ELSEVIER

Journal of Molecular Structure (Theochem) 455 (1998) 275–301

THEO CHEM

Characteristics of Ramachandran maps of L-alanine diamides as computed by various molecular mechanics, semiempirical and ab initio MO methods. A search for primary standard of peptide conformational stability☆

Ana M. Rodríguez[a,*], Héctor A. Baldoni[a], Fernando Suvire[a], Rodolfo Nieto Vázquez[a], Graciela Zamarbide[a], Ricardo D. Enriz[a], Ödön Farkas[b,c], András Perczel[b,d], Michael A. McAllister[e], Leslie L. Torday[f], Julius G. Papp[f], Imre G. Csizmadia[g]

[a]*Department of Chemistry, National University of San Luis, Chacabuco 971, 5700 San Luis, Argentina*
[b]*Department of Chemistry, Loránd Eötvös University, Pázmány Péter sétány 2, 1117 Budapest, Hungary*
[c]*Department of Chemistry, Wayne State University, Detroit, MI 48202, USA*
[d]*Department of Biochemistry, University of Oxford, Oxford OX1 2HG, UK*
[e]*Department of Chemistry, P.O. Box 305070, University of North Texas, Denton, TX 76203, USA*
[f]*Department of Pharmacology and Pharmacotherapy, Albert Szent-Györgyi Medical University, H-6701 Szeged, Hungary*
[g]*Department of Chemistry, University of Toronto, 80 St George str, Toronto, Ontario, Canada M5S 1A1*

Received 15 March 1998; accepted 15 April 1998

Abstract

The optimized geometries and relative energies obtained by four force field and two semi-empirical methods were compared with ab initio results computed for formyl-L-alaninamide. Not all methods yielded the same number of minimum energy conformers. Furthermore, while the optimized geometries of the conformers found were comparable, the computed relative energies varied substantially. Also, the force field calculations produced Ramachandran maps that did not even have the appearance of the ab initio Ramachandran map.

Correlating the ab initio relative energies (ΔE) or free energy (ΔG) with the log of relative populations, $\ln(p_x/p_{\gamma_L})$, led to linear relationships from which four conformers deviated; two of them (α_L and ε_L) were overly destabilized and two of them (γ_L and γ_D) were over-stabilized. It is suggested that, after such deviations are corrected, a primary standard may be obtained that might be useful in further investigations related to force-field parametrization as well as protein folding.

Keywords: Empirical force fields; Semi-empirical and ab initio MO methods; Alanine diamides; HCONH–CHMe–$CONH_2$; MeCONH–CHMe–CONHMe

1. Introduction

The Ramachandran map, associated with the conformational behaviour of an amino acid residue,

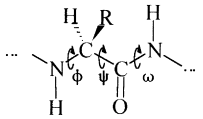
☆ In honor of Professor Árpád Kucsman on the occasion of his 70th birthday.

* Corresponding author.

0166-1280/98/$ - see front matter
PII: S0166-1280(98)00221-8

within a peptide or a protein molecule, is one of the corner stones on which peptide and protein folding is defined. Three torsional angles (ϕ, ψ and ω), as shown in **I**, define a potential energy hypersurface [1] which is associated with a peptide residue

$$E = E(\phi, \psi, \omega) \quad (1)$$



I

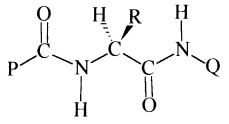
For *trans*-peptide bonds, ω = 180°, therefore the potential energy hypersurface (PEHS) is reduced to a conventional potential energy surface (PES)

$$E = E(\phi, \psi) \quad (2)$$

where the energy is a function of two independent variables: ϕ and ψ Eq. (2). When Eq. (2) is plotted in terms of energy contours, the graphical representation of this PES is normally referred to as the Ramachandran map of **I**.

Numerous force fields have been proposed for the conformational analysis of peptides and proteins. The success of these molecular mechanics packages is based on their different parametrizations which, in turn, are expected to lead to different results. Some comparisons have already been published [1,2] but they are far from being exhaustive.

Semi-empirical MO methods, such as MINDO/3 [3], MNDO [4], AM1 [5] and PM3 [6] have been used extensively to study molecules that are too large for ab initio investigations. However, they are not the generally accepted research tool for peptide conformational analysis.



II

Ab initio computations have been carried out on small peptides (**II**) with P = Q = –H or P = Q = $-CH_3$, for glycine R = –H [2,7,8], as well as for alanine R = $-CH_3$ [2,7,8], valine R = $-CH(Me)_2$ [9], serine R = $-CH_2OH$ [10] and phenylalanine R = $-CH_2Ph$ [11].

In most of these previous studies, computations were carried out at the RHF/3-21G level of theory although extensive basis set computations with and without the inclusion of electron correlation have also been reported for R = –H, –CH3 with P = Q = –H [12,13].

In the present paper we wish to compare the results obtained by a variety of empirical molecular mechanics (MM) and semiempirical MO methods with those obtained by ab initio methods in the past, with and without the inclusion of electron correlation.

2. Methods

In the molecular mechanics (MM) approach, the following force fields were used, as adapted by HYPERCHEM [14]: MM+ [15], AMBER [16], BIO+ (CHARMM) [17], and OPLS [18].

In the semiempirical MO approach, AM1 [5] and PM3 [6] were used as adapted by MOPAC 6.0 [19]. Computations were carried out with MM corrections (MMOK) and without MM correction (NOMM).

Most of the ab initio computed results were taken from previously published papers [2,12,20]. These results were obtained at three levels of theory: RHF/3-21G, RHF/6-311++G(d, p) and MP2/6-311++G(d, p). However, some additional geometries needed to be computed. These computations were carried out at the RHF/3-21G level of theory using Gaussian 94 [21].

3. Results and discussion

3.1. Molecular energetics of conformations

It has been demonstred in numerous papers, as has been reviewed recently [22], that the Ramachandran map for a single L-amino acid diamide has nine discrete conformers. These are labelled as α_D, α_L, β_L, δ_D, δ_L, ε_D, ε_L, γ_D, γ_L and they form a certain pattern as shown in Fig. 1. Earlier ECEPP/2 force field calculations [2] showed the existence of these nine minima. The ab initio potential energy surface (PES) shows that the α_L and ε_L conformations have been annihilated due to the presence of a high mountain ridge which is oriented more or less along the disrotatory

mode of motion as may be seen in the contour diagram (Fig. 2) or the pseudo 3D-representation (Fig. 3) of the Ramachandran landscape calculated for HCO–L–Ala–NH_2 [20].

The optimized torsional angles (ϕ, ψ) with the computed relative energies obtained by the four different MM (MM+, AMBER, BIO+ and OPLS), as well as the ECEPP results reported earlier [2], and two semiempirical MO (AM1 and PM3) computational results, without and with MM corrections for HCONH–CHMe–$CONH_2$, are summarized in Table 1. Ab initio results, obtained previously [2,12] at three levels of theory [RHF/3-21G, RHF/6-311 + + G(d,p) and MP2/6-311 + + G(d,p)], are also included in Table 1 for the sake of comparison. Table 2 shows similar conformational and energetic data obtained for MeCONH–CHMe-CONHMe.

