Quantifying the Intrinsic Effects of Two Point Mutation Models of Pro-Pro-Pro Triamino 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 & 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 25, 2007; In Final Form: August 21, 2007

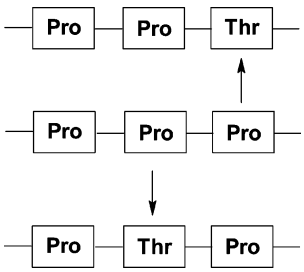
Stabilities and conformational properties of two Pro → Thr point mutation models were computed at the B3LYP/6-31G(d) level of theory for the parent triamino acid diamide Pro-Pro-Pro (HCO-Pro-Pro-Pro-NH_2). Geometrical parameters for the amino acid sequences, used in the molecular orbital computations for Pro-Pro-Thr and Pro-Thr-Pro, were retrieved from the Protein Data Bank. Thermodynamic functions (*S*, *H*, *G*) were computed for the fully optimized geometries. To assess the stabilization energetics of these mutant models, relative to the parent Pro-Pro-Pro reference conformer $\epsilon_L\epsilon_L\gamma_L$, isodesmic reactions were constructed to calculate ΔS, ΔH, and ΔG. The importance of intramolecular hydrogen bonds involving the −OH group of the Thr side chain, which emerged after the point mutations, was also examined to determine the internal stabilization of these peptide models. This study describes an approach to analyzing a point mutation at the center of a peptide chain and compares its stability to that of a point mutation at a terminal end in a small peptide model.

1. Introduction

It is well known that the physical and chemical properties of proteins such as molecular conformations and thermodynamic stabilities can be affected by mutations in the primary structure, the replacement or substitution of the wild-type amino acid residue by another residue. Therefore, it is natural to assume that the biological effects of these mutations will also be site-dependent. Recent experimental results, involving point mutations at various sites of oligoalanine, provide supporting evidence for such an *a priori* assumption.[1] The replacement of proline residues in proteins has also been shown to have significant biological consequences.[2] The effect of such mutations has previously been studied at two terminal sites of a short proline chain.[3] Therefore, as a continuation to the study, the present article explores the mutation at the center of a proline chain and compares its stability to that of a point mutation carried out at the C-terminal proline residue (Scheme 1). The component single amino acid (Pro and Thr) diamides have also been previously published.[4,5]

It is hoped that this study as well as the related previous studies[3−5] will provide an underlying foundation to explain the molecular structure of a key Pro-Pro-Thr-Pro tetrapeptide fragment of the hinge region of human immunoglobulin A1 (IgA1) when it is cleaved by an extracellular IgA1 protease.[6,7]

SCHEME 1: Pro-Pro-Pro Tripeptide and Its Thr Single-Point Mutant Models



2. Methods

2.1. Conformational and Configurational Specifications. Numeric definitions of the relative spatial orientation of all constituent atoms of HCO-Pro-Pro-Pro-NH_2, HCO-Pro-Pro-Thr-NH_2, and HCO-Pro-Thr-Pro-NH_2 follow a standard protocol,[8,9] as illustrated in Figure 1. As a result, amino acid residues in the tripeptides, as well as the protecting end groups, are exclusively defined using the *z*-matrix internal coordinate system to characterize molecular structure, geometry, and stereochemistry.

The number of possible conformational states that the tripeptide mutants can exist in, on the basis of backbone (BB) and side-chain (SC) geometries, is large. The magnitude of the conformational search can simply be grasped by just considering the BB geometries as described by multidimensional confor-

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

† 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/jp074046r CCC: $37.00
Published on Web 10/20/2007

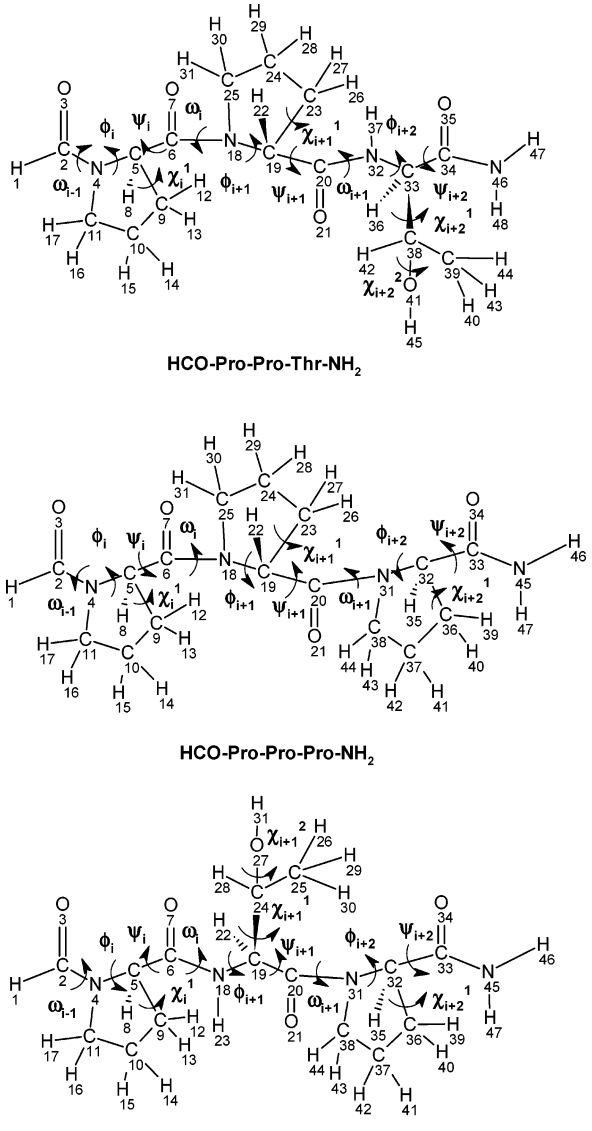


Figure 1. Numbering system employed for HCO-Pro-Pro-Thr-NH_2, HCO-Pro-Pro-Pro-NH_2, and HCO-Pro-Thr-Pro-NH_2 with definitions of backbone and side-chain dihedral angles.

