cataract*. In this case, blindness is caused by mutations in galactokinase[2,3] where Pro28 is replaced by Thr in the mutation process. This paper focuses on the geometrical and energetic consequences of such a Pro → Thr mutation in a simple diamino acid model (Scheme 1).

Clearly, the position in the sequence that leads to different isomers will influence the stability of the different mutant models and can be measured by the change in the various thermodynamic functions. In addition to the topological position of the point mutation, the side chain of the incoming new amino acid may be involved in different intramolecular interactions such as hydrogen bonding. Both of these effects may be determined by first-principle molecular computations. It may therefore be conceded that this will lead to a novel approach to analyzing the stability of point mutants in peptide models.

SCHEME 1: The Pro-Pro Dipeptide and its Thr Single-Point Mutants



SCHEME 2: Backbone Hydrogen Bonding Involving the Terminal NH_2 Group in HCO−Pro-Pro−NH_2



2. Method

2.1. Conformational and Configurational Specifications. Numeric definitions of the relative spatial orientation of all constituent atoms of HCO−Pro-Pro−NH_2, HCO−Pro-Thr−NH_2, and HCO−Thr-Pro−NH_2 follow an established standard,[4,5] shown explicitly in Figure A in the Supporting Information. As a result, amino acid residues in the dipeptides, as well as the protecting end groups, were exclusively defined using the *z*-matrix internal coordinate system to characterize molecular structure, geometry, and stereochemistry. Both trans and cis peptide bonds were considered in the case of HCO−Pro-Pro−NH_2 leading to a total of four geometrical isomers:[1]

cis−cis	cis−trans	
trans−cis	trans−trans	(1)

* Corresponding authors. Tel.: (01) 416 581 7548 (M.A.S. and I.G.C.). Fax: (01) 416 581 7546 (M.A.S. and I.G.C.). E-mail: michelle.sahai@utoronto.ca (M.A.S.); imre.csizmadia@utoronto.ca (I.G.C.).

† Ontario Cancer Institute.

‡ Department of Medical Biophysics, University of Toronto.

§ University of Szeged.

⊥ Department of Biochemistry and Molecular and Medical Genetics, University of Toronto.

Department of Chemistry, University of Toronto.

10.1021/jp073471h CCC: $37.00
Published on Web 09/08/2007

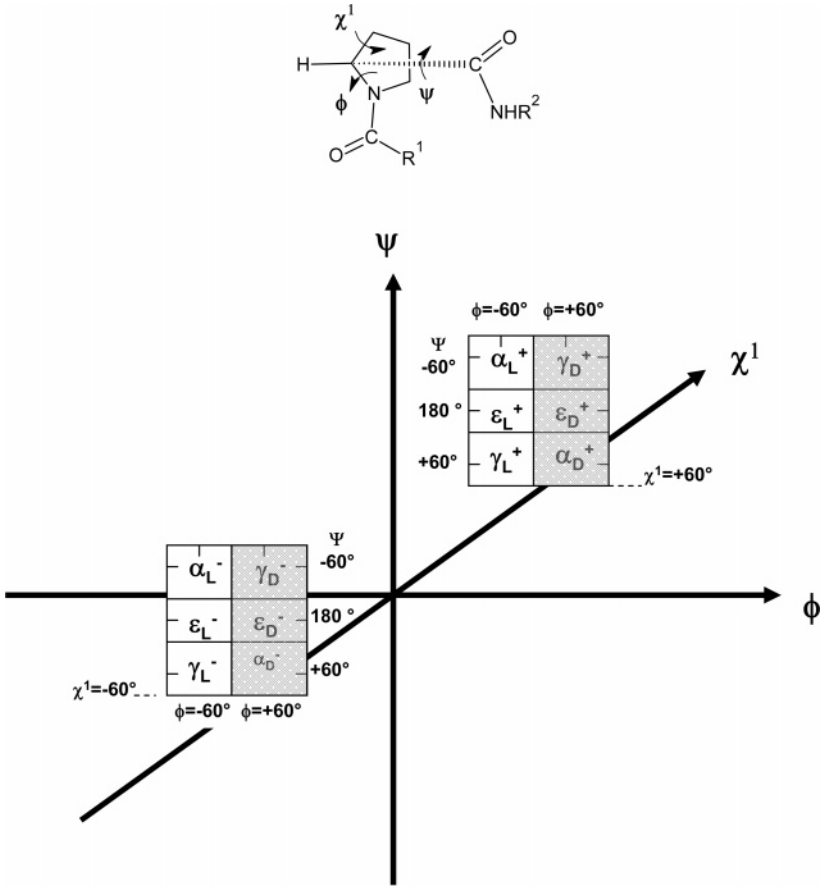


Figure 1. Conformational space of L-proline diamide.

The conformers to be studied in detail were specifically chosen from the Protein Data Bank (PDB).

2.2. Selecting Geometry Parameters from the Brookhaven PDB. A total of 2485 nonhomologous polypeptide chains were retrieved from the PDB SELECT[6−8] (October 2004 update) whose structures have been determined by solution NMR and X-ray crystallography. These polypeptide chains contained a total of 392 257 amino acid residues. Single polypeptide chains were then chosen from the X-ray data, since the results of NMR analysis constitute an ensemble of alternative models, in contrast to the unique model obtained by X-crystallography. From this data set only those chains with X-ray structure resolution better than 3.0 Å were chosen since they resolved more atomic positions with greater certainty. Atomic coordinates of the residues in the representative chains were then converted to torsional angle values. Computed torsional angles (ω, ϕ, ψ, χ^1, and χ^2) for the chosen dipeptides were derived from the retrieved angles (Tables A−C in the Supporting Information) from proteins from the PDB to reduce the MDCA[9] search. In searching for cis peptide bonds, ω was limited within a threshold value of ±30°.

2.3. Molecular Computations of Structures and Energies. All computations were carried out using the Gaussian 03[10] program package (G03). Each structure was initially optimized using the ab initio[11] restricted Hartree−Fock (RHF)[12] method with the split valence 3-21G basis set.[13−15] Multidimensional conformational analysis (MDCA)[9] was used to define the topologically possible set of conformers represented by a grid-defined set of catchment regions (Figures 1 and 2). Presently, it is possible to accurately characterize the topologically probable set of stable conformers emerging from the larger set of topologically possible conformers.[16]

The RHF/3-21G geometry optimized structural parameters were then used as the input in a subsequent theoretical refinement step with the inclusion of electron correlation effects at the B3LYP/6-31G(d) level of theory to obtain more reliable geometry and stability data. Here, B3LYP[17] denotes the combination of Becke's three-parameter exchange functional with the Lee−Yang−Parr (LYP)[18] correlation functional and also employs the mathematically more complete 6-31G(d) basis set. Energies of this type are labeled as $E^{\text{uncorrected}}$. Total energies are given in hartrees, and the relative energies are given in kilocalories per mole (with the conversion factor: 1 hartree = 627.5095 kcal·mol^{-1}).

Additionally, each stable conformer was subjected to frequency calculations at the B3LYP/6-31G(d) level of theory in order to confirm their identity as being true minima. The results also provided zero-point energy (ZPE) values, which were scaled using a correction factor and added to the total energy of each conformer to provide more accurate energetic characterization of the conformers as well as the vibrational frequency of each of the normal modes. Corrected energies for these geometries are labeled as $E^{\text{corrected}}$.

2.4. Potential Energy Surfaces. Cross sections of the potential energy surface (PES) for Pro-Thr and Thr-Pro were made at the RHF/3-21G level of theory by varying the ψ dihedral angle for the Pro residue in 30° increments. Simultaneously, the ψ dihedral angle for the Thr residue, in each of these dipeptides, was kept fixed at 150°. For the Pro-Pro diamide two similar cross sections were generated, where at any particular instant one Pro ψ dihedral angle was varied and the other kept frozen. In addition to the cross sections for Pro-Pro, a PES of the type $E = f(\psi_1,\psi_2)$ was generated and presented as surface and contour diagrams in the range of −360° to +360°.