The empirical force field (MM) surfaces for MeCONH–CHMe–CONHMe are shown in Fig. 4. Three PES(s) out of the four force field methods used, three PES (MM+, AMBER and BIO+) did not result in smooth E(ϕ, ψ) functions. However, the fourth PES, calculated by OPLS, looks more like as expected (Fig. 4) with smooth features similar to the relevant ab initio surface (Fig. 3). These discrepancies exhibited by the empirical PES, are also present in the data reported in Tables 1 and 2. While the ECEPP, AMBER, BIO+, OPLS methods located all 9 backbone minima, the MM+ could find only 7 stationary points. The α_L and δ_L minima were annihilated in the case of formyl alanine amide (Fig. 4 and Table 1). However, in the case of acetyl alanine methylamide (Table 2) all four methods located all nine conformers. Overall, we noticed that the number of the zero gradient points differed as a function of the applied force field. We also noticed that the various methods did not have the same conformer as their global minima. As reported in Table 1 for the HCONH–CHMe–$CONH_2$ molecule, MM+ and ECEPP/2 had the γ_L, AMBER and OPLS had the α_L and BIO+ had the δ_D as the most stable backbone structure. Similarly, for the MeCONH–CHMe–NHMe model compound, the MM+ and

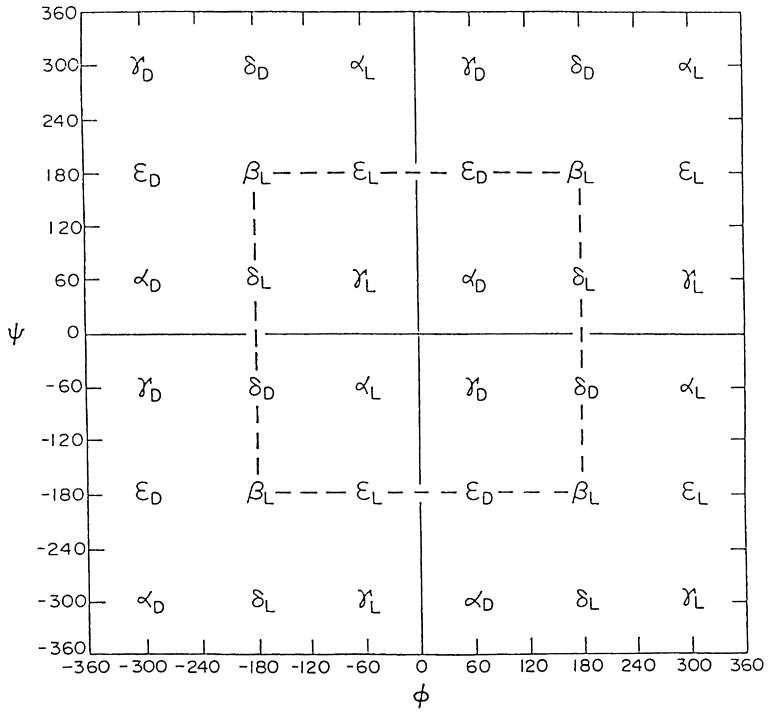


Fig. 1. Idealized PES topology for a single amino acid residue involving two complete cycles of rotation in both ϕ and ψ leading to four equivalent quadrants. Location of the minima are specified by their names in terms of subscripted greek letters. The central square, marked by broken lines, specifies the cut of the PES accepted by IUPAC convention: $-180^o \leq \phi \leq +180^o$ and $-180^o \leq \psi \leq +180^o$.

Table 1
Molecular mechanics, semiempirical MO and ab initio SCF results computed for $HCONH–CHMe–CONH_2$

		ϕ	ψ	ΔE^a (kcal/mol)
		Molecular mechanics		
MM+				
	α_D	61.5	31.6	2.53
	α_L		not found	
	β_L	-158.4	177.7	4.46
	δ_D	-179.2	-58.4	3.82
	δ_L		not found	
	ε_D	72.1	-171.8	5.60
	ε_L	-90.9	166.8	4.33
	γ_D	73.3	-59.7	1.23
	γ_L	**-87.6**	**42.8**	**0.00**
AMBER				
	α_D	56.3	55.2	0.46
	α_L	**-64.3**	**-58.9**	**-0.18**
	β_L	-163.1	177.5	0.46
	δ_D	-165.6	-59.8	0.22
	δ_L	-163.8	67.7	0.37
	ε_D	61.5	-176.7	1.22
	ε_L	-71.0	171.4	0.26
	γ_D	71.9	-64.0	0.96
	γ_L	-80.2	70.9	0.00
BIO+				
	α_D	57.6	59.7	4.11
	α_L	-91.5	-36.9	-0.78
	β_L	-139.2	160.0	-0.22
	δ_D	**-140.6**	**-37.5**	**-0.93**
	δ_L	-123.0	122.1	-0.15
	ε_D	62.1	167.2	3.36
	ε_L	-91.0	154.9	-0.23
	γ_D	64.6	-60.9	2.41
	γ_L	-97.5	115.9	0.00
OPLS				
	α_D	58.1	62.9	1.23
	α_L	**-75.2**	**-55.4**	**-0.34**
	β_L	-156.7	178.8	0.03
	δ_D	-158.9	-57.5	-0.19
	δ_L	-157.9	68.1	-0.01
	ε_D	66.1	-174.8	2.16
	ε_L	-83.7	167.2	-0.16
	γ_D	76.3	-68.4	2.42
	γ_L	-92.0	77.2	0.00
ECEPP/2				
	α_D	54.9	46.3	2.73
	α_L	-73.3	-35.6	1.14
	β_L	-154.8	157.7	1.00
	δ_D	-159.5	-57.9	2.29
	δ_L	-158.5	47.8	1.29
	ε_D	64.5	-178.0	5.64
	ε_L	-77.1	146.4	1.51
	γ_D	79.9	-63.3	8.27
	γ_L	**-80.5**	**76.2**	**0.00**
		Semiempirical MO without MM correction		
AM1				
	α_D		not found	
	α_L		not found	
	β_L		not found	
	δ_D	-118.3	-54.9	3.19
	δ_L		not found	
	ε_D		not found	
	ε_L	-112.9	145.2	1.49
	γ_D	75.1	-61.8	0.54
	γ_L	**-82.2**	**65.7**	**0.00**
PM3				
	α_D	61.5	41.7	2.05
	α_L	-92.7	-52.2	0.43
	β_L		not found	
	δ_D	-151.8	-59.8	0.26
	δ_L		not found	
	ε_D	63.0	-169.1	0.89
	ε_L	**-75.2**	**145.1**	**-2.63**
	γ_D	54.5	-78.8	1.38
	γ_L	-61.6	87.9	0.00

ECEPP/2 had the γ_L, AMBER and OPLS had the α_L and BIO + had the δ_L as the most stable orientation.

The two semiempirical MO surfaces generated without and with MM corrections are shown in Figs. 5 and 6 respectively. Some of them are similar to the relevant ab initio surface (Fig. 3). However, not all of these minima are located at the expected ϕ, ψ values and some minima were clearly annihilated. On the other hand, the geometry optimization results, reported in Tables 1 and 2, revealed that the AM1 and PM3 approaches yielded geometrical parameters comparable to those of ab initio computations.

The energetics, summarized in Tables 1 and 2, are illustrated graphically in Figs. 7 and 8 for *N*-formyl-L-alaninamide and *N*-acetyl-L-alaninmethylamide respectively. As it has been noted previously [2,7], the order of energy levels obtained by MM is not necessarily parallel with the ab initio results. All ab initio results reported so far [2,7,8,12], indicated that the γ_L conformer is the global minimum. We may

Table 1 (*continued*)

	ϕ	ψ	ΔE[a] (kcal/mol)
	Semiempirical MO with MM correction		
AM1			
α_D		not found	
α_L	-107.4	-26.4	3.06
β_L	-120.5	134.5	1.57
δ_D	-120.3	-58.2	3.20
δ_L		not found	
ε_D		not found	
ε_L		not found	
γ_D	74.9	-64.2	0.54
γ_L	**-81.9**	**67.8**	**0.00**
PM3			
α_D	60.3	36.4	3.31
α_L		not found	
β_L		not found	
δ_D	-136.1	-62.7	0.61
δ_L		not found	
ε_D		not found	
ε_L	**-89.6**	**146.2**	**-2.28**
γ_D	72.1	-81.3	0.69
γ_L	-70.7	67.8	0.00
	Ab initio		
HF/3-21G			
α_D	63.8	32.7	5.95
α_L		not found	
β_L	-168.3	170.6	1.25
δ_D	-178.6	-44.1	7.31
δ_L	-128.1	29.8	3.83
ε_D	67.2	-171.9	8.16
ε_L		not found	
γ_D	74.0	-57.4	2.53
γ_L	**-84.5**	**67.3**	**0.00** [b]
HF/6-311++G(d,p)			
α_D	69.0	26.9	4.56
α_L		not found	
β_L	-155.1	161.0	0.11
δ_D	-165.2	-42.1	5.39
δ_L	-112.8	13.2	2.22
ε_D		not found	
ε_L		not found	
γ_D	75.3	-55.4	2.54
γ_L	**-86.2**	**78.8**	**0.00** [c]
MP2/6-311++G(d,p)			
α_D	63.1	35.5	3.88
α_L		not found	
β_L	-157.1	163.2	1.21
δ_D	-166.0	-39.9	5.45
δ_L		not found	
ε_D		not found	
ε_L		not found	
γ_D	74.4	-49.1	2.19
γ_L	**-82.8**	**80.6**	**0.00** [d]

[a] Global minima are printed in bold.