mational analysis (MDCA).[10−17] Each of the two Pro (P) residues can assume three BB conformations (α_L, ϵ_L, γ_L) and Thr (T) nine BB conformations (α_L, α_D, β_L, δ_L, δ_D, ϵ_L, ϵ_D, γ_L, γ_D). Considering trans → cis (c/t) isomerism in the peptide bonds, as well as the side-chain orientation for each of the residues, would yield 23 328 potential optimizations for either PPT or PTP, calculated as follows:

$$\overbrace{[9]}^{729} \times [3^2] \times \overbrace{[9]}^{4} \times \overbrace{[2^2] \times [2^3]}^{8} = 23,328$$

T P T P *c/t*

BB SC

For this reason, the conformers of these two model tripeptide mutant sequences, PPT and PTP, were specifically chosen from the Brookhaven Protein Database (PDB) to reduce the complexity of the conformational search by selecting conformers existing in nature.

2.2. Structural Features and the Selection of Geometry Parameters from the Brookhaven PDB. Computed torsional angles (ω, ϕ, ψ, χ^1, and χ^2) for the chosen tripeptide models were derived from retrieved dihedral angles (Tables A and B in the Supporting Information) from proteins, listed in the PDB, to reduce the MDCA[18] search (Figure 2). This method was described in detail elsewhere.[3]

Only a few of the 38 structures retrieved from the PDB for the two tripeptides were of cis configurations of the type ctt, tct, and ttc. However, since a substantial amount of the retrieved structures was completely of the trans (ttt) configuration the focus of the computations was mostly on structures existing in the full trans (ttt) configuration.

The ttt $\epsilon_L\epsilon_L\epsilon_L$ conformer was considered as the parent conformer for Pro-Pro-Pro, on the basis of structural features and trends from the PDB results, where the ϕ and ψ backbone dihedral angles for Pro residues had a preference for values around −60° and 150°, respectively.

2.3. Molecular Computations. Traditionally, protein and peptide models were studied using force-field methods that can also produce thermodynamic functions. The novelty of the present article is that *first-principle* molecular computations are used to study the phenomena of point mutations.

First-principle methods are based on the many-electron Schrödinger equation, $H\Psi = E\Psi$.[19,20] It encompasses ab initio[21] Hartree Fock[22]-type calculations and density functional theory (DFT)[23,24] in which the energy (E) is a function of density (ρ) which in turn is a function of structure (eq I):

$$E = f[\rho(\text{structure})] \qquad \text{(I)}$$

The retrieved geometries of the 38 tripeptide mutant structures were used as starting structures for full optimizations using the ab initio RHF/3-21G level of theory followed by the standard DFT employing the B3LYP hybrid function[25] at the 6-31G(d) level of theory. Here, B3LYP denotes the combination of Becke's three-parameter exchange functional with the Lee−Yang−Parr (LYP)[26] correlation functional and also employs the mathematically more complete 6-31G(d) basis set. All computations were carried out using the Gaussian 03[27] program package (G03). 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}). Thermodynamic functions were also computed using the DFT method. From the optimized geometries, the hydrogen bondings due to the introduced threonine were also determined.

2.4. Potential Energy Surfaces. Potential energy surfaces (PESs) of the type $E = f(\psi_2,\psi_3)$ for Pro-Pro-Pro were computed at the RHF/3-21G level of theory, by varying the ψ_1 dihedral angle for the first Pro residue in 30° increments in the range 0° ≥ ψ_1 ≤ 360°. The scientific graphing software Origin 7.5[28] was used to present the data as spline and contour diagrams in the range of 0° to +360°.

2.5. Isodesmic Reactions. Isomeric compounds, meaning compounds having the same number and type of atoms, have comparable energy values and thermodynamic functions. However, even if there is the same number of atoms but there is one atom type difference between the two compounds, such as in the cases of proline and threonine where instead of a carbon atom there is an oxygen, their energetic values are no longer comparable. To have comparable values to establish relative stabilities, an isodesmic (ID) reaction has to be constructed. An isodesmic reaction will then have the same number and type of bonds on both sides of the equation, and the difference in energetics in the thermodynamic sense, the isodesmic energy [E], enthalpy [H], entropy [S], or Gibbs free energy [G] values,