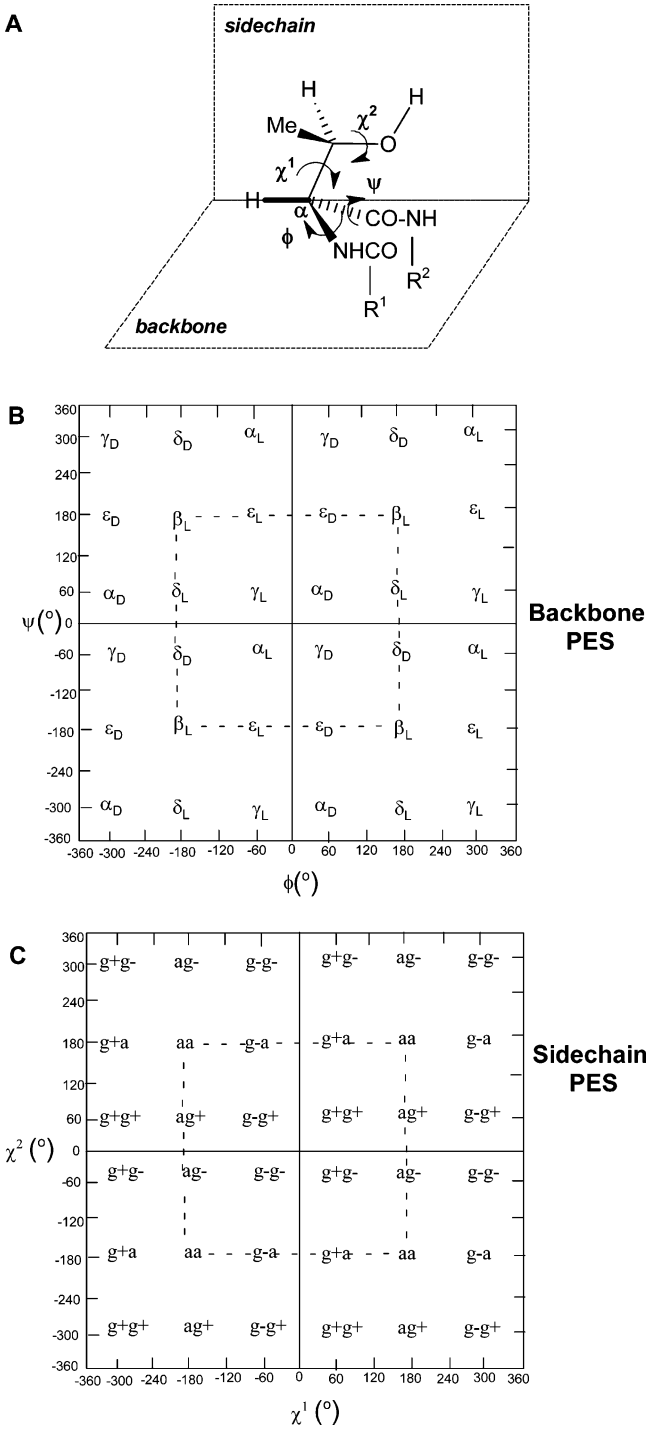


Figure 2. Conformational space of L-threonine diamide. (A) Partitioning the four independent variables (ϕ, ψ, χ^1, χ^2) to backbone (ϕ and ψ) and side-chain (χ^1 and χ^2) domains. Unless otherwise stated, $R^1 = R^2 = H$ in the present study. (B) A schematic illustration of the topology of the Ramachandran backbone PES. (C) A schematic illustration of the topology of the side-chain PES. For both, $\chi^1 = \chi^2 = 180°$.

2.5. Thermodynamic Functions (ΔH, ΔG, and ΔS) and Hydrogen-Bonding Interactions. Besides the relative energy ($\Delta E^{\text{uncorrected}}$ and $\Delta E^{\text{corrected}}$) values thermodynamic functional changes of enthalpy (ΔH), entropy (ΔS), and Gibbs free energy (ΔG) were also calculated for the stable conformations. The values of total enthalpy (H) and Gibbs free energy (G) given in hartrees were converted to their respective relative values, ΔH and ΔG, in kcal·mol^{-1} as described above, whereas the values for entropy (S) and entropy change (ΔS) were given in cal·mol^{-1}·K^{-1} units.

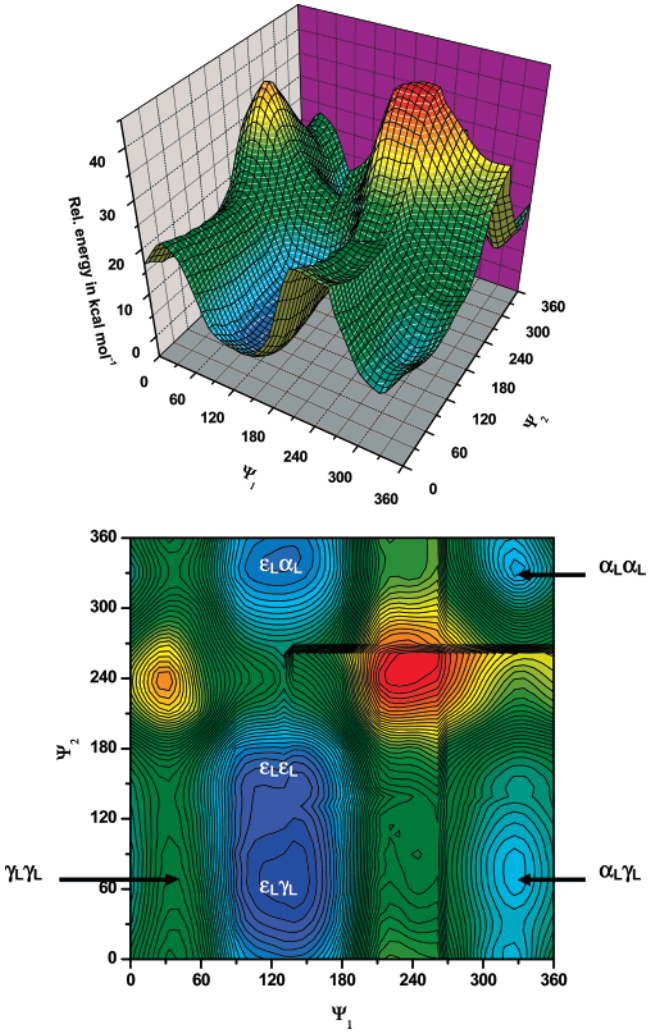


Figure 3. PES $E = f(\psi_1,\psi_2)$ at $\chi^1 = \chi^2 = 30.0°$ for tt HCO−Pro-Pro−NH_2 computed at the RHF/3-21G level of theory. Top: spline representation. Bottom: contour map showing the six minima ($\gamma_L\gamma_L$, $\epsilon_L\gamma_L$, $\alpha_L\gamma_L$, $\epsilon_L\epsilon_L$, $\epsilon_L\alpha_L$, $\alpha_L\alpha_L$).