[b] E (total) = - 412.474780 hartree

[c] E (total) = - 414.909271 hartree

[d] E (total) = - 416.384696 hartree

now conclude that the same inconsistency applies for the semiempirical MO calculations as for the MM approaches.

The reason why certain MM methods favour the a_L conformer (which is not even a minimum on the ab initio PES) is simply due to the fact that these empirical force fields are catering for the right-handed α-helical backbone orientation, i.e. $(\alpha_L)_n$, which is the most frequently occurring secondary structural motif in proteins.

It might be added, parenthetically, that, as has been pointed out previously [12], the increase of basis set size without the inclusion of electron correlation is not a recommended practice. It is clear from Fig. 7 that for certain conformations of HCONH–CHMe–$CONH_2$ the relative stabilities at the HF/6-311++G(d,p) level are low with respect to the values obtained at the MP2/6-311++G(d,p) level of theory. The corresponding energetics, associated with MeCONH–CHMe–CONHMe, are shown in Fig. 8.

In order to ever hope to obtain, at least, a nearly perfect force field, which is a requirement for the correct study of protein folding, it would be desirable to compare the computed results to some kind of a primary standard. However, no such primary standard exists at this time.

3.2. Population of molecular conformations

Let's assume that in proteins, conformer (x) is populated (p_x) according to its relative stability. Choosing γ_L as the reference conformer, the relative population (p_x/p_{γ_L}) may be assumed to be related to the relative stabilities (Eq. (3)) in a Boltzman type exponential distribution (Eq. (4))

$$\Delta E = E_x - E_{\gamma_L} \tag{3}$$

$$\left(\frac{p_x}{p_{\gamma_L}}\right) = e^{+\frac{\Delta E}{m}} \cdot e^{-\frac{b}{m}} \tag{4}$$

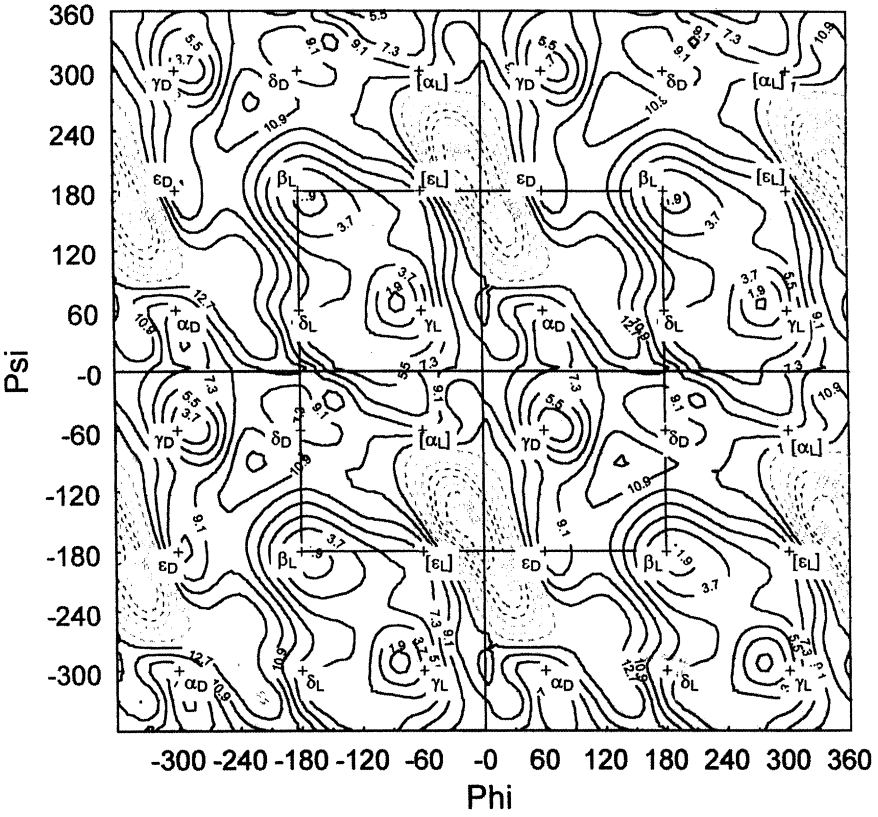


Fig. 2. Energy contour diagram of the Ramachandran map computed for HCONH–CHMe–$CONH_2$ at the HF/3-21G level of theory showing two complete cycles of rotation in both ϕ and ψ.

or its equivalent logarithmic form (Eq. (5))

$$\Delta E = m.\ln\left(\frac{p_x}{p_{\gamma_L}}\right) + b \quad (5)$$

Of course the more negative is ΔE the higher the value of $\log(p_x/p_{\gamma_L})$. This implies that m is expected to be a negative quantity.

3.3. Single amino acid diamides

The populations of the nine discrete conformations have recently been published [2], based on a study of 73 proteins containing 11 793 amino acid residues. These data, essential for the present study, are given in Table 3. Fig. 9 shows how the MM relative energy values

Table 2
Molecular mechanics, semiempirical MO and ab initio SCF results computed for MeCONH–CHMe–CONHMe

		ϕ	ψ	ΔE[a] (kcal mol^{-1})
Molecular mechanics				
MM +				
	α_D	62.0	58.8	8.64
	α_L	–63.0	–58.6	8.05
	β_L	–165.0	179.8	8.23
	δ_D	–179.3	–58.7	7.41
	δ_L	–179.3	60.9	7.75
	ε_D	71.0	–174.0	10.04
	ε_L	–58.7	179.7	10.73
	γ_D	61.7	–58.6	4.42
	γ_L	**–60.0**	**59.3**	**0.00**
AMBER				
	α_D	58.9	60.4	–1.77
	α_L	**–56.8**	**–62.2**	**–2.73**
	β_L	–179.3	–177.9	–0.81
	δ_D	179.2	–59.5	–1.89
	δ_L	–177.4	65.5	–1.73
	ε_D	62.5	–177.4	–1.79
	ε_L	–62.3	178.2	– 1.85
	γ_D	58.1	–57.4	4.97
	γ_L	–60.8	59.6	0.00
BIO +				
	α_D	61.4	55.7	–7.14
	α_L	–65.2	–54.1	–5.49
	β_L	–179.9	–177.9	–6.75
	δ_D	–179.8	–59.1	–3.03
	δ_L	**–178.1**	**64.7**	**–10.57**
	ε_D	63.0	168.8	–1.49
	ε_L	–63.1	178.2	–2.41
	γ_D	57.4	–53.9	–0.98
	γ_L	–62.2	56.9	0.00
OPLS				
	α_D	59.9	59.0	–4.50
	α_L	**–72.8**	**–54.4**	**–7.21**
	β_L	–179.2	–177.9	–4.65
	δ_D	179.6	–59.3	–5.34
	δ_L	–177.3	65.6	–5.34
	ε_D	66.9	–174.0	–4.57
	ε_L	–63.8	178.4	–1.63
	γ_D	60.1	–55.3	–1.02
	γ_L	–62.0	58.9	0.00

computed for MeCONH–CHMe–CONH–Me relate to $\ln(p_x/p_{\gamma_L})$. None of the MM results follow Eq. (5). Furthermore some of the lines between adjacent points do not have negative slopes. However, a segment here and there hints that there might be considerable linearity if the forcefields were optimally parametrized. Such hints might be sufficient to reexamine the question of MM parametrization.