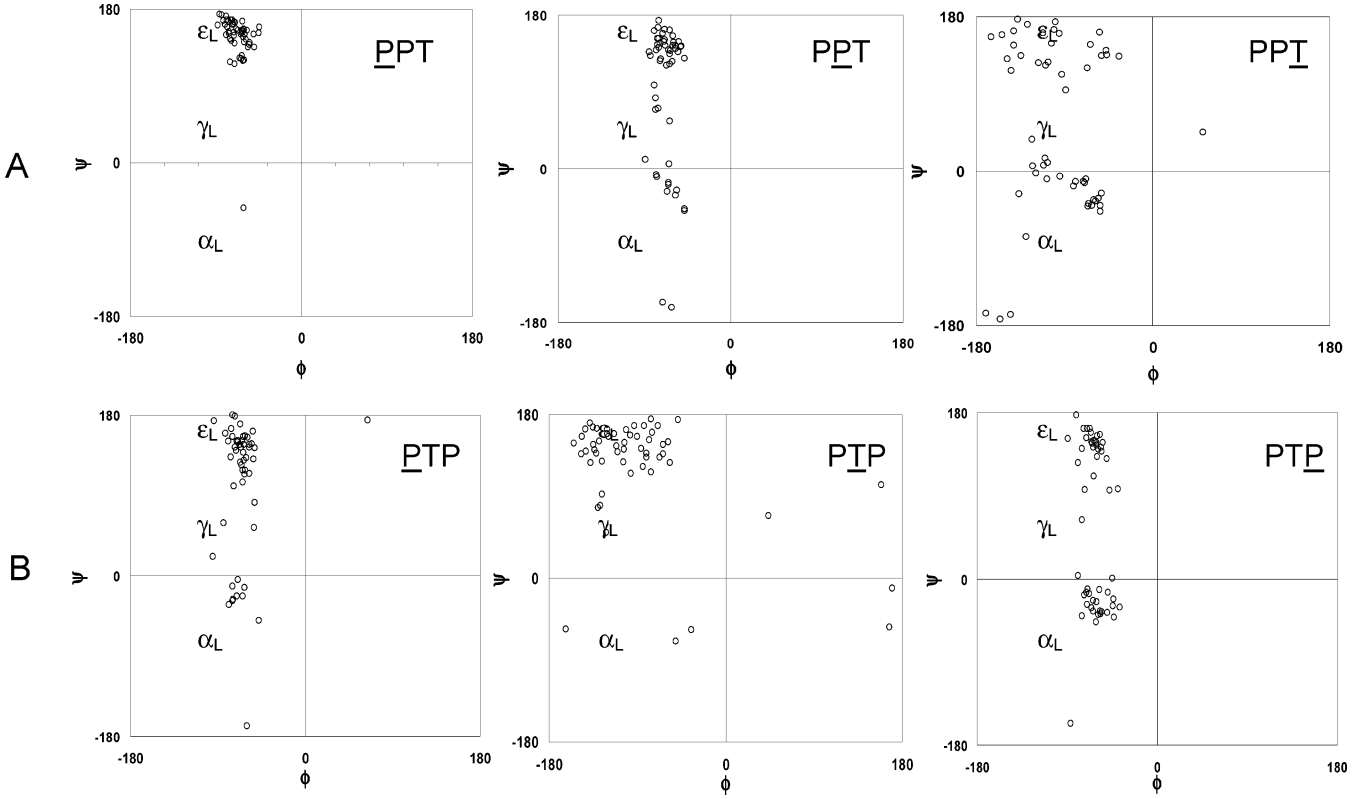


Figure 2. 2D Ramachandran plots showing the ϕ,ψ distribution of the residues in the two tripeptides, (A) PPT and (B) PTP, retrieved from the PDB.

will give the relative change with respect to the primary standard, which in this study is two Pro → Thr mutations (PPT and PTP) of the Pro-Pro-Pro tripeptide [eqs II and III].

$$\text{PPP} + \text{T} \rightarrow \text{PTP or PPT} + \text{P} \quad \text{(II)}$$

$$\Delta X_{\text{ID}} = [X(\text{PPT or PTP}) + X(\text{P})] - [(X(\text{PPP}) + X(\text{T})] \quad \text{(III)}$$

where $X = E$, H, S, or G.

All calculations for the isodesmic reactions utilized energy values and thermodynamic functions calculated at the B3LYP/6-31G(d) level of theory.

3. Results and Discussion

3.1. Results of the PDB Analysis. Tables A and B in the Supporting Information present selected parameters (ϕ, ψ, χ dihedral angles) from the PDB for the Pro-Pro-Thr and Pro-Thr-Pro tripeptide sequences. These conformational properties of the tripeptides, specifically the various ϕ and ψ backbone dihedral angles, are better described and visualized using 2D (ϕ,ψ) type Ramachandran plots, as shown in Figure 2.

The 2D Ramachandran plots aimed to characterize each of these tripeptides by showing the ϕ,ψ distribution of each of the residues in the sequence. Pro-Pro-Thr and Pro-Thr-Pro are presented in Figure 2A and 2B, respectively, while residue names are underlined as corresponding to their individual plots. A substantial distribution is noticed around values of $\phi = -60°$ and $\psi = 150°$ (corresponding to the ϵ_L conformation) for all residues. Ab initio quantum mechanical calculations were then carried out to further assess the energetic preferences of these conformers.

3.2. Molecular Structures and Energetics. *3.2.1. Pro-Pro-Pro Tripeptide.* To rationalize the conformational space available for the first Pro residue in the Pro-Pro-Pro tripeptide, 12 PESs of the type $E = f(\psi_2,\psi_3)$, were generated for the ttt HCO-Pro-Pro-Pro-NH_2 at the RHF/3-21G level of theory. These 12 PESs are invaluable as they provide information for subsequent calculations and are presented in Figure A in the Supporting Information. It is interesting to note that the diagrams appear strikingly similar, and only upon closer inspection does $\psi_1 = 60°$ for the N-terminal Pro residue show some small variation at $\psi_2 = 180°$ and $\psi_3 = 60°$. As such, the N-terminal Pro residue in the Pro-Pro-Pro tripeptide was neglected in the point mutations.

TABLE 1: Selected Parameters for the HCO-Pro-Pro-Pro-NH_2 + HCO-Thr-NH_2 Initial State Using Thermodynamic Functions Computed at the B3LYP/6-31G(d) Level of Theory[a]

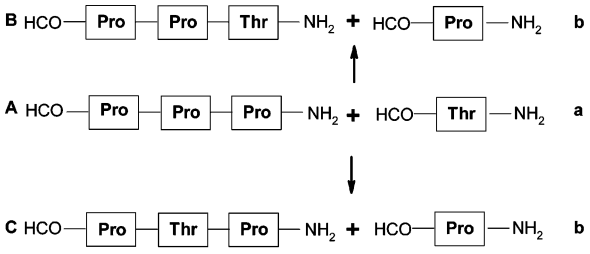
conformers	enthalpy	entropy	Gibbs free energy
ttt ϵ_L [−] ϵ_L [−] γ_L [−]	−1674.716842	261.96	−1674.841950

[a] Global minima of HCO-Thr-NH_2 = tt γ_L [− +]: ΔH = −531.030809, ΔS = 98.65, ΔG = −531.078324. Total energies are given in hartrees.

For the trans−trans−trans (ttt) isomer, the $\epsilon_L\epsilon_L\epsilon_L$ conformer was considered to be the parent reference state for Pro-Pro-Pro, by following the trends from the PDB results, with values of ϕ and ψ at −60° and 150°, respectively. However, ttt $\epsilon_L\epsilon_L\epsilon_L$ converged to the ttt $\epsilon_L\epsilon_L\gamma_L$ conformer during initial geometry optimizations at the RHF/3-21G level of theory. Subsequent energetic refinement was performed on ttt $\epsilon_L\epsilon_L\gamma_L$ at the B3LYP/6-31G(d) level of theory. These results are summarized in Table C in the Supporting Information. The ttt $\epsilon_L\epsilon_L\gamma_L$ was consequently used as the reference conformer, particularly since the γ_L backbone geometry affecting the C-terminal Pro residue was also shown to exist based on the PPT and PTP results from the PDB (Figure 2B).