3. Results and Discussion

3.1. Molecular Structures and Energetics. *3.1.1. Pro-Pro Dipeptide.* A PES of the type $E = f(\psi_1,\psi_2)$ was generated for trans−trans HCO−Pro-Pro−NH_2 at the RHF/3-21G level of theory (Figure 3). Six minima, $\gamma_L\gamma_L$, $\epsilon_L\gamma_L$, $\alpha_L\gamma_L$, $\epsilon_L\epsilon_L$, $\epsilon_L\alpha_L$, $\alpha_L\alpha_L$, appeared on the PES even though on the basis of the one-dimensional (1D) cross sections (Figure 4, parts a and b) one would have predicted a $3 \times 3 = 9$ minima. These six minima, suggested by the PES (Figure 4) of trans−trans HCO−Pro-Pro−NH_2 were also located by direct optimization at the RHF/3-21G level of theory (Figure B of the Supporting Information). However, the density functional theory (DFT) computations revealed only two minima $\epsilon_L\gamma_L$, $\alpha_L\alpha_L$ (Figure B of the Supporting Information). For the other geometrical isomers of Pro-Pro, with the exception of cis−cis, which revealed only the $\epsilon_L\alpha_L$ at the DFT level, the cis−trans and trans−cis isomers also contained the $\epsilon_L\gamma_L$ at the DFT level of theory. This later structure ($\epsilon_L[+]\gamma_L[+]$) was kept as the reference conformer throughout this work. It is interesting to note that $\epsilon_L\gamma_L$ is preferred over the $\epsilon_L\epsilon_L$ conformer because the

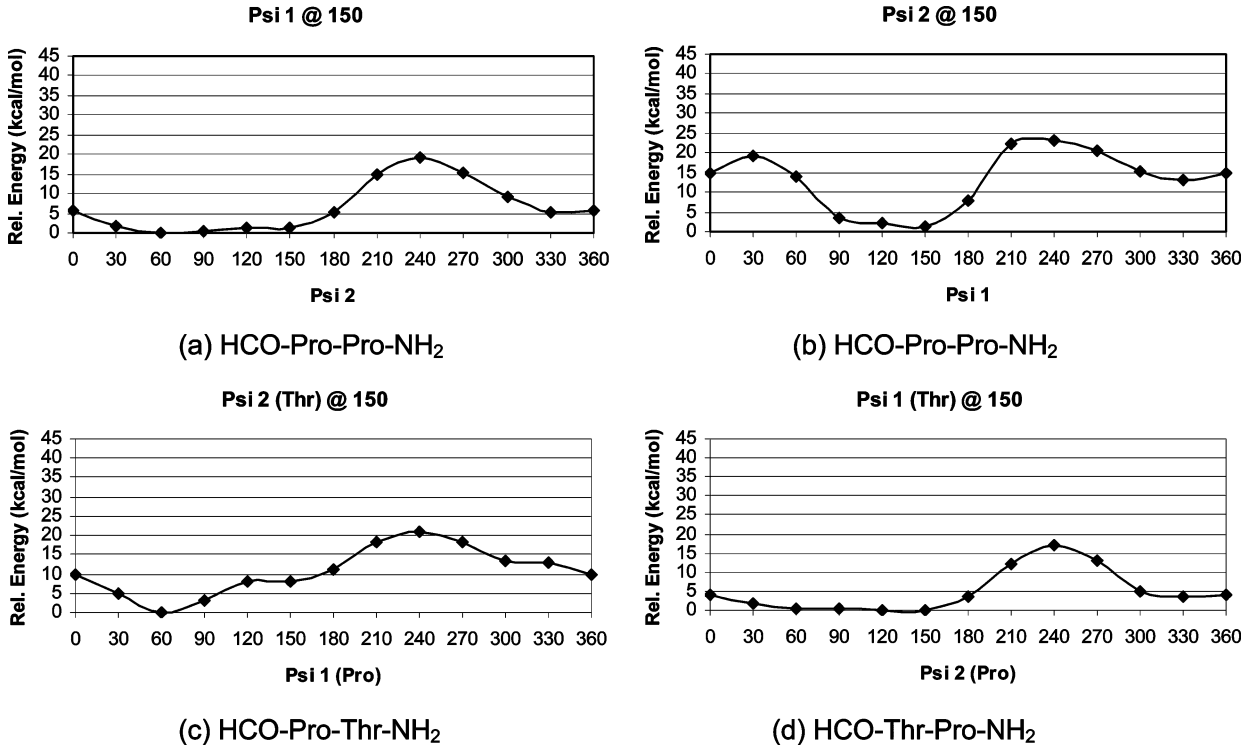
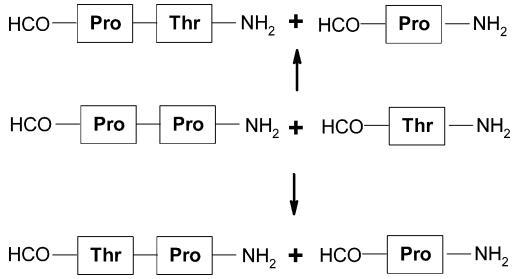
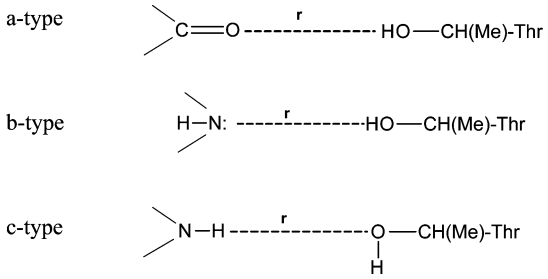


Figure 4. Cross sections of the PEHS for the three dipeptides computed at the RHF/3-21G level of theory. (a−c) $E = f(\psi_1)$ and variable ψ_2 computed; (d) PEHS of HCO−Pro-Pro−NH_2 where $E = f(\psi_2)$ and variable ψ_1 computed.

SCHEME 3: Schematic Representation of the Isodesmic Reactions Involving the Three Corresponding Dipeptides



SCHEME 4: Three Types of Hydrogen Bonds and Their Hydrogen-Bond Lengths (r) Occurring in the Pro-Thr and Thr-Pro Mutants



terminal NH_2 is the only proton donor to form a hydrogen bond. This is manifested in a C_7 or γ_L conformation of the C-terminal amino acid (Scheme 2).

The RHF/3-21G results are listed in Table D, and the B3LYP/6-31G(d) results are summarized in Table E of the Supporting Information, for all the isomers studied. The B3LYP/6-31G(d) optimized HCO−Pro-Pro−NH_2 structures are shown in Figure C of the Supporting Information.

3.1.2. Pro-Thr and Thr-Pro Mutant Models. The structures of HCO−Pro-Thr−NH_2 and HCO−Thr-Pro−NH_2 optimized at the RHF/3-21G level of theory are summarized in Tables F and

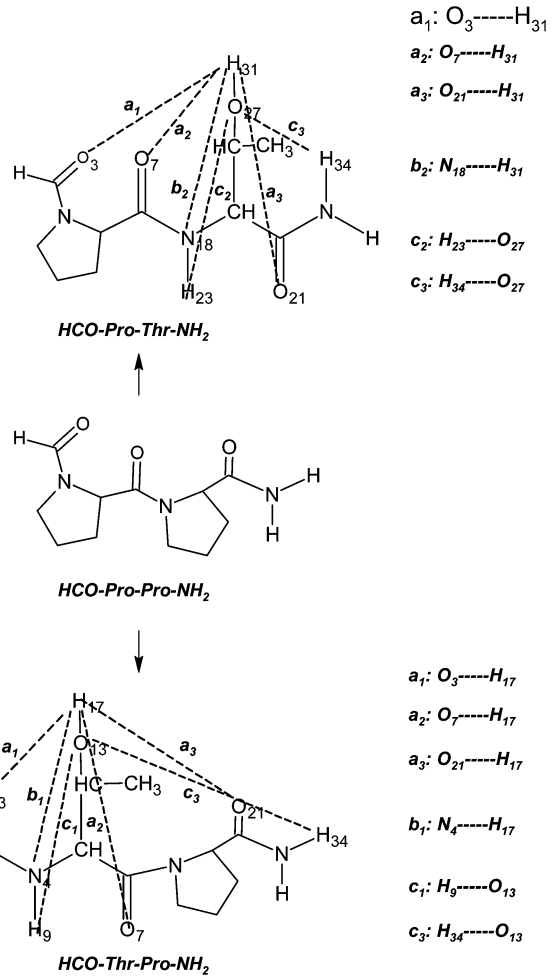


Figure 5. Side-chain−backbone hydrogen-bonding networks available with the two HCO−Pro-Pro−NH_2 dipeptides mutants: HCO−Pro-Thr−NH_2 and HCO−Thr-Pro−NH_2. The hydrogen-bond designation and atoms involved in hydrogen bonding are indicated for each mutant.