Similar plots were prepared for the results obtained by the two semiempirical methods. These correlations are shown in Fig. 10.

The ab initio relative energies (ΔE), as computed at

Table 2 (*continued*)

		ϕ	ψ	ΔE^a (kcal mol^{-1})
Molecular mechanics				
ECEPP/2				
	α_D	54.7	46.0	2.37
	α_L	–73.6	–34.9	0.81
	β_L	–154.7	157.2	0.71
	δ_D	–158.2	–57.5	1.73
	δ_L	–150.7	45.6	1.10
	ε_D	63.7	–174.8	5.07
	ε_L	–75.5	139.0	1.12
	γ_D	71.7	–64.3	7.24
	γ_L	**–80.4**	**75.8**	**0.00**
Semiempirical MO without MM correction				
AM1				
	α_D		not found	
	α_L		not found	
	β_L		not found	
	δ_D		not found	
	δ_L		not found	
	ε_D		not found	
	ε_L	–106.4	145.2	1.58
	γ_D	76.2	–63.0	0.71
	γ_L	**–83.9**	**68.0**	**0.00**
PM3				
	α_D	63.4	46.0	2.48
	α_L	–90.6	–52.7	0.79
	β_L	–162.5	127.8	–1.06
	δ_D	–154.3	–61.1	0.65
	δ_L		not found	
	ε_D	62.2	–171.1	1.32
	ε_L	**–85.5**	**152.1**	**–1.80**
	γ_D	62.7	–45.8	2.09
	γ_L	–62.6	88.7	0.00
Semiempirical MO with MM corrections				
AM1				
	α_D		not found	
	α_L		not found	
	β_L	–112.5	111.9	1.86
	δ_D	–111.3	–50.8	2.96
	δ_L		not found	
	ε_D		not found	
	ε_L	–111.4	133.6	1.78
	γ_D	76.5	–64.0	0.62
	γ_L	**–83.6**	**69.6**	**0.00**

Table 2 (*continued*)

		ϕ	ψ	ΔE[a] (kcal mol^{-1})
PM3				
	α_D	61.1	39.1	3.48
	α_L	–110.1	–39.2	1.31
	β_L	–144.1	132.6	–1.03
	δ_D	–134.3	–58.6	1.05
	δ_L		not found	
	ε_D	63.2	–161.1	2.37
	ε_L	**–90.2**	**145.6**	**–1.73**
	γ_D	69.9	–60.4	1.02
	γ_L	–74.6	73.9	0.00
Ab initio				
HF/3-21G				
	α_D	63.4	32.7	5.80
	α_L		not found	
	β_L	–167.1	169.4	1.13
	δ_D	–175.9	–45.2	7.24
	δ_L	–127.6	27.8	3.61
	ε_D	64.3	–171.8	7.95
	ε_L		not found	
	γ_D	74.4	–58.2	2.81
	γ_L	**–85.8**	**69.0**	**0.00**[b]

[a]Global minima are printed in bold.
[b]E(total) = –490.1199478 hartree.

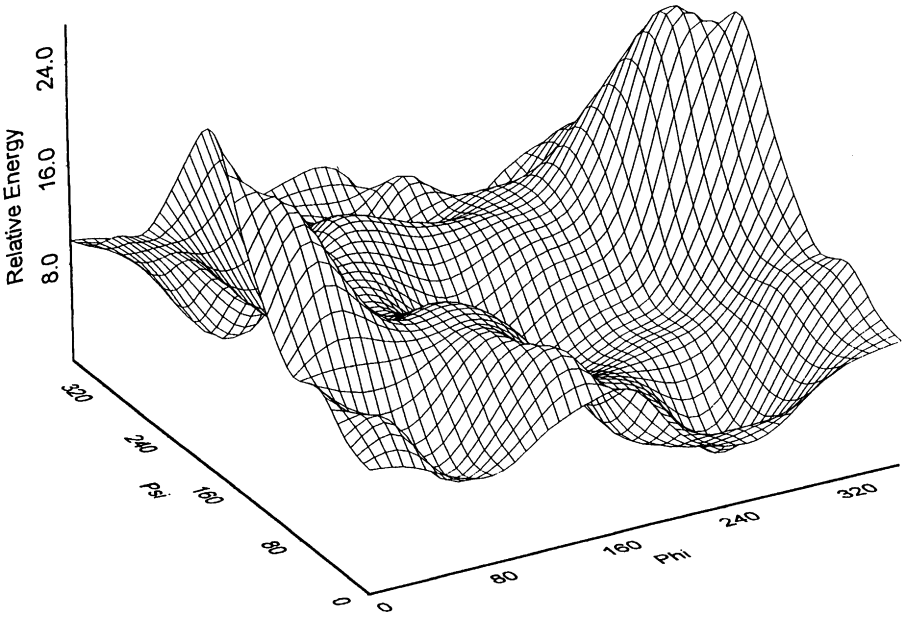


Fig. 3. Pseudo-3D representation of the Ramachandran landscape computed for HCONH–CHMe–$CONH_2$ at the HF/3-21G level of theory. The PES shown corresponds to the upper right hand quadrant of Fig. 1 and Fig. 2.

Table 3
Population of amino acid residue conformations in proteins[a]

Conformation x	p_x	(p_x/p_{γ_L})	$\ln(p_x/p_{\gamma_L})$
α_L	4599	4.8870	+1.59
ε_L	2743	2.9150	+1.07
β_L	1581	1.6801	+0.52
γ_L	941	1.0000	± 0.00
δ_L	691	0.7343	–0.31
α_D	515	0.5473	–0.60
δ_D	314	0.3337	–1.10
ε_D	274	0.2912	–1.23
γ_D	135	0.1435	–1.94

[a]Data taken from [2]. It consists of 11 793 amino acid residues in 73 proteins carefully selected for their accuracy.

the HF/3-21G and MP2/6-311++G(d, p) levels of theory, were also correlated with the log of relative populations: i.e. $\ln(p_x/p_{\gamma_L})$. The results are shown in Fig. 11. The hint of linearity is clearly present here, much more so than was observed in the case of MM and semiempirical results (Figs. 9 and 10). It is clear from Fig. 11 that we need to focus our attention to two sets of conformations: (α_L,ε_L) and (γ_L,γ_D).

In the case of α_L and ε_L, it should be noted that the PES of a single amino acid diamide does not have these conformers as stable minima. Thus, nothing is stabilizing these conformers when they are isolated from the polypeptide chain. However, both the α_L [20,23] and ε_L [23,24] minima reappear if there is a stabilizing effect that originates from outside of the single amino acid residue. We have computed the energy of the α_L and ε_L conformations of HCNH–CHMe–$CONH_2$ at RHF/3-21G level of theory (Table 4), at fixed ϕ and ψ values, and as may be seen from Fig. 11A, they are way above the line fitted to the five points. Thus, in the isolated peptides these conformers are understabilized with respect to the protein environment.

Table 4
Total molecular energies of HCONH–CHR–$CONH_2$ computed at the HF/3-21G level of theory for fixed ϕ and ψ torsional angle values associated with the α_L and ε_L conformations

Conformation	ϕ	ψ	E(HF/3-21G)
α_L	–69.8	–36.8	–412.456752*
ε_L	–74.7	+167.8	–412.458130*

*Interpolated from the grid points of the $E = E(\phi,\psi)$ surface using a cubic spline function.