3.2.2. Pro-Pro-Thr and Pro-Thr-Pro Mutant Models. The structures of HCO-Pro-Pro-Thr-NH_2 and HCO-Pro-Thr-Pro-NH_2 optimized at the B3LYP/6-31G(d) level of theory are sum-

SCHEME 2: Schematic Representation of the Isodesmic Reactions (Eqs IV and V), Involving the Three Corresponding Tripeptides



$$\Delta E_{ID}(PPT) = [E(B) + E(b)] - [(E(A) + E(a)] \qquad [IV]$$

$$\Delta E_{ID}(PTP) = [E(C) + E(b)] - [(E(A) + E(a)] \qquad [V]$$

marized in Tables D and E in the Supporting Information, respectively. In addition to optimized geometries and energies (*E*), thermodynamic functions (*H*, *S*, and *G*) and their relative values (ΔE, ΔH, ΔS, and ΔG) were also computed at that level of theory and summarized in Tables D and E in the Supporting Information.

The RHF/3-21G results are listed in the Tables F and G in the Supporting Information. Some of the conformers obtained at the RHF/3-21G level of theory shifted to other conformations when reoptimized at the B3LYP/6-31G(d) level of theory. These structures are listed in Table H in the Supporting Information. The optimized molecular structures of these two mutants are depicted in Figures B and C in the Supporting Information.

3.3. Isodesmic Reactions and Thermodynamic Functions. Since not all of the computed thermodynamic functions are directly comparable, the values obtained were transformed via the isodesmic reactions (eqs IV and V) shown in Scheme 2.

Only the trans−trans−trans conformers of the tripeptides were chosen for calculating the isodesmic reactions with their corresponding trans global mimima monomers. The numerical values are summarized in Tables 1, 2, and 3 for Pro-Pro-Pro, Pro-Pro-Thr, and Pro-Thr-Pro, respectively.

3.4. Intramolecular Hydrogen Bondings. Three classes of intramolecular hydrogen bonds are recognized in HCO-Pro-Pro-Thr-NH_2 and HCO-Pro-Thr-Pro-NH_2 (Scheme 3).

Details of the H-bond types are also specified in Figure 3. The associated H-bond distances are summarized in Tables 2 and 3, respectively.

The H-bond lengths (*r*) have been shown to be related to the electron density (ρ_b) at the bond critical point,[29] and therefore they may be, at least semiquantitatively, related to the strength of the hydrogen bond. Accordingly, the following relationship

TABLE 2: Selected Parameters for the HCO-Pro-Pro-Thr-NH_2 + HCO-Pro-NH_2 Final Mutated State Using Thermodynamic Functions Computed at the B3LYP/6-31G(d) Level of Theory[a]

	conformers	B3LYP/6-31G(d) enthalpy total	enthalpy relative	entropy total	entropy relative	Gibbs free energy total	Gibbs free energy relative	H-bonding distances (Å) a_1	a_2	a_3	a_4	b_1	b_2	c_1	c_2	c_3	$\sum\rho_b$
1	ttt α_L [−] γ_L [−] β_L [++]	−1674.713177	2.30	262.25	0.29	−1674.838351	2.26									2.05	0.0198
2	ttt ϵ_L [−] α_L [−] α_L [a−]	−1674.715539	0.82	262.84	0.88	−1674.840993	0.60								2.22		0.0132
3	ttt ϵ_L [−] α_L [−] δ_D [− −]	−1674.709184	4.81	263.80	1.84	−1674.83 5094	4.30									2.02	0.0215
4	ttt ϵ_L [−] α_L [−] γ_L [a+]	−1674.714239	1.63	261.73	−0.23	−1674.839166	1.75						2. 29				0.0111
5	ttt ϵ_L [−] α_L [−] γ_L [−+]	−1674.721987	−3.23	263.52	1.56	−1674.847764	−3.65				2. 20						0.01 38
6	ttt ϵ_L [−] ϵ_L [−] β_L [a+]	−1674.718827	−1.25	266.72	4.76	−1674.846129	−2.62		1.88								0.02 94
7	ttt ϵ_L [−] ϵ_L [−] β_L [++]	−1674.719203	−1.48	264.35	2.39	−1674.845378	−2.15									2.01	0.0217
8	ttt ϵ_L [−] ϵ_L [−] δ_L [a−]	−1674.716971	−0.08	267.02	5.06	−1674.844414	−1.55		1.94						2.05		0.0453
9	ttt ϵ_L [−] γ_L [−] α_D [−+]	−1674.713811	1.90	263.22	1.26	−1674.839449	1.57				2.16						0.0153
10	ttt ϵ_L [−] γ_L [−] δ_D [− −]	−1674.719723	−1.81	259.89	−2.07	−1674.843778	−1.15		1.92							2.01	0.0485
11	ttt ϵ_L [−] γ_L [−] δ_L [+−]	−1674.721059	−2.65	263.44	1.48	−1674.846798	−3.04				1.93						0.0267
12	ttt ϵ_L [−] γ_L [−] γ_L [a+]	−1674.722132	−3.32	259.77	−2.19	−1674.846130	−2.62		1.96								0.0247
13	ttt ϵ_L [−] γ_L [−] γ_L [−+]	−1674.726594	−6.12	262.71	0.75	−1674.851988	−6.30								2.53		0.0063

[a] Global minima of HCO-Pro-Pro-Pro-NH_2 = tt ϵ_L [−] ϵ_L [−] γ_L [−]: $\Delta H = -1674.716842$, $\Delta S = 261.96$, $\Delta G = -1674.841950$. Global minima of HCO-Pro-NH_2 = tt γ_L [+]: $\Delta H = -493.918627$, $\Delta S = 94.16$, $\Delta G = -493.963938$. Total energies are given in hartrees, and relative energies are given in kilocalories per mole.