TABLE 1: Selected Parameters for the HCO−Pro-Pro−NH_2 + HCO−Thr−NH_2 Initial State[a]

	B3LYP/6-31G(d)					
	enthalpy		entropy		Gibbs free energy	
conformers	total	relative	total	relative	total	relative
tt α_L [+] α_L [+]	−1350.093706	6.81	227.34	−1.17	−1350.202367	7.16
tt ϵ_L [+] α_D [−]	−1350.083721	13.08	228.14	−0.38	−1350.192758	13.19
tt ϵ_L [+] γ_L [+]	−1350.104549	0.01	227.64	−0.88	−1350.213349	0.27
tt ϵ_L [−] γ_L [+]	−1350.104561	0.00	228.51	0.00	−1350.213777	0.00

[a] Global minima of HCO−Thr−NH_2 = *tt* γ_L[− +]: $\Delta H = -531.030809$, $\Delta S = 98.65$, $\Delta G = -531.078324$.

TABLE 2: Selected Parameters for the HCO−Pro-Thr−NH_2 + HCO−Pro−NH_2 Final Mutated State[a]

		B3LYP/6-31G(d)						H-bonding								
		enthalpy		entropy		Gibbs free energy		distances (Å)								
	conformers	total	relative	total	relative	total	relative	a_1	a_2	a_3	b_1	b_2	c_1	c_2	c_3	$\sum\rho_b$
1	*tt* α_L [−] α_D [a −]	−1350.087428	10.75	231.95	3.44	−1350.198210	9.77							2.33		0.0102
2	*tt* α_L [−] α_L [a −]	−1350.101319	2.03	230.22	1.71	−1350.211278	1.57							2.25		0.0123
3	*tt* α_L [−] β_L [+ +]	−1350.103765	0.50	230.67	2.16	−1350.213935	−0.10								2.04	0.0203
4	*tt* α_L [−] ϵ_D [− −]	−1350.104759	−0.12	228.57	0.05	−1350.213930	−0.10		1.79						1.94	0.0626
5	*tt* α_L [−] γ_L [− a]	−1350.104444	0.07	230.99	2.47	−1350.214766	−0.62							2.34		0.0100
6	*tt* α_L [−] γ_L [a +]	−1350.097959	4.14	231.75	3.24	−1350.208642	3.22					2.28				0.0115
7	*tt* α_L [−] γ_L [− +]	−1350.107951	−2.13	229.89	1.38	−1350.217750	−2.49			2.20						0.0139
8	*tt* α_L [−] γ_L [+ −]	−1350.107308	−1.72	228.52	0.00	−1350.216456	−1.68			1.90						0.0284
9	*tt* ϵ_L [−] α_D [− +]	−1350.101890	1.68	227.94	−0.58	−1350.210762	1.89			2.16						0.0153
10	*tt* ϵ_L [−] α_L [a −]	−1350.102608	1.23	231.11	2.60	−1350.212987	0.50	1.97						2.02		0.0453
11	*tt* ϵ_L [−] γ_L [− −]	−1350.103621	0.59	227.60	−0.91	−1350.212336	0.90		1.82							0.0343
12	*tt* γ_L [−] α_D [− +]	−1350.101908	1.66	228.57	0.06	−1350.211082	1.69			2.15						0.0156
13	*tt* γ_L [−] β_L [a +]	−1350.106710	−1.35	229.49	0.98	−1350.216321	−1.60	1.94								0.0258
14	*tt* γ_L [−] β_L [+ +]	−1350.106707	−1.35	229.47	0.95	−1350.216306	−1.59								2.05	0.0198
15	*tt* γ_L [−] δ_L [− −]	−1350.105376	−0.51	229.78	1.26	−1350.215123	−0.84	2.07						2.37		0.0282
16	*tt* γ_L [−] δ_L [+ −]	−1350.108436	−2.43	229.16	0.65	−1350.217891	−2.58			1.93						0.0264
17	*tt* γ_L [−] γ_L [a +]	−1350.107936	−2.12	226.08	−2.43	−1350.215928	−1.35	2.09								0.0180
18	*tt* γ_L [−] γ_L [− +]	−1350.113711	−5.74	225.70	−2.81	−1350.221522	−4.86			1.96				2.53		0.0309

[a] Global minima of HCO−Pro-Pro−NH_2 = *tt* ϵ_L[−]γ_L[+]: $\Delta H = -1350.104561$, $\Delta S = 228.51$, $\Delta G = -1350.213777$. Global minima of HCO−Pro−NH_2 = *tt* γ_L[+]: $\Delta H = -493.918627$, $\Delta S = 94.16$, $\Delta G = -493.963938$.

TABLE 3: Selected Parameters for the HCO−Thr-Pro−NH_2 + HCO−Pro−NH_2 Final Mutated State[a]

		B3LYP/6-31G(d)						H-bonding								
		enthalpy		entropy		Gibbs free energy		distances (Å)								
	conformers	total	relative	total	relative	total	relative	a_1	a_2	a_3	b_1	b_2	c_1	c_2	c_3	$\sum\rho_b$
1	*tt* α_D [− +] γ_L [−]	−1350.097615	4.36	227.80	−0.71	−1350.206424	4.61		1.98							0.0234
2	*tt* α_D [+ −] γ_L [−]	−1350.093584	6.89	226.97	−1.54	−1350.201997	7.39		1.97							0.0240
3	*tt* α_L [a −] α_L [−]	−1350.095255	5.84	227.87	−0.65	−1350.204094	6.08						2.16			0.0153
4	*tt* α_L [a −] γ_L [+]	−1350.095599	5.62	228.85	0.34	−1350.204907	5.57							2.16		0.0153
5	*tt* β_L [+ +] α_L [−]	−1350.098064	4.08	230.18	1.67	−1350.208002	3.62								2.18	0.0146
6	*tt* β_L [− +] γ_L [−]	−1350.099756	3.02	229.85	1.33	−1350.209536	2.66									
7	*tt* β_L [+−] γ_L [+]	−1350.103269	0.81	229.84	1.32	−1350.213044	0.46	2.05								0.0198
8	*tt* β_L [+ +] γ_L [+]	−1350.102695	1.17	228.45	−0.07	−1350.211810	1.23									
9	*tt* β_L [+ +] γ_L [−]	−1350.100730	2.40	229.08	0.57	−1350.210147	2.28									
10	*tt* β_L [a +] γ_L [−]	−1350.102699	1.17	227.63	−0.88	−1350.211426	1.48	1.86								0.0312
11	*tt* δ_D [− −] ϵ_L [+]	−1350.091473	8.21	226.42	−2.09	−1350.199627	8.88	1.94								0.0258
12	*tt* δ_D [− −] γ_L [−]	−1350.096903	4.81	225.62	−2.89	−1350.204676	5.71	1.84								0.0327
13	*tt* δ_L [+ −] α_L [−]	−1350.095579	5.64	229.62	1.11	−1350.205251	5.35		2.27						2.41	0.0202
14	*tt* δ_L [+ −] γ_L [+]	−1350.099667	3.07	227.33	−1.18	−1350.208254	3.47		2.02							0.0213
15	*tt* ϵ_D [− −] α_L [−]	−1350.099465	3.20	224.76	−3.75	−1350.206829	4.36	1.78							2.08	0.0562
16	*tt* ϵ_L [a +] α_L [−]	−1350.093769	6.77	231.30	2.79	−1350.204240	5.98				2.36					0.0095
17	*tt* ϵ_L [a +] γ_L [−]	−1350.100606	2.48	230.17	1.66	−1350.210542	2.03				2.35					0.0097
18	*tt* γ_L [− +] α_L [−]	−1350.108879	−2.71	227.69	−0.83	−1350.217633	−2.42		1.91				2.28			0.0392
19	*tt* γ_L [− +] γ_L [+]	−1350.109960	−3.39	227.86	−0.65	−1350.218796	−3.15		1.92				2.27			0.0388
20	*tt* γ_L [− +] γ_L [−]	−1350.108880	−2.71	227.68	−0.84	−1350.217629	−2.42		1.92				2.27			0.0388

[a] Global minima of HCO−Pro-Pro−NH_2 = *tt* ϵ_L[−]γ_L[+]: $\Delta H = -1350.104561$, $\Delta S = 228.51$, $\Delta G = -1350.213777$. Global minima of HCO−Pro−NH_2 = *tt* γ_L[+]: $\Delta H = -493.918627$, $\Delta S = 94.16$, $\Delta G = -493.963938$.