In contrast to the above, the γ_D and γ_L conformations are overstabilized in their isolated state with respect to the protein environment. In order to mimic their stability in a protein environment, we need to destabilize γ_D by about 8 kcal mol^{-1} and γ_L by about 3 kcal mol^{-1} at the HF/3-21G level of theory (Fig. 11A). Of course, the needed destabilization values are method dependent. For example at MP2/6-311++G(d, p) level of theory, these values turn out to be 5.6 and 2.6 kcal mol^{-1} (Fig. 11B) for γ_D and γ_L

Table 5
Scaled stabilities (ΔE^{scaled}) for the conformations of HCONH–CHMe–$CONH_2$ and MeCONH–CHMe–CONHMe computed at the HF/3-21G level of the theory

Conformation x	HCONH-CHMe-$CONH_2$				MeCONH-CHMe-CONHMe			
	ΔE^{HF} [a]	$\Delta\Delta E^{shift}$	$\Delta E^{HF} + \Delta\Delta E^{shift}$	ΔE^{scaled}	ΔE^{HF} [a]	$\Delta\Delta E^{shift}$	$\Delta E^{HF} + \Delta\Delta E^{shift}$	ΔE^{scaled}
α_L	[11.31][b]	–14.48	–3.13	–6.28				
ε_L	[10.45][b]	–11.53	–1.08	–4.23				
β_L	1.25		1.25	–1.90	1.13		1.13	–1.87
γ_L	0.00	3.15	3.15	0.00	0.00	+3.00	3.00	0.00
δ_L	3.83		3.83	0.68	3.61		3.61	+0.61
α_D	5.95		5.95	2.80	5.80		5.80	+2.80
δ_D	7.31		7.31	4.16	7.24		7.24	+4.24
ε_D	8.16		8.16	5.01	7.95		7.95	+4.95
γ_D	2.53	8.26	10.79	7.64	2.81	+7.83	10.64	+7.64

[a]Computed at the HF/3-21G level of theory.
[b]Optimized with fixed ϕ, ψ angle pairs since these conformations are located at the side of a mountain ridge. Fixed torsional angles were taken from [20]. For α_L: $\phi = -69.8$ and $\psi = -36.8$. For ε_L: $\phi = -74.7$ and $\psi = 167.8$.

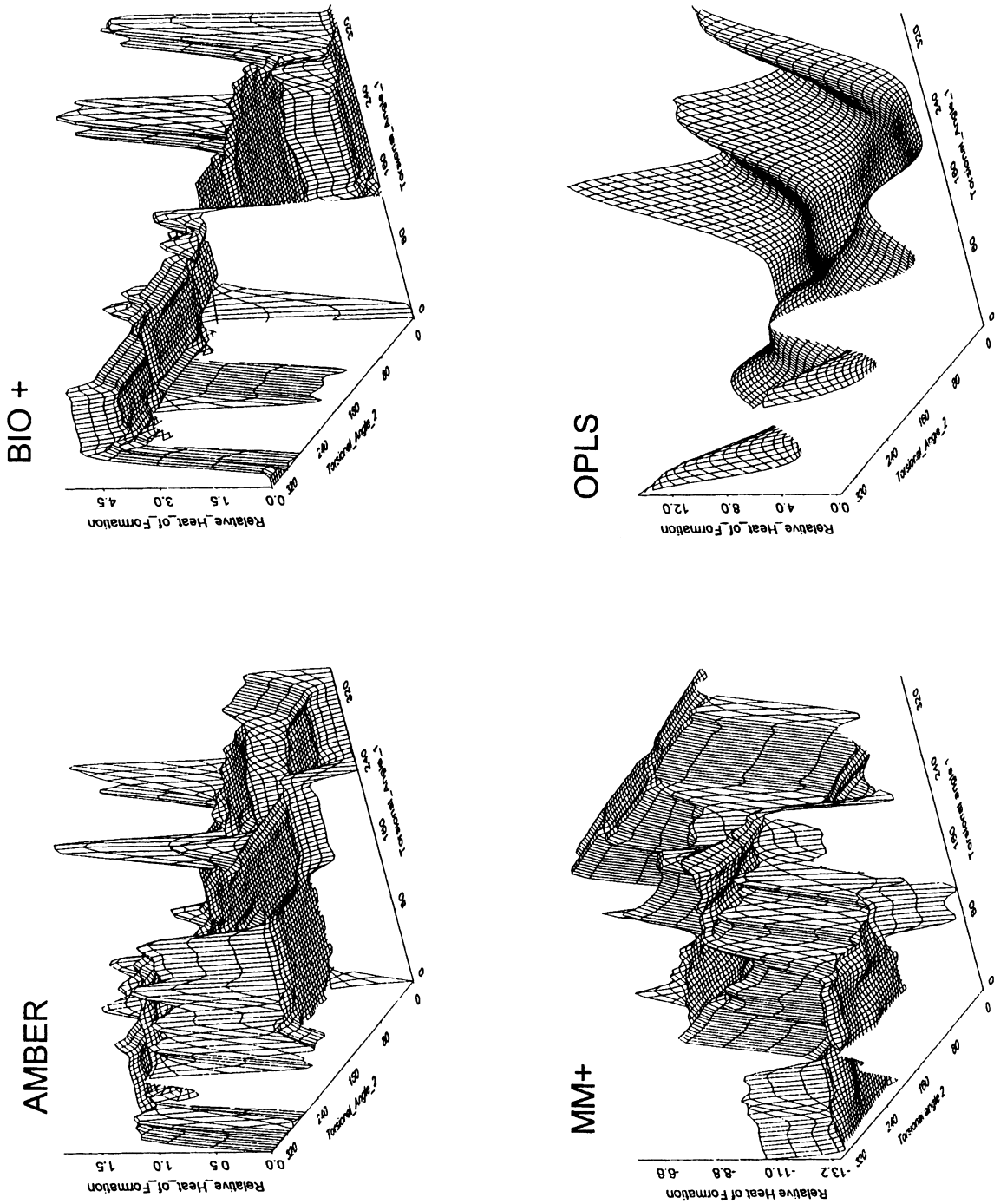


Fig. 4. Ramachandran landscape computed for MeCONH–CHMe–CONHMe using four different force fields [MM + , AMBER, BIO + (CHARMM), OPLS]. Torsional angle 1 is ϕ and torsional angle 2 is ψ.