TABLE 3: Selected Parameters for the HCO-Pro-Thr-Pro-NH_2 + HCO-Pro-NH_2 Final Mutated State Using Thermodynamic Functions Computed at the B3LYP/6-31G(d) Level of Theory[a]

	conformers	B3LYP/6-31G(d) enthalpy total	enthalpy relative	entropy total	entropy relative	Gibbs free energy total	Gibbs free energy relative	H-bonding distances (Å) a_1	a_2	a_3	a_4	b_1	b_2	c_1	c_2	c_3	$\sum\rho_b$
1*	ttt α_L [−] β_L [++] γ_L [−]	−1674.714573	1.42	264.75	2.79	−1674.840938	0.64										
2*	ttt α_L [−] δ_D [− −] γ_L [−]	−1674.711418	3.40	260.94	−1.02	−1674.835970	3.75		1.81								0.0349
3*	ttt α_L [+] ϵ_L [a+] γ_L [−]	−1674.716003	0.53	263.47	1.51	−1674.841761	0.12						2.33				0.0103
4*	ttt α_L [+] ϵ_L [a+] γ_L [+]	−1674.716758	0.05	263.87	1.91	−1674.842701	−0.47						2.33				0.0103
5*	ttt ϵ_L [−] β_L [++] α_L [−]	−1674.715752	0.68	263.16	1.19	−1674.841358	0.37									2.15	0.0157
6*	ttt ϵ_L [−] β_L [+−] ϵ_L [−]	−1674.721598	−2.98	261.74	−0.22	−1674.846529	−2.87				2.03						0.0210
7*	ttt γ_L [−] α_D [−+] γ_L [−]	−1674.719285	−1.53	260.49	−1.47	−1674.843626	−1.05			1.99							0.0231
8*	ttt γ_L [−] α_L [a−] γ_L [−]	−1674.714478	1.48	262.53	0.57	−1674.839787	1.36	2.03						2.04			0.0411
9*	ttt γ_L [−] β_L [++] γ_L [−]	−1674.718634	−1.12	263.30	1.34	−1674.844311	−1.48										
10*	ttt γ_L [−] δ_D [a−] γ_L [−]	−1674.715468	0.86	262.91	0.95	−1674.840955	0.62		2.06								0.0195
11*	ttt γ_L [+] ϵ_L [a+] α_L [−]	−1674.716275	0.36	261.63	−0.33	−1674.841155	0.50	2.01									0.0220
12*	ttt γ_L [−] ϵ_L [a+] γ_L [−]	−1674.723011	−3.87	261.25	−0.71	−1674.847714	−3.62	2.01									0.0219
13*	ttt γ_L [−] γ_L [−+] γ_L [−]	−1674.725878	−5.67	261.60	−0.36	−1674.850747	−5.52			1.90				2.41			0.0366

[a] Global minima of HCO-Pro-Pro-Pro-NH_2 = tt ϵ_L [−] ϵ_L [−] γ_L [−]: $\Delta H = -1674.716842$, $\Delta S = 261.96$, $\Delta G = -1674.841950$. Global minima of HCO-Pro-NH_2 = tt γ_L [+]: $\Delta H = -493.918627$, $\Delta S = 94.16$, $\Delta G = -493.963938$. Total energies are given in hartrees, and relative energies are given in kilocalories per mole.

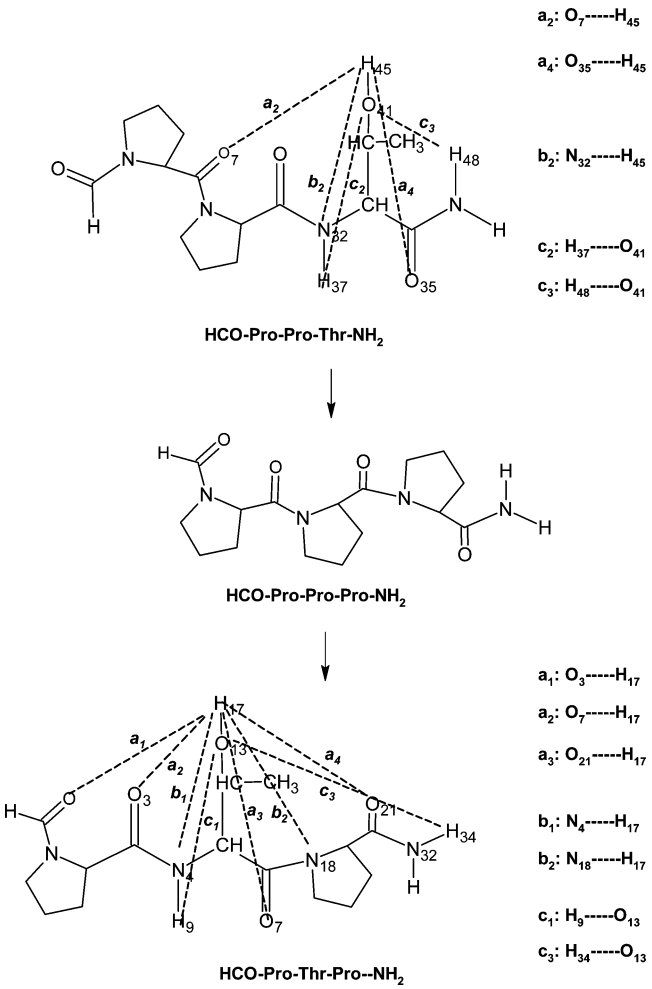
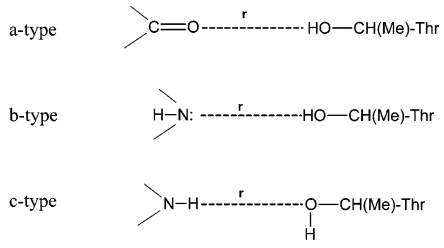


Figure 3. Side chain/backbone hydrogen-bonding networks available with the two HCO-Pro-Pro-Pro-NH_2 tripeptide mutants: HCO-Pro-Pro-Thr-NH_2 and HCO-Pro-Thr-Pro-NH_2. The hydrogen bond designation and atoms involved in hydrogen bonding are indicated for each mutant.