G of the Supporting Information, respectively. The B3LYP/6-31G(d) results are listed in the Tables H and I of the Supporting Information, respectively. The calculated 1D cross sections for Pro-Thr and Thr-Pro are represented in Figure 4, parts c and d,

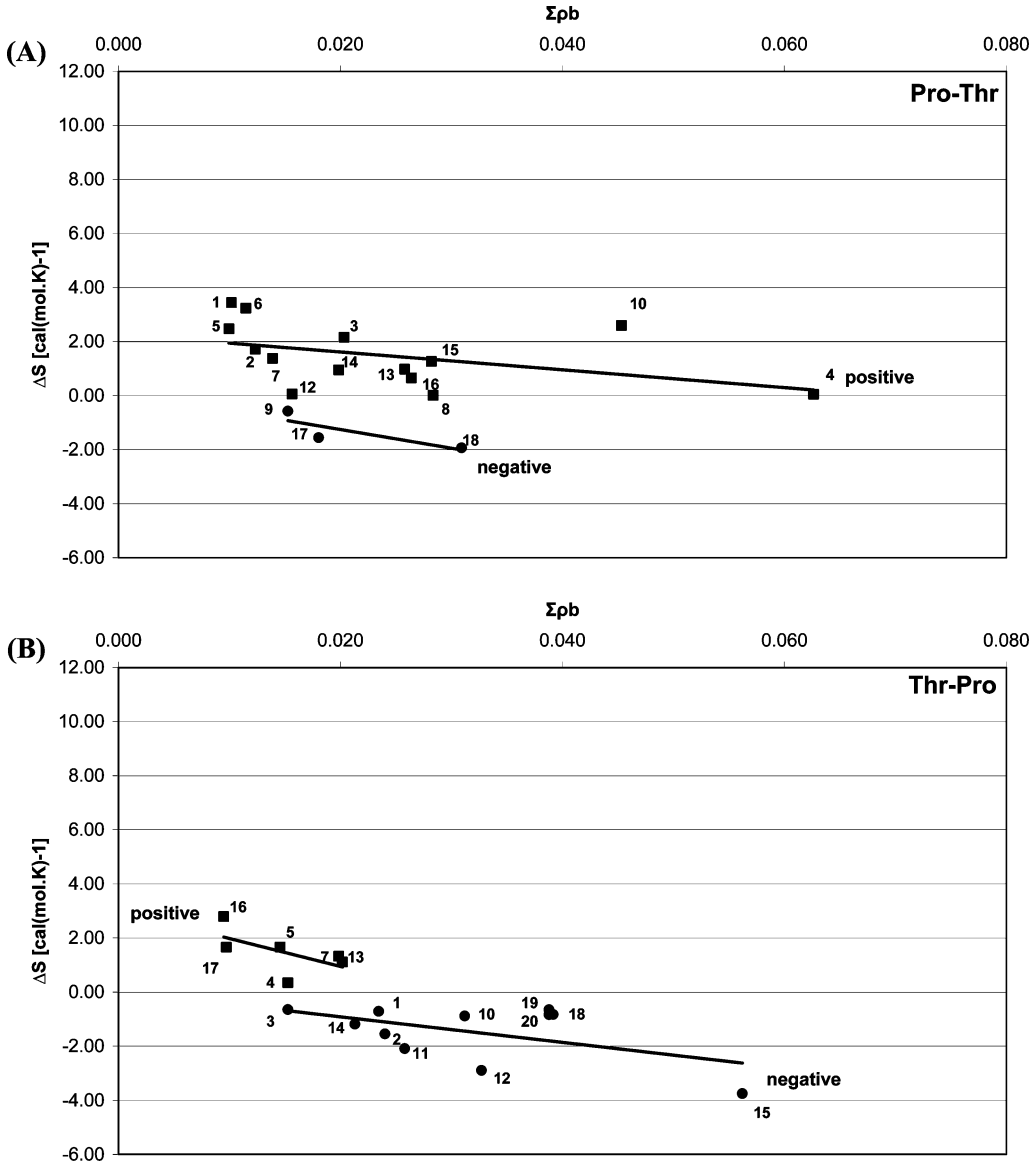


Figure 6. Correlation of computed ΔS of (A) HCO−Pro-Thr−NH_2 and (B) HCO−Thr-Pro−NH_2 with $\sum\rho_b$.

and their three-dimensional (3D) topological arrangements are shown in Figures D and E of the Supporting Information, respectively. The B3LYP/6-31G(d) optimized HCO−Pro-Thr−NH_2 and HCO−Thr-Pro−NH_2 structures are shown in Figures F and G of the Supporting Information, respectively.

Table J of the Supporting Information lists the conformations that migrated away from their initial geometries at the B3LYP/6-31G(d) level of theory for the three dipeptides.

The HCO−Pro-Thr−NH_2 and HCO−Thr-Pro−NH_2 dipeptides are structural isomers since their amino acid sequences are different. Consequently, all their thermodynamic functions are comparable. However, the thermodynamic functions of these mutants are not directly comparable to those of the HCO−Pro-Pro−NH_2 parent compound. Intuitively one would have thought that the entropy may be comparable since Pro and Thr residues have the same number of atoms and therefore the same number of vibrational frequencies. However, Pro is cyclic, whereas Thr has an open side chain; therefore, as such there is a residual difference between their entropy values, as indicated in (2).

$$\text{HCO}-\text{Pro}-\text{NH}_2 \qquad S = 94.16 \quad \text{(ref 19)}$$

$$\text{HCO}-\text{Thr}-\text{NH}_2 \qquad S = 98.65 \quad \text{(ref 20)}$$

$$\Delta S = 4.49 \quad [\text{cal}(\text{mol}\cdot\text{K})^{-1}] \qquad (2)$$

For this reason direct comparison can only be made if they are subjected to some appropriately chosen isodesmic reactions.

3.2. Isodesmic Reactions and Thermodynamic Functions. Since not all the computed thermodynamic functions are directly comparable, the values obtained were transformed via the following isodesmic reactions (Scheme 3).

Only the trans−trans conformers of the dipeptides were chosen for calculating the isodesmic reactions with their corresponding trans global minima monomers. The numerical values are summarized in Tables 1−3 for Pro-Pro, Pro-Thr, and Thr-Pro, respectively.