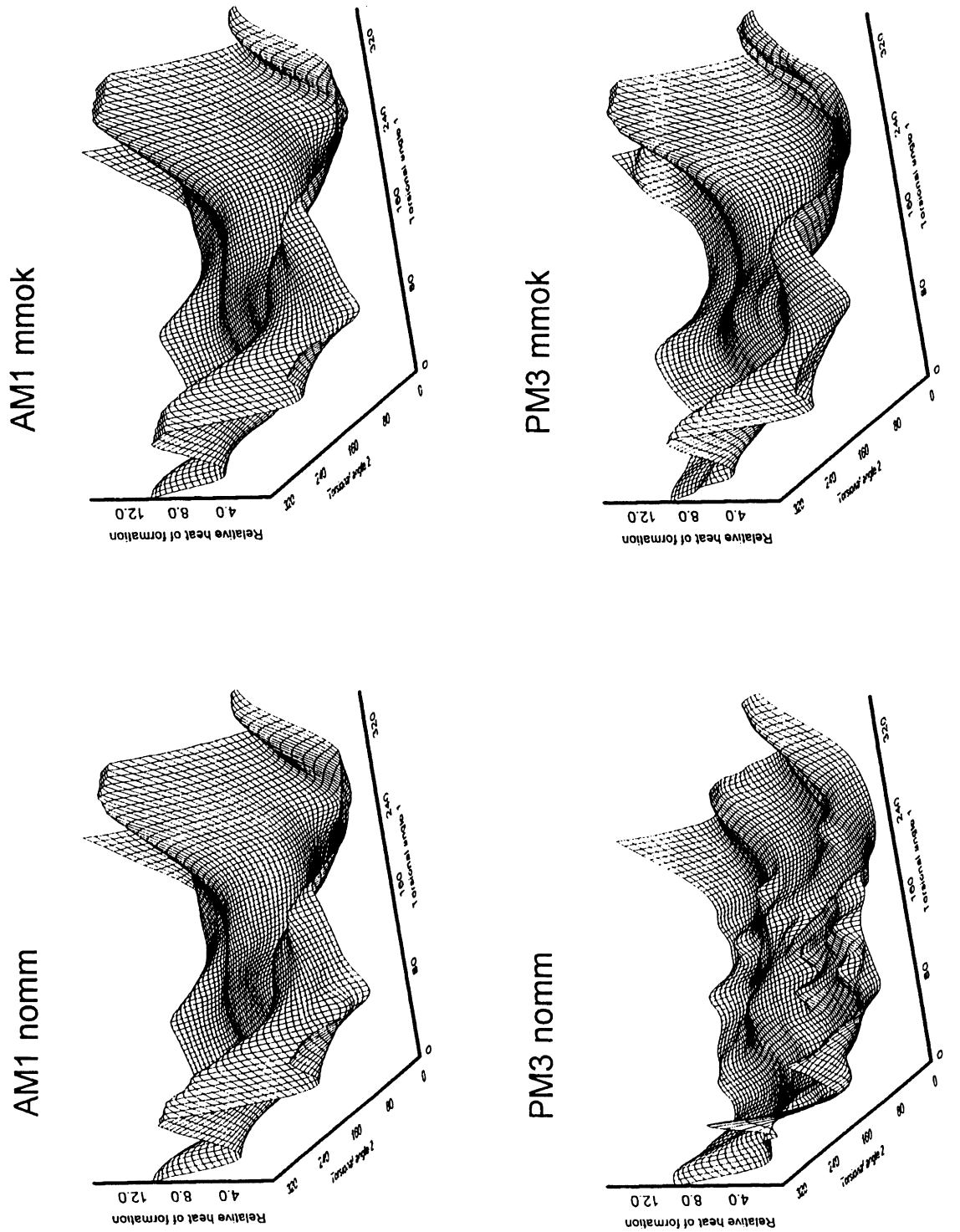


Fig. 5. Ramachandran landscape computed for HCONH–CHMe–$CONH_2$ using two different semiempirical methods (AM1 and PM3) with (MMOK) and without (NOMM) MM correction. Torsional angle 1 is ϕ and torsional angle 2 is ψ.

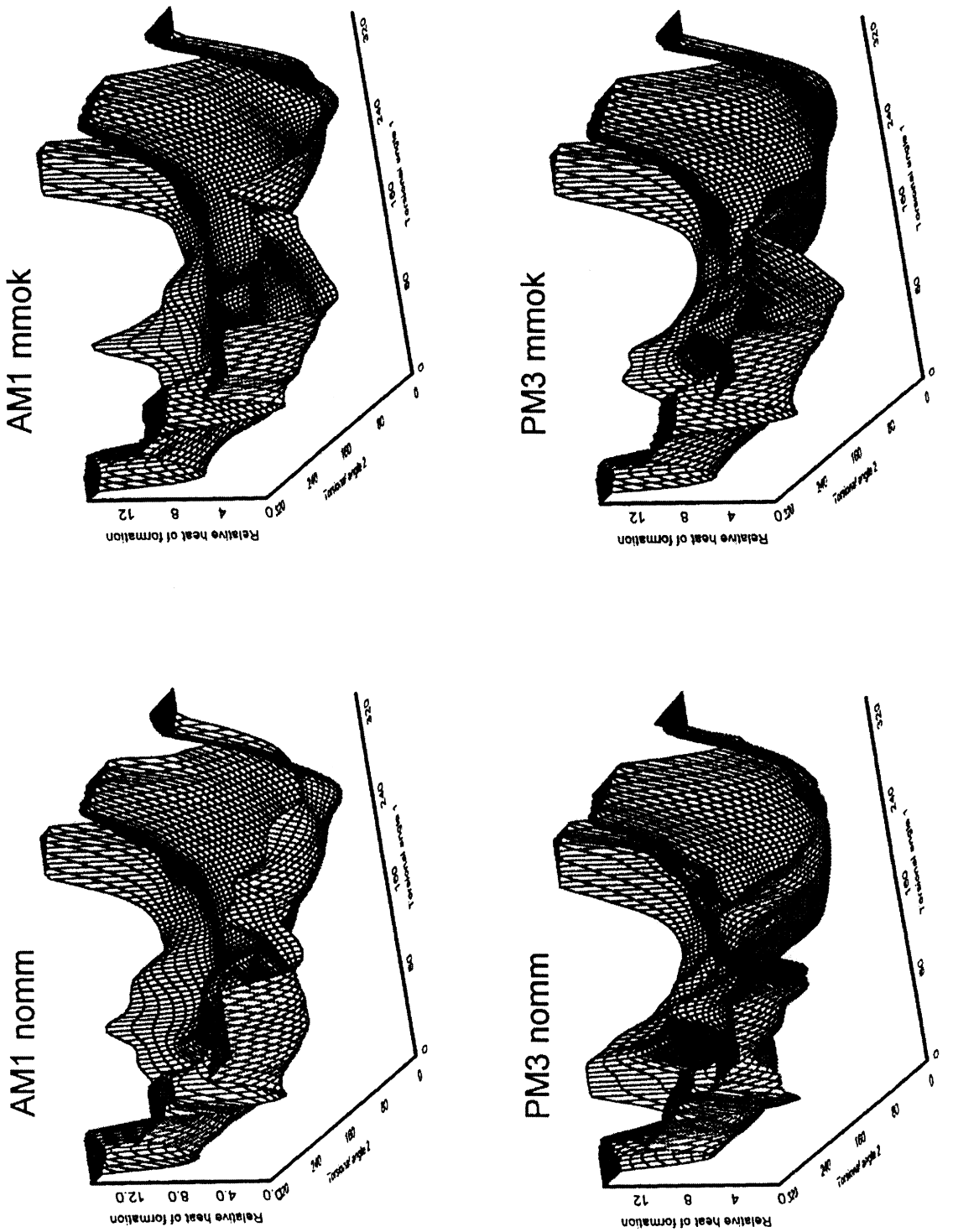


Fig. 6. Ramachandran landscape computed for MeCONH–CHMe–CONHMe using two different semiempirical methods (AM1 and PM3) with (MMOK) and without (NOMM) MM correction. Torsional angle 1 is ϕ and torsional angle 2 is ψ.

respectively. One may wonder why the γ_D and γ_L conformations are stabilized so excessively in these isolated forms of the peptide. Alternatively we may ask what process is responsible for the destabilization of these conformations on going from the isolated amino acid to the protein environment.

If the excessive stabilization in the isolated state is caused by C=O.....H–N type hydrogen bonding, which holds the seven-membered ring together in both the γ_L (C_7^{eq}) and γ_D (C_7^{ax}) conformations, then in proteins the destabilizing process must work against this type of intramolecular hydrogen bonding. At any rate, it is clear that something is different in the case of single amino acid diamides in vacuum concerning these four conformers (γ_D, γ_L, ε_L and α_L) in comparison to other conformers of the amino acid residues in proteins. A similar anomaly is not observed for the remaining five conformations: (β_L, δ_L, α_D, δ_D and ε_L).

Table 6
Scaled stabilities (ΔE^{scaled}) for the conformations of HCONH–CHMe–$CONH_2$ computed at the MP2/6-311++G(d,p) level of theory

	ΔE^{HFa}	$\Delta\Delta E^{shift}$	$\Delta E^{HF} + \Delta\Delta E^{shift}$	ΔE^{scaled}
α_L	[–1.60][b]		–1.60	–4.10
ε_L	[–0.26][b]		–0.26	–2.76
β_L	1.21		1.21	–1.29
γ_L	0.00	2.50	2.50	0.00
δ_L	[3.30][b]		3.30	+0.80
α_D	3.88		3.88	+1.33
δ_D	5.45		5.45	+2.95
ε_D	[5.67][b]		5.67	+3.17
γ_D	2.19	5.31	7.50	+5.00

[a]Computed at the MP2/6 – 311++G(d,p) level of theory.
[b]Obtained by interpolation or extrapolation of the plot given in Fig. 11B.