SCHEME 3: Three Types of Hydrogen Bonds and Its Hydrogen Bond Length (*r*) Occurring in the Pro-Pro-Thr and Pro-Thr-Pro Mutant Models



(eq VI),[29] which was described previously,[3] was used to calculate ρ_b from the bond length (r) of the hydrogen bond.

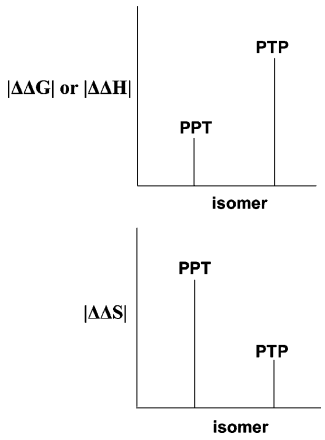
$$\rho_b\,(\text{au}) = [2.61]e^{-(2.38)r(\text{Å})} \qquad \text{(VI)}$$

This equation is reminiscent of the Pauling bond energy bond order (Pauling-BEBO) correlation equation[30−32] (eq VII), where bond order (n) is related exponentially to the bond length (r)

$$n = [e^{-(1/a)r_s}]e^{-(1/a)r} \qquad \text{(VII)}$$

and r_s is the reference bond length.

SCHEME 4: Schematic Illustration of the Pattern of Thermodynamic Functions of Tripeptide Mutants



Applying this to the ttt ϵ_L [−] ϵ_L [−] δ_L [a−] conformer of Pro-Pro-Thr where two hydrogen bonds exist at 1.94 and 2.05 Å, we can calculate the corresponding $\Sigma\rho_b$ values:

$$\rho_b(1.94\ \text{Å}) = 0.0256\ (57\%)$$
$$\rho_b(2.05\ \text{Å}) = 0.0197\ (43\%)$$
$$\Sigma\rho_b = 0.0453\ (100\%) \qquad \text{(VIII)}$$

On the basis of these and previous findings,[3] the shortest hydrogen bond is always considered to have a more significant stabilizing effect. Nevertheless, in the case of multiple hydrogen bondings, $\Sigma\rho_b$ is a better global representation of the overall strength of H-bond stabilization than a single H-bond length.

3.5. Correlations of Thermodynamic Functions and Hydrogen Bonding. A series of correlations of thermodynamic functions (ΔH, ΔS, and ΔG) with $\Sigma\rho_b$ for Pro-Pro-Thr and Pro-Thr-Pro, as represented in Tables 2 and 3, were made 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.

Scheme 4 shows that for PPT the change in ΔH or ΔG is small while the change in ΔS is large. On the other hand, for PTP the change in ΔH or ΔG is large while the change in ΔS is small. Figures D−F in the Supporting Information clearly show the maximum values of the $\Sigma\rho_b$ for the change in the thermodynamic functions (ΔH, ΔS, and ΔG).

Figure 4A shows two arbitrarily drawn lines enclosing the two domains for the conformers of the Pro-Pro-Thr and Pro-Thr-Pro mutants with respect to the reference conformer: ϵ_L [−] ϵ_L [−] γ_L [−] of Pro-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. This was also shown previously to be true.[3] Figure 4A also indicates that negative entropy change ($\Delta S < 0$) represents information accumulation while positive entropy change ($\Delta S > 0$) represents information depletion according to the following relationship:[33,34]

$$\ln(I/I_0) = -\frac{\Delta S}{R} \qquad \text{(IX)}$$

where (I/I_0) is the relative information content with respect to the reference state chosen. Thus, $I > I_0$ (accumulation) when