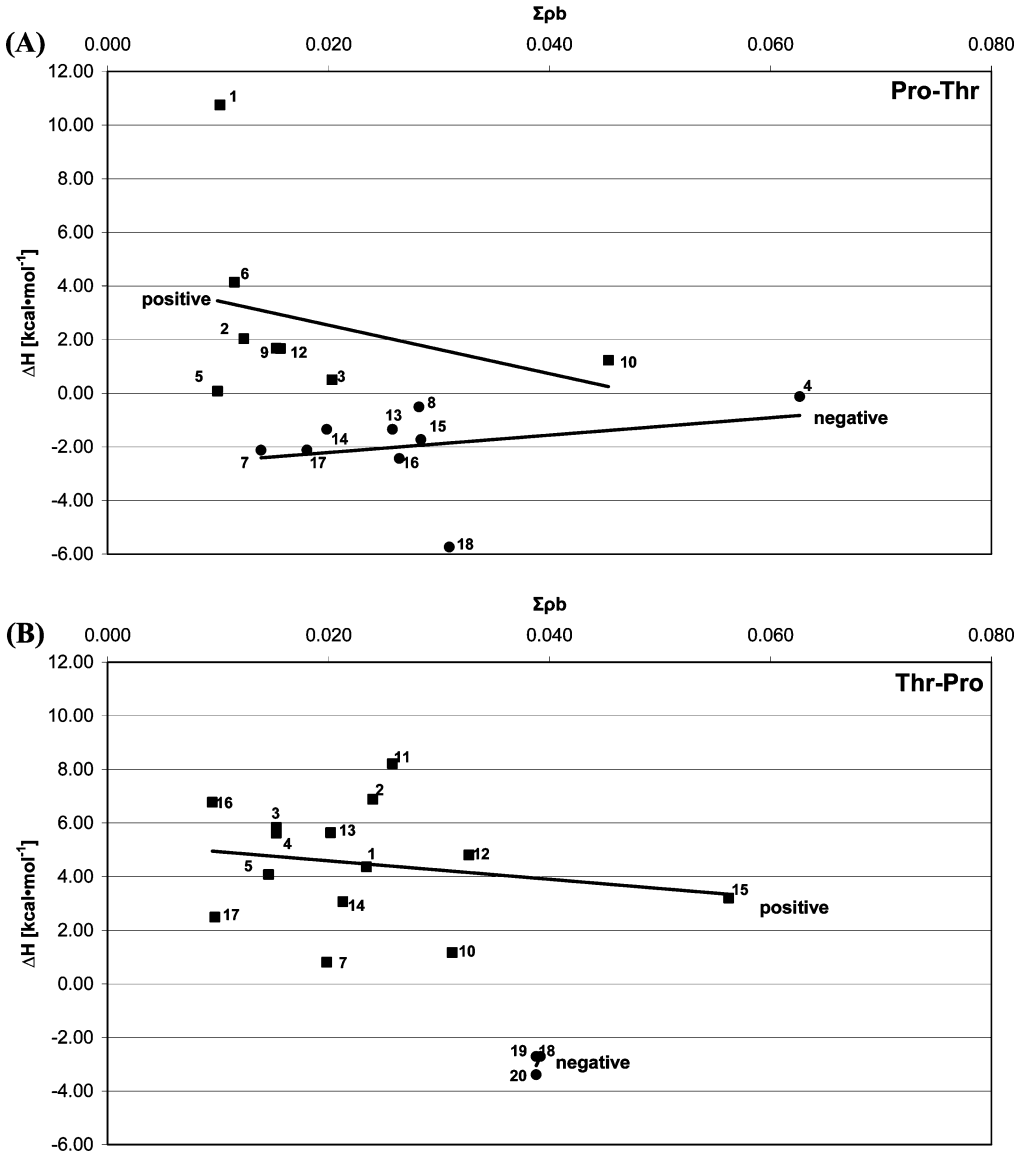


Figure 7. Correlation of computed ΔH of (A) HCO−Pro-Thr−NH_2 and (B) HCO−Thr-Pro−NH_2 with $\sum\rho_b$.

An analysis was made for the Pro-Pro + Thr isodesmic reaction by comparing the values of the trans−trans conformers in Table 1 and Table E of the Supporting Information. A general trend is observed where the ΔH and ΔG values for Pro-Pro, before and after adding the thermodynamic functions for Thr, remained exactly the same. However, the ΔS for these four conformers became negative except for the $\epsilon_L[-]\gamma_L[+]$ conformer, which is incidentally the global minima for this dipeptide at the B3LYP/6-31G(d) level of theory. This is due to the fact that different reference states were used for ΔH, ΔG, and ΔS.

3.3. Intramolecular Hydrogen Bonding. Three classes of intramolecular hydrogen bonds are recognized in HCO−Pro-Thr−NH_2 and HCO−Pro-Thr−NH_2 (Scheme 4).

Details of the H-bond types are also specified in Figure 5. The associated H-bond distances are summarized in Tables H and I of the Supporting Information, respectively. It should perhaps be emphasized that b-type hydrogen bonds are considered rather rare.[20−23] However, this study shows an interesting trend, where even the rather weak b-type H-bond is utilized to stabilize more of the cis isomers of the proline peptide bond rather than the trans forms.

The H-bond lengths (r) have been shown to be related to the electron density (ρ_b) at the bond critical point,[24] and therefore they may be, at least semiquantitatively, related to the strength of the hydrogen bond. Accordingly, the following relationship[3,24] may be used to calculate ρ_b from the bond length (r) of the hydrogen bond.

$$\rho_b(\text{a.u.}) = [2.61]e^{-(2.38)r(\text{Å})} \tag{3}$$

Applying this to the $\gamma_L[-]\gamma_L[-+]$ conformer of Pro-Thr, where two hydrogen bonds exist at 1.96 and 2.53 Å, we can calculate the corresponding $\sum\rho_b$ values.

$$\begin{aligned} \rho_b\,(1.96\ \text{Å}) &= 0.0246 \quad (80\%) \\ \rho_b\,(2.53\ \text{Å}) &= 0.0063 \quad (20\%) \\ \textstyle\sum\rho_b \quad 0.0309 \quad &(100\%) \end{aligned} \tag{4}$$

On the basis of these previous findings, the shortest hydrogen bond is always considered to have a more significant stabilizing effect. Nevertheless, $\sum\rho_b$ is a better global representation of the overall strength of H-bond stabilization than a single H-bond length.