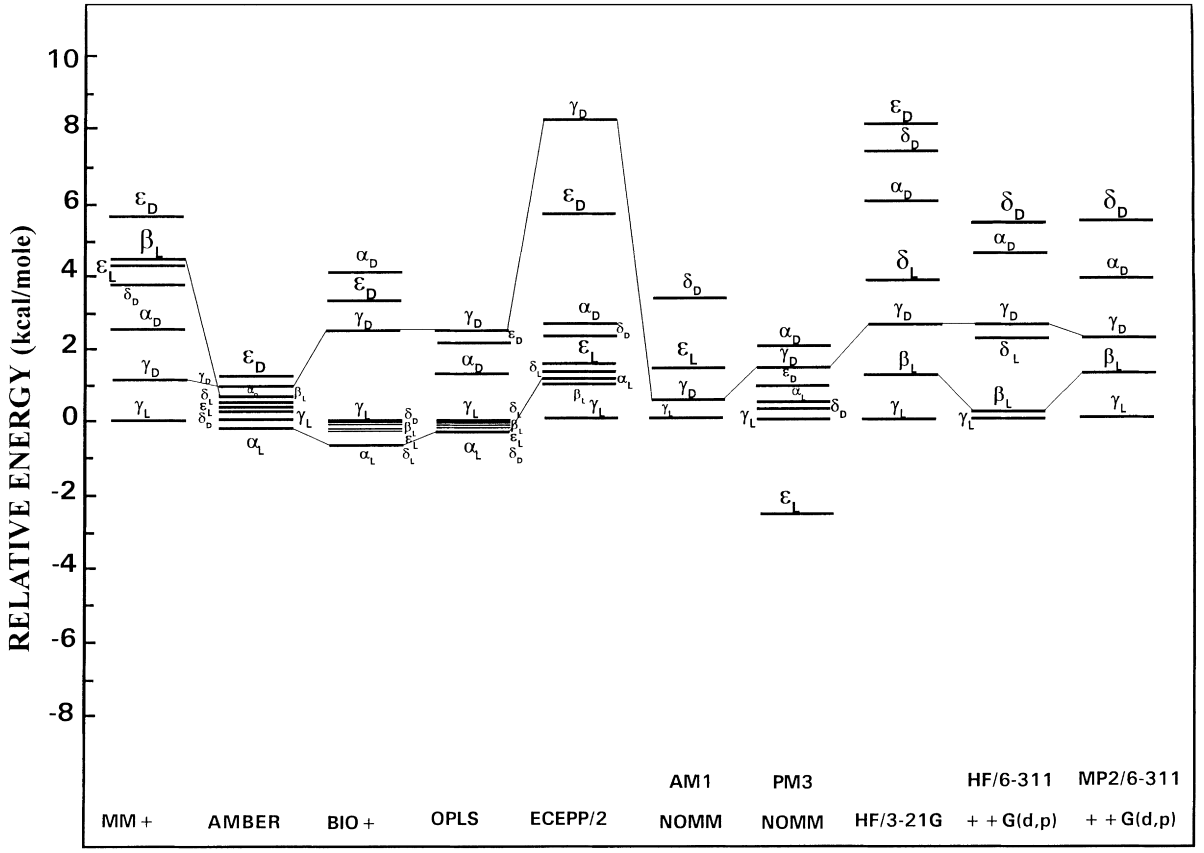


Fig. 7. Relatives energies of alanine diamide: HCONH–CHMe–$CONH_2$ conformations as obtained by various empirical, semi-empirical and non-empirical (ab initio) methods.

If we are willing to shift γ_D, γ_L, ε_L and α_L to the correlated line of slope -3.94 as shown in Fig. 11A, for the RHF/3-21G results, then we have a new set of stability values which are summarized in Table 5. Using this set of unified ΔE^{scaled} values, for the nine conformations, one gets a reasonable straight line (Fig. 12) that correlates ab initio ΔE values, now scaled for the protein environment (ΔE^{scaled}), with the $\ln(p_x/p_{\gamma_L})$ obtained from protein X-ray. A set of ΔE^{scaled} values were also determined on the basis of the MP2/6-311++G(d,p) as summarized in Table 6. In the absence of anything better, we may accept, at least tentatively, the ΔE^{scaled} values, obtained from the RHF/3-21G computation for HCONH–CHMe–$CONH_2$, as the primary standard for conformational stability with respect to which the MM and semiempirical MO results may be assessed.

3.4. Diamino acid diamides

Since the proposed primary standard is based on the $\ln(p_x/p_{\gamma_L})$ values, perhaps one should look for the transferability of these $\ln(p_x/p_{\gamma_L})$ values. One may make an assessment on the basis of the recently published [24] 49 dipeptide (i.e. diamino acid diamide) results. The log relative population of dipeptide conformation *xy*, denoted here as $\ln(p_{xy}/p_{\gamma_L\gamma_L})$ for the 49 cases found by ab initio molecular study [24] are summarized in Table 7. Alternatively we may construct a set of approximate log relative populations to mimic these $\ln(p_{xy}/p_{\gamma_L\gamma_L})$ values according to the following definition

$$\ln\left[\left(\frac{p_x}{p_{\gamma_L}}\right)\left(\frac{p_y}{p_{\gamma_L}}\right)\right] = \ln\left(\frac{p_x}{p_{\gamma_L}}\right) + \ln\left(\frac{p_y}{p_{\gamma_L}}\right) \tag{6}$$

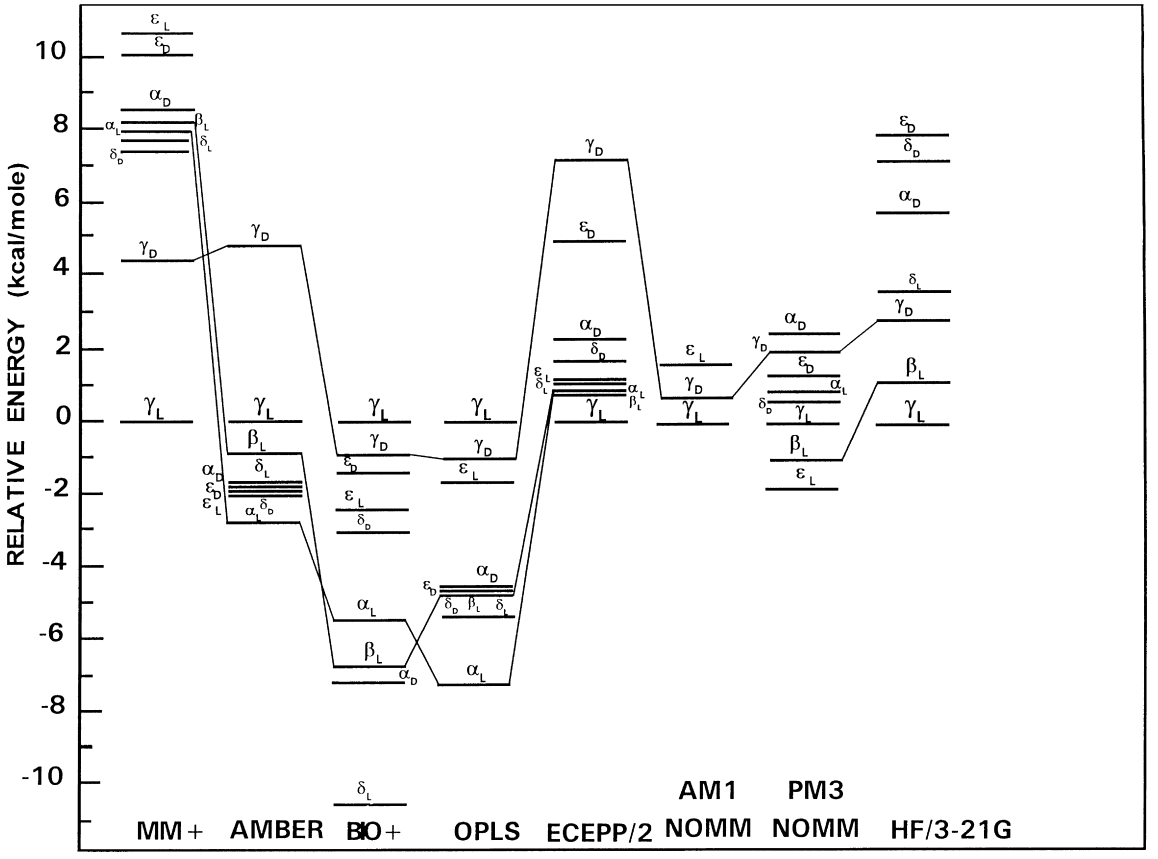


Fig. 8. Relative energies of alanine diamide: MeCONH–CHMe–CONHMe conformations as obtained by various empirical, semi-empirical and non-empirical (ab initio) methods.