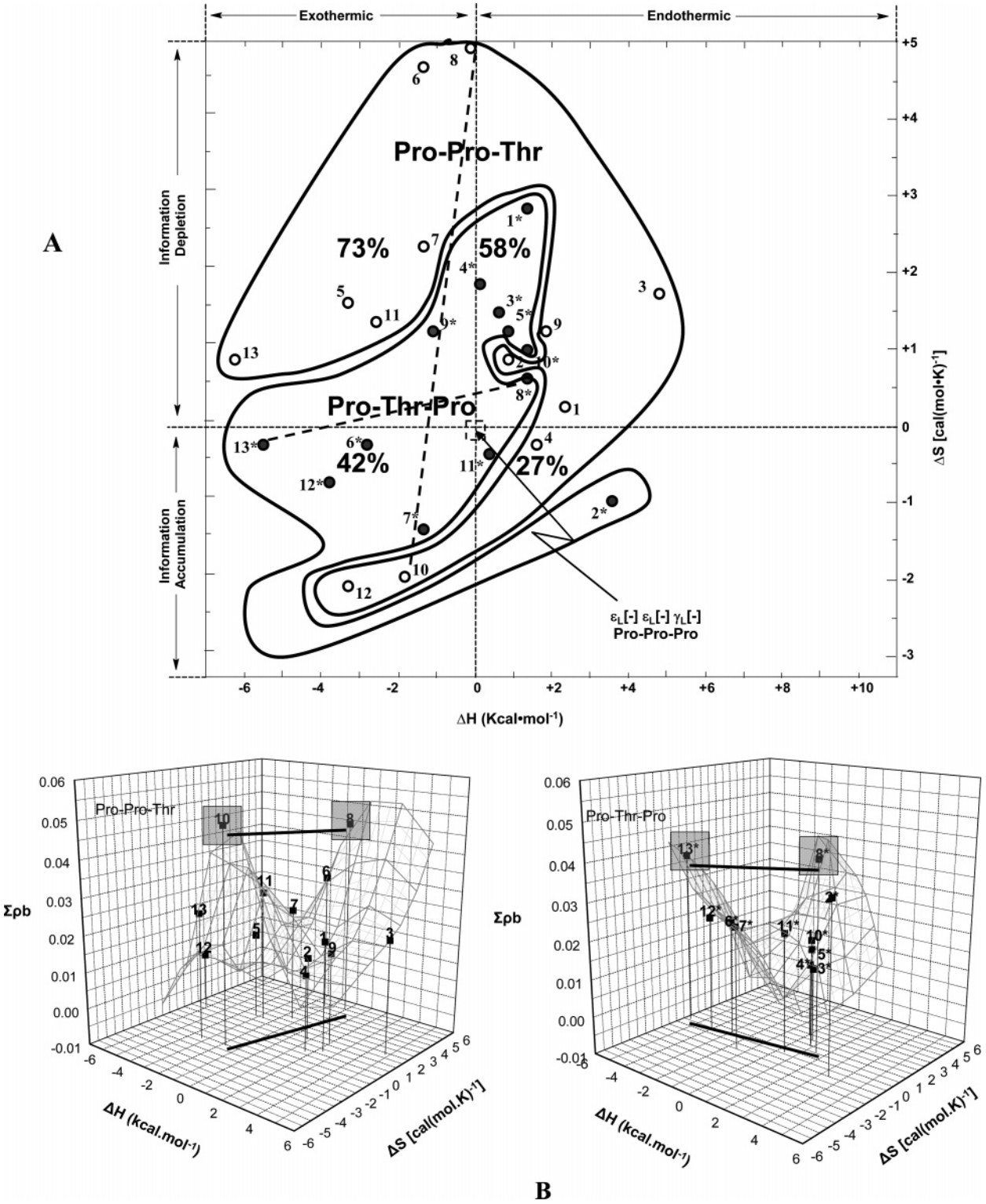


Figure 4. (A) Isodesmic ΔS versus ΔH thermodynamic domains of HCO-Pro-Pro-Thr-NH_2 and HCO-Pro-Thr-Pro-NH_2 conformers. (B) 3D representation of $\Sigma\rho_b = f[\Delta H,\Delta S]$ for HCO-Pro-Pro-Thr-NH_2 and HCO-Pro-Thr-Pro-NH_2.

$\Delta S < 0$ (ordered) and $I < I_0$ (depletion) when $\Delta S > 0$ (disordered).

It is clear from Figure 4A that upon mutation most of the conformers suffer information depletion and only some of them experience information accumulation. This is the situation in 73% of the cases studied for Pro-Pro-Thr and 58% of the cases studied for Pro-Thr-Pro.

A mutation involving Pro → Thr creates the possibility of several intramolecular hydrogen bonds that depend on the site of the mutation as well as the conformations involved. The $\Sigma\rho_b$ values are regarded to be proportional to the total hydrogen bond stabilization energy,[29] which is characteristic for a given conformation of a specific mutation. The highest $\Sigma\rho_b$ values would then represent a measure of the Pro → Thr mutational perturbation of the peptide sequence.

If we take this [ΔH, ΔS] domain shown in Figure 4A as independent variables for the following function (eq X), we obtain Figure 4B for Pro-Pro-Thr and Pro-Thr-Pro.

$$\Sigma\rho_b = f[\Delta H, \Delta S] \quad \text{(X)}$$

Both of these plots are reminiscent of "tents" containing two "poles", corresponding to the highest $\Sigma\rho_b$ values in the two mutants. For Pro-Pro-Thr the highest points correspond to structures 8 and 10, and for the Pro-Thr-Pro mutant the highest points correspond to structures 8* and 13*. These two pairs of points with high $\Sigma\rho_b$ values are connected by two lines in Figure 4A. These two lines are not exactly orthogonal but are oriented in noticeably different directions, indicating that the site of mutation does make a fundamental thermodynamic difference. Such a thermodynamic difference may well be related to a biological difference between the two mutants.

4. Conclusion

In the 2D coordinate system of ΔS and ΔH, the two model peptides Pro-Pro-Thr and Pro-Thr-Pro fell into two mutually exclusive domains with respect to the parent Pro-Pro-Pro triamino acid diamide, 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 compared to $\Sigma\rho_b$. The computed ΔS value revealed that mutations in most of the cases studied led to structural information depletion, implying a disordered state ($\Delta S > 0$): 73% for Pro-Pro-Thr and 58% for Pro-Thr-Pro. In contrast to that, only a relatively few conformers experienced structural information accumulation, implying an ordered state ($\Delta S < 0$) for the Pro → Thr point mutations: 27% for Pro-Pro-Thr and 42% for Pro-Thr-Pro.

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 $\Sigma\rho_b$, may be an indicator for the degree of perturbation caused by the mutation.

This study describes a novel *first-principle* quantum mechanical approach to analyzing a point mutation at the center of a peptide chain and comparing its stability to that of a point mutation at a terminal end. It is hoped that the results herein and those described previously[3] will be useful in future studies to understand the full effects of point mutations on the thermodynamic stability, as a function of the mutation site, in longer peptide models.

Acknowledgment. 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 Príve for his help in the analysis of the PDB results. We also thank Szilárd N. Fejér for his kind help while preparing this manuscript.

Supporting Information Available: Retrieved geometrical parameters from the PDB for the studied tripeptides as well as data obtained from ab initio and DFT quantum mechanical computations and the subsequent molecular structures. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Heinz, D. W.; Baase, W. A.; Matthews, B. W. Folding and function of a T4 lysozyme containing 10 consecutive alanines illustrate the redundancy of information in an amino acid sequence. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89* (9), 3751−3755.

(2) Thoden, J. B.; Timson, D. J.; Reece, R. J.; Holden, H. M. Molecular structure of human galactokinase: Implications for type II galactosemia. *J. Biol. Chem.* **2005**, *280* (10), 9662−9670.

(3) Sahai, M. A.; Viskolcz, B.; Pai, E. F.; Csizmadia, I. G. Quantifying the intrinsic effects of two point mutation models of Pro-Pro diamino acid diamide: A First-Principle Computational Study. *J. Phys. Chem. B* [Online early access]. DOI: 10.1021/jp073471h. Published Online: Sept 8, 2007. http://pubs.acs.org/cgi-bin/abstract.cgi/jpcbfk/asap/abs/jp073471h.html.