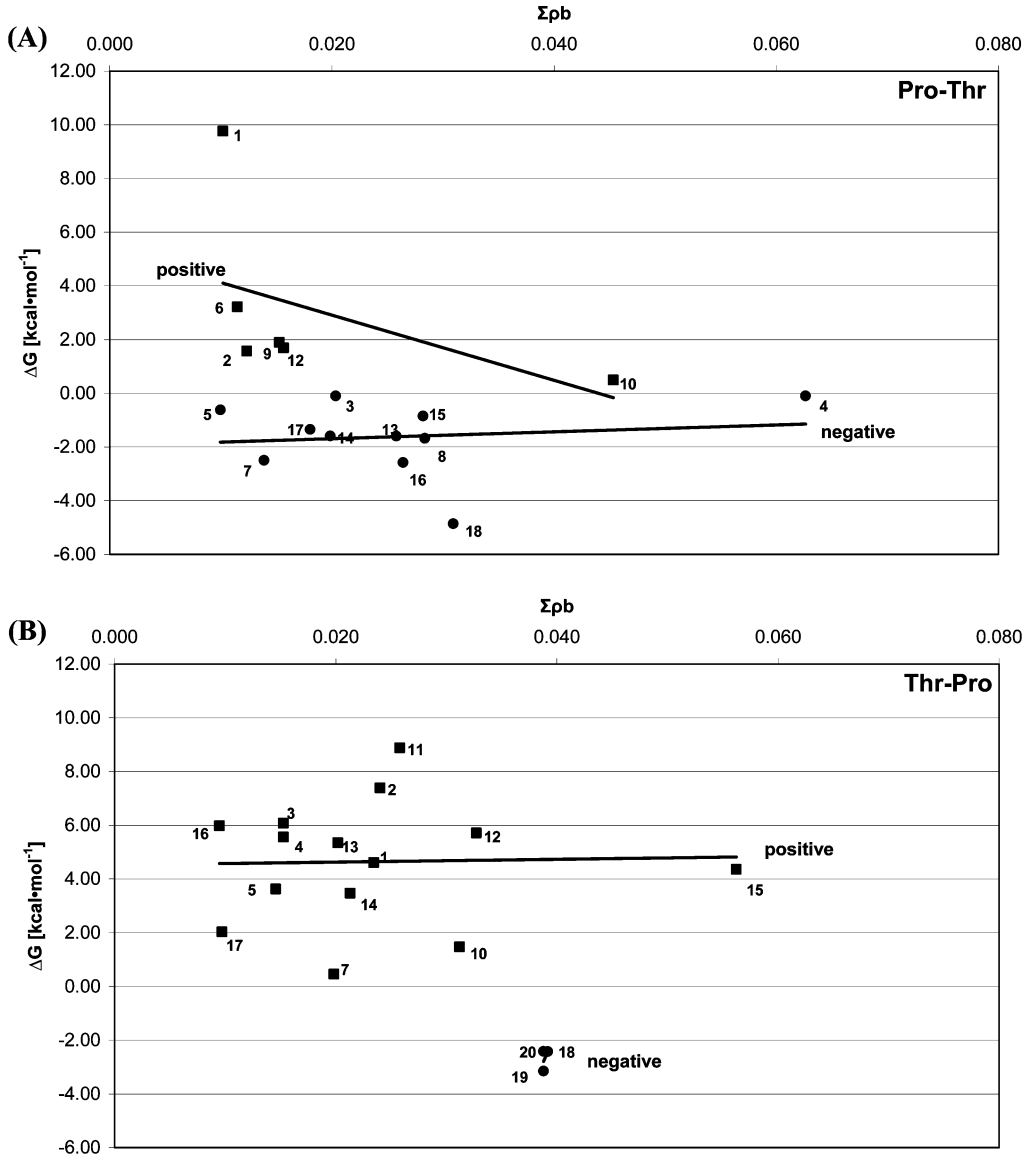


Figure 8. Correlation of computed ΔG of (A) HCO−Pro-Thr−NH_2 and (B) HCO−Thr-Pro-NH_2 with $\sum\rho_b$.

At this stage it should be emphasized that although hydrogen bonds have strong stabilizing effects there are other intramolecular interactions, such as steric effects and dipole−dipole repulsions, which counteract with the stabilizing effect of the hydrogen bond. Consequently, it is not always possible to explain trends of thermodynamic functions by direct comparison. One should allow for deviations in certain cases.

3.4. Correlations of Thermodynamic Functions and Hydrogen Bonding. It would be convenient to explain all aspects of the computed thermodynamic functions in terms of hydrogen-bond strength or in terms of the shortness of the hydrogen-bond length. As a result, a series of correlations of thermodynamic functions (ΔH, ΔS, and ΔG) with $\sum\rho_b$ for Pro-Thr and Thr-Pro, as represented in Tables 2 and 3, were made.

However, sometimes certain points do not fit to the trend because of other intramolecular interactions, which are repulsive and therefore counteract the stabilizing effect of the hydrogen bond. One such notorious conformation is *tt* $\epsilon_L[-]\gamma_L[--]$ of Pro-Thr. Therefore, this conformer was excluded from the correlations.

Figures 6−8 show the variation in both the negative and positive changes to the thermodynamic functions (ΔS, ΔH, and ΔG) for Pro-Thr and Thr-Pro, respectively. Figure 6 compares only the computed ΔS values. Interestingly a higher frequency of negative relative entropy (ΔS) values are found for Thr-Pro, which implies a more ordered state, whereas a higher frequency of positive relative entropy (ΔS) values were found for Pro-Thr, implying a disordered state. Clearly, the larger the total density ($\sum\rho_b$) the more negative the entropy change becomes. In terms of the computed ΔH values (Figure 7), Pro-Thr and Thr-Pro show some similarity for the positive values ($\Delta H > 0$). However, ΔG versus $\sum\rho_b$ (Figure 8) shows different trends when comparing the two models, possibly because of the lack of data points to see a substantial correlation.

In contrast to the foregoing, Figure 9A shows two domains for the conformers of the Pro-Thr and Thr-Pro mutants with respect to the reference conformer: $\epsilon_L[-]\gamma_L[+]$ of Pro-Pro which is at the origin ($\Delta S = \Delta H = 0.0$) of the plot. The fact that the points are clustering into two domains is a clear indication that it does make a fundamental difference at which point in the sequence the mutation is introduced. Figure 9A also indicates that negative entropy change ($\Delta S < 0$) represents information accumulation, whereas positive entropy change (ΔS

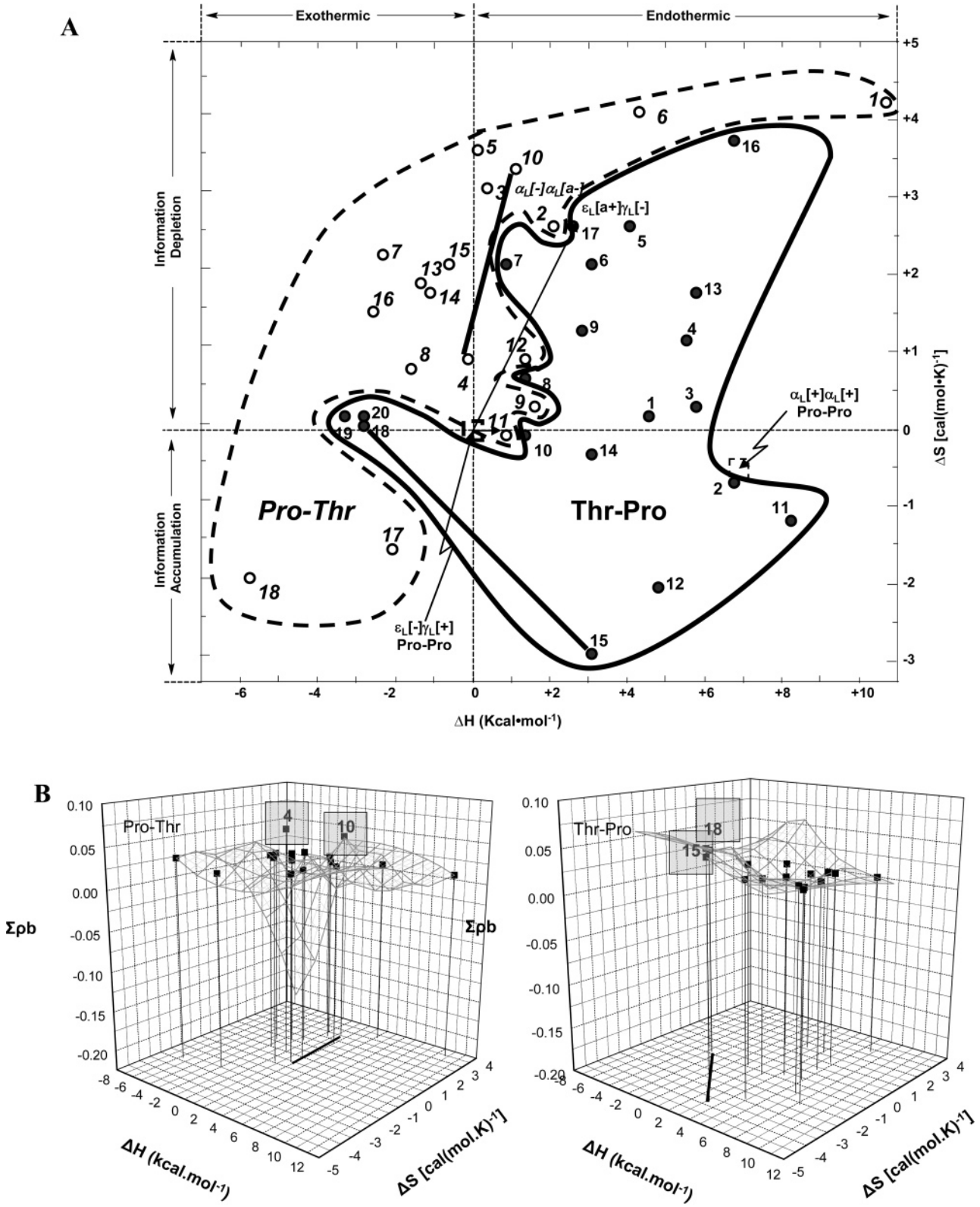


Figure 9. (A) Isodesmic ΔS vs ΔH thermodynamic domains of HCO−Pro-Thr−NH_2 and HCO−Thr-Pro−NH_2 conformers. (B) Three-dimensional representation of $\sum\rho_b = f[\Delta H, \Delta S]$ for HCO−Pro-Thr−NH_2 and HCO−Thr-Pro−NH_2.

> 0) represents information depletion according to the following relationship:[25,26]

$$\ln(I/I_0) = -\frac{\Delta S}{R} \tag{5}$$

where (I/I_0) is relative information content with respect to the reference state chosen. Thus, $I > I_0$ (accumulation) when $\Delta S < 0$ (ordered) and $I < I_0$ (depletion) when $\Delta S > 0$ (disordered).

It is clear from Figure 9A that upon mutation most of the conformers suffer information depletion and only some of them experience information accumulation.

If we take this [ΔH, ΔS] domain shown in Figure 9A as independent variables for the following function[6] we obtain Figure 9B for the Pro-Thr and Thr-Pro mutants.

$$\sum \rho_b = f[\Delta H, \Delta S] \quad (6)$$

Both of these plots are reminiscent to "tents" containing two "poles", corresponding to the highest $\sum\rho_b$ values in the two mutants. For Pro-Thr the highest points correspond to structures 4 and 10, and for the Thr-Pro mutant the highest points correspond to structures 15 and 18. These two pairs of points with high $\sum\rho_b$ values are connected by two lines in Figure 9A. These two lines are not exactly orthogonal but are noticeably oriented in different directions indicating that the site of mutation makes a fundamental thermodynamic difference. Such a thermodynamic difference may well be related to a biological difference between the two mutants.

4. Conclusion

The conformers studied for the two mutant models, Pro-Thr and Thr-Pro diamides, chosen from the PDB, were compared to the parent Pro-Pro diamide. The thermodynamic functions (S, H, and G) of the two models were computed at the B3LYP/6-31G(d) level of theory, and their relative values (ΔS, ΔH, and ΔG) were calculated with respect to the global minimum energy conformations: $\epsilon_L[-]\gamma_L[+]$ of Pro-Pro and $\gamma_L[-+]$ of Thr diamides. In the ΔS and ΔH two-dimensional (2D) coordinate system, the two models fell into two mutually exclusive domains, indicating that the site of the point mutation makes the two isomers fundamentally different.

The −OH group of the Thr side chain exhibited hydrogen bonding to the backbone, and therefore the mutants showed certain trends when the thermodynamic functions, ΔS, ΔH, and ΔG, were plotted against $\sum\rho_b$. It was observed that the larger the total density ($\sum\rho_b$), the more negative the entropy change (ΔS) becomes; however, for the other thermodynamic functions (ΔH and ΔG), no clear correlation was observed. Most of the relative entropy (ΔS) values, computed for the two mutant models, turned out to be positive, and only a few of the conformers exhibited negative ΔS with respect to the parent Pro-Pro + Thr diamides. This implied that mutations, in most of the studied cases led to structural information depletion and only a few conformers experienced structural information accumulation for the two-site Pro → Thr point mutations. Generally, more negative relative entropy (ΔS) values were found for Thr-Pro, implying a more ordered state, whereas a higher frequency of positive relative entropy (ΔS) values found for Pro-Thr, implied a disordered state.

Overall, a Pro → Thr mutation introduced in oligoprolines creates a situation whereby side-chain−backbone intramolecular hydrogen bonding becomes possible. Consequently the maximum extent, and in some sense the maximum strength of hydrogen bonds, as measured by $\sum\rho_b$, may be an indicator for the degree of perturbation caused by the mutation.

Much of this study relies on inferences made on a small conformational space provided by only two amino acids. It would be worthwhile in the future to extrapolate the findings herein to larger peptide systems (≥3 residues) while exploring mutations at the center of the peptide chain, to understand the full effects of point mutations on the thermodynamic stability of peptide models.

Acknowledgment. One of us (E.F.P.) gratefully acknowledges support from the Canada Research Chairs Program, the Natural Sciences and Engineering Research Council of Canada, and the Canadian Institutes of Health Research. We gratefully thank Dr. Zoltán Gáspári for his kind help in retrieving geometry parameters from the PDB and Dr. Gilbert G. Príve for his help in the analysis of the PDB results. We also thank Szilárd N. Fejér and Milán Szöri for their helpful discussions while preparing this manuscript.

Supporting Information Available: Retrieved geometrical parameters from the PDB for the studied dipeptides as well as important geometry parameters computed at various levels of theory and their subsequent molecular structures. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamentals of Biochemistry*, upgraded ed.; John Wiley and Sons Inc.: New York, 2002.

(2) Kalaydjieva, L.; Perez-Lezaun, A.; Angelicheva, D.; Onengut, S.; Dye, D.; Bosshard, N. U.; Jordanova, A.; Savov, A.; Yanakiev, P.; Kremensky, I.; Radeva, B.; Hallmayer, J.; Markov, A.; Nedkova, V.; Tournev, I.; Aneva, L.; Gitzelmann, R. *Am. J. Hum. Genet.* **1999**, *65*, 1299.

(3) Thoden, J. B.; Timson, D. J.; Reece, R. J.; Holden, H. M. *J. Biol. Chem.* **2005**, *280*, 9662.

(4) Sahai, M. A.; Lovas, S.; Chass, G. A.; Penke, B.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **2003**, *666*, 169.

(5) Chass, G. A.; Sahai, M. A.; Law, J. M. S.; Lovas, S.; Farkas, O.; Perczel, A.; Rivail, J. L.; Csizmadia, I. G. *Int. J. Quantum Chem.* **2002**, *90*, 933.

(6) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.

(7) Hobohm, U.; Sander, C. *Protein Sci.* **1994**, *3*, 522.

(8) Hobohm, U.; Scharf, M.; Schneider, R.; Sander, C. *Protein Sci.* **1992**, *1*, 409.

(9) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1991**, *113*, 6256.

(10) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, H.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukada, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Maliek, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.01; Gaussian Inc.: Wallingford, CT, 2004.

(11) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Theory*; John Wiley & Sons: New York, 1986.

(12) Roothan, C. C. *Rev. Mod. Phys.* **1951**, *23*, 69.

(13) Hehre, W. J.; Ditchfie, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

(14) Ditchfie, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.

(15) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(16) Berg, M. A.; Chasse, G. A.; Deretey, E.; Fuzery, A. K.; Fung, B. M.; Fung, D. Y. K.; Henry-Riyad, H.; Lin, A. C.; Mak, M. L.; Mantas, A.; Patel, M.; Repyakh, I. V.; Staikova, M.; Salpietro, S. J.; Tang, T. H.; Vank, J. C.; Perczel, A.; Csonka, G. I.; Farkas, O.; Torday, L. L.; Szekely, Z.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **2000**, *500*, 5.

(17) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.

(18) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(19) Sahai, M. A.; Kehoe, T. A. K.; Koo, J. C. P.; Setiadi, D. H.; Chass, G. A.; Viskolcz, B.; Penke, B.; Pai, E. F.; Csizmadia, I. G. *J. Phys. Chem. A* **2005**, *109*, 2660.

(20) Sahai, M. A.; Fejer, S. N.; Viskolcz, B.; Pai, E. F.; Csizmadia, I. G. *J. Phys. Chem. A* **2006**, *110*, 11527.

(21) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97.

(22) Stickle, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D. *J. Mol. Biol.* **1992**, *226*, 1143.

(23) Mcdonald, I. K.; Thornton, J. M. *J. Mol. Biol.* **1994**, *238*, 777.
(24) Tang, T. H.; Deretey, E.; Jensen, S. J. K.; Csizmadia, I. G. *Eur. Phys. J. D* **2006**, *37*, 217.
(25) Fejer, S. N.; Csizmadia, I. G.; Viskolcz, B. *J. Phys. Chem. A* **2006**, *110*, 13325.
(26) Viskolcz, B.; Fejer, S. N.; Csizmadia, I. G. *J. Phys. Chem. A* **2006**, *110*, 3808.