(4) 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. First principle computational study on the full conformational space of L-proline diamides. *J. Phys. Chem. A* **2005**, *109* (11), 2660−2679.

(5) Sahai, M. A.; Fejer, S. N.; Viskolcz, B.; Pai, E. F.; Csizmadia, I. G. First-principle computational study on the full conformational space of L-threonine diamide, the energetic stability of cis and trans isomers. *J. Phys. Chem. A* **2006**, *110* (40), 11527−11536.

(6) Siemion, I. Z.; Pedyczak, A.; Burton, J. Conformational preferences of the sequential fragments of the hinge region of the human Iga1 immunoglobulin molecule. *Biophys. Chem.* **1988**, *31* (1−2), 35−44.

(7) Wood, S. G.; Lynch, M.; Plaut, A. G.; Burton, J. Tetrapeptide inhibitors of the Iga1 proteinases from type I Neisseria gonorrhoeae. *J. Med. Chem.* **1989**, *32* (10), 2407−2411.

(8) Chass, G. A.; Sahai, M. A.; Law, J. M. S.; Lovas, S.; Farkas, O.; Perczel, A.; Rivail, J. L.; Csizmadia, I. G. Toward a computed peptide structure database: The role of a universal atomic numbering system of amino acids in peptides and internal hierarchy of database. *Int. J. Quantum Chem.* **2002**, *90* (2), 933−968.

(9) Sahai, M. A.; Lovas, S.; Chass, G. A.; Penke, B.; Csizmadia, I. G. A modular numbering system of selected oligopeptides for molecular computations: using pre-computed amino acid building blocks. *J. Mol. Struct.: THEOCHEM* **2003**, *666−667*, 169−218.

(10) Borics, A.; Chass, G. A.; Csizmadia, I. G.; Murphy, R. F.; Lovas, S. The benefits of a pre-computed amino acid structure database in quantum chemical geometry optimizations of β-turns of peptides. *J. Mol. Struct.: THEOCHEM* **2003**, *666−667*, 355−359.

(11) Chass, G. A.; Sahai, M. A.; Law, J. M. S.; Lovas, S.; Farkas, O.; Perczel, A.; Rivail, J. L.; Csizmadia, I. G. Toward a computed peptide structure database: The role of a universal atomic numbering system of amino acids in peptides and internal hierarchy of database. *Int. J. Quantum Chem.* **2002**, *90* (2), 933−968.

(12) Chasse, G. A.; Rodriguez, A. M.; Mak, M. L.; Deretey, E.; Perczel, A.; Sosa, C. P.; Enriz, R. D.; Csizmadia, I. G. Peptide and protein folding. *J. Mol. Struct.: THEOCHEM* **2001**, *537* (1−3), 319−361.

(13) Kehoe, T. A. K.; Peterson, M. R.; Chass, G. A.; Viskolcz, B.; Stacho, L.; Csizmadia, I. G. The fitting and functional analysis of a double rotor potential energy surface for the *R* and *S* enantiomers of 1-chloro-3-fluoro-isobutane. *J. Mol. Struct.: THEOCHEM* **2003**, *666−667*, 79−87.

(14) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G. Peptide models. 1. Topology of selected peptide conformational potential energy surfaces (glycine and alanine derivatives). *J. Am. Chem. Soc.* **1991**, *113* (16), 6256−6265.

(15) Peterson, M. R.; Csizmadia, I. G. Analysis of the topological features of conformational hypersurface of n-butane. *J. Am. Chem. Soc.* **1978**, *100* (22), 6911−6916.

(16) Peterson, M. R.; Csizmadia, I. G. Analytic equations for conformation energy surfaces. *Prog. Theor. Org. Chem.* **1982**, *3*, 190−266.

(17) Sahai, M. A.; Lovas, S.; Chass, G. A.; Penke, B.; Csizmadia, I. G. A modular numbering system of selected oligopeptides for molecular computations: using pre-computed amino acid building blocks. *J. Mol. Struct.: THEOCHEM* **2003**, *666−667*, 169−218.

(18) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G. Peptide Models .1. Topology of selected peptide conformational potential energy surfaces (glycine and alanine derivatives). *J. Am. Chem. Soc.* **1991**, *113* (16), 6256−6265.

(19) Schrodinger, E. Quantification of the eigen-value problem. *Ann. Phys.* **1926**, *79* (6), 489−527.

(20) Schrodinger, E. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.* **1926**, *28* (6), 1049−1070.

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

(22) Roothan, C. C. New developments in molecular orbital theory. *Rev. Mod. Phys.* **1951**, *23*, 69−89.

(23) Becke, A. D. Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.* **1996**, *104* (3), 1040−1046.

(24) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37* (2), 785−789.

(25) Becke, A. D. Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.* **1996**, *104* (3), 1040−1046.

(26) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37* (2), 785−789.

(27) 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.

(28) *Origin*, version 7.5; OriginLab Corporation: Northampton, MA, 2003.

(29) Tang, T. H.; Deretey, E.; Jensen, S. J. K.; Csizmadia, I. G. Hydrogen bonds: relation between lengths and electron densities at bond critical points. *Eur. Phys. J. D* **2006**, *37* (2), 217−222.

(30) Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ithaca, NY, 1960.

(31) Pauling, L. Atomic radii and interatomic distances in metals. *J. Am. Chem. Soc.* **1947**, *69* (3), 542−553.

(32) Agmon, N. Estimation of the hydrogen-bond lengths to H_3O^+ and $H_5O_2^+$ in liquid water. *J. Mol. Liq.* **1997**, *73*−*74*, 513−520.

(33) Fejer, S. N.; Csizmadia, I. G.; Viskolcz, B. Thermodynamic functions of conformational changes: Conformational network of glycine diamide folding, entropy lowering, and informational accumulation. *J. Phys. Chem. A* **2006**, *110* (49), 13325−13331.

(34) Viskolcz, B.; Fejer, S. N.; Csizmadia, I. G. Thermodynamic functions of conformational changes. 2. Conformational entropy as a measure of information accumulation. *J. Phys. Chem. A* **2006**, *110* (10), 3808−3811.