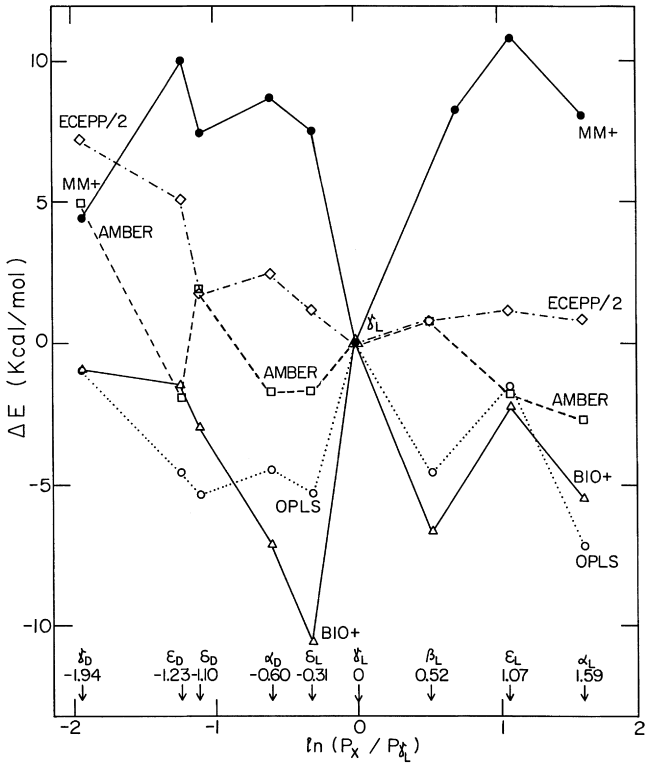


Fig. 9. Correlation of relative energies (ΔE) computed with the use of five different force fields [MM + , AMBER, BIO + (CHARMM), OPLS and ECEPP/2] with the log of relative populations [$\ln(p_x/p_{\gamma_L})$] for MeCONH–CHMe–CONHMe.

The resulting 81 approximate populations are summarized in Table 8. Of course no identity can be expected between the data presented in Tables 7 and 8, but it would be nice to see a linear relationship of the type:

$$\ln\left(\frac{p_{xy}}{p_{\gamma_L\gamma_L}}\right) = M\,\ln\left[\left(\frac{p_x}{p_{\gamma_L}}\right)\left(\frac{p_y}{p_{\gamma_L}}\right)\right] + B \qquad (7)$$

Table 7
Natural log of relative populations, $\ln(p_{xy}/p_{\gamma_L\gamma_L})$, of dipeptide conformations (xy) in proteins[a]

	p_y	α_L	ε_L	β_L	γ_L	δ_L	α_D	δ_D	ε_D	γ_D
p_x										
α_L						0.9628				
ε_L										
β_L		–1.3398		–1.2321	0.3895	–0.8109		–1.1987	–0.9651	–2.4384
γ_L			0.9658	–0.0741	0.00	–0.6168	–1.0986	–1.8918	–1.8405	–2.1972
δ_L				–0.2616	–0.9045		–0.5188	–1.5041	–1.0521	–2.1282
α_D				–0.5458	–1.0521	–1.5041	–0.7587	–2.5337		–2.1282
δ_D			–1.4023	–1.2254	–2.0637		–2.3514	0.1125	– 2.7568	–3.4499
ε_D				–0.8109	–1.5041	–1.7918	–2.5337	–2.8904	–2.0637	–3.0445
γ_D			–0.8299	–2.6391	–2.6391	–2.1282		–3.0445	–3.2268	–3.4499

[a]Although 81 dipeptide conformations are expected to be present in protein, here only those 49 conformations are included which were found by ab initio geometry optimization [24] for HCONH–Ala–Ala–NH_2.

Table 8
Natural log of constructed relative populations, $\ln[(p_x/p_{\gamma_L})(p_y/p_{\gamma_L})]$, of single peptide conformations (x and y) in proteins

	p_y	α_L	εe_L	β_L	γ_L	δ_L	α_D	δ_D	ε_D	γ_D
p_x		1.59	1.07	0.52	0.00	–0.31	–0.60	–1.10	–1.23	–1.94
α_L	1.59	3.18	2.66	2.11	1.59	1.28	0.99	0.49	0.36	–0.35
ε_L	1.07	2.66	2.14	1.59	1.07	0.76	0.47	–0.03	–0.16	–0.87
β_L	0.52	2.11	1.59	1.04	0.52	0.21	–0.08	–0.58	–0.71	–1.42
γ_L	0.00	1.59	1.07	0.52	0.00	–0.31	–0.60	–1.10	–1.23	–1.94
δ_L	–0.31	1.28	0.76	0.21	–0.31	–0.62	–0.91	– 1.41	– 1.54	–2.25
α_D	–0.60	0.99	0.47	–0.08	–0.60	–0.91	–1.20	–1.70	–1.83	–2.54
δ_D	–1.10	0.49	–0.03	–0.58	–1.10	–1.41	–1.70	–2.20	–2.33	–3.04
ε_D	–1.23	0.36	–0.16	–0.71	–1.23	–1.54	–1.83	–2.33	–2.46	–3.17
γ_D	–1.94	–0.35	–0.87	–1.42	–1.94	–2.25	–2.54	–3.04	–3.17	–3.88

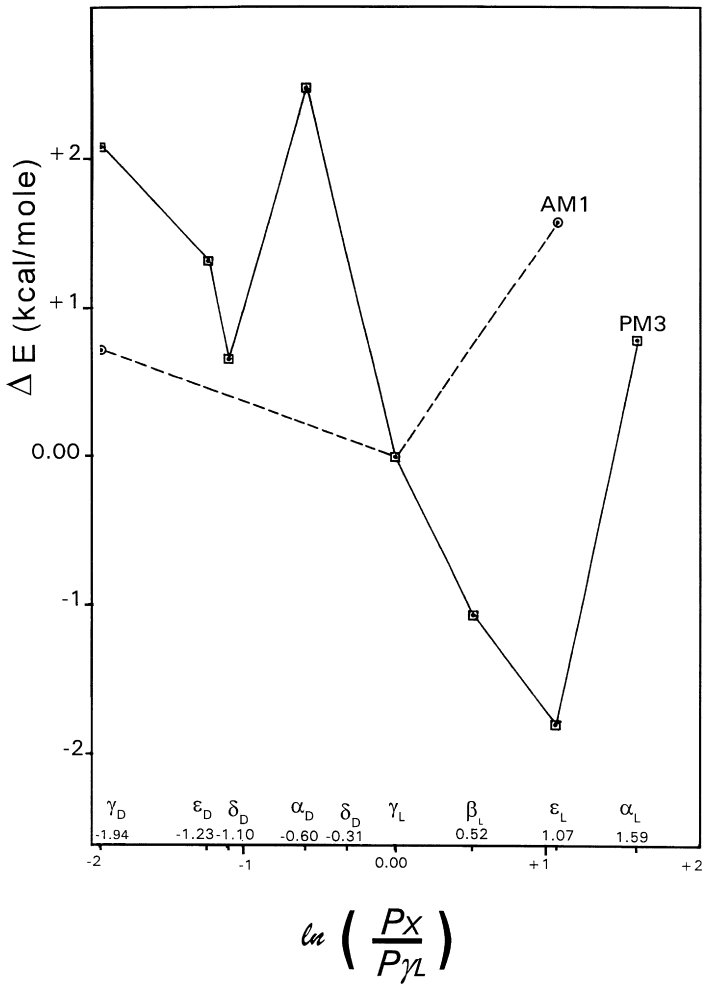


Fig. 10. Correlation of relative energies (ΔE) computed with the use of two different semi-empirical methods (AM1 and PM3) with the log of relative populations $[\ln(p_x/p_{\gamma_L})]$ for MeCONH–CHMe–CONHMe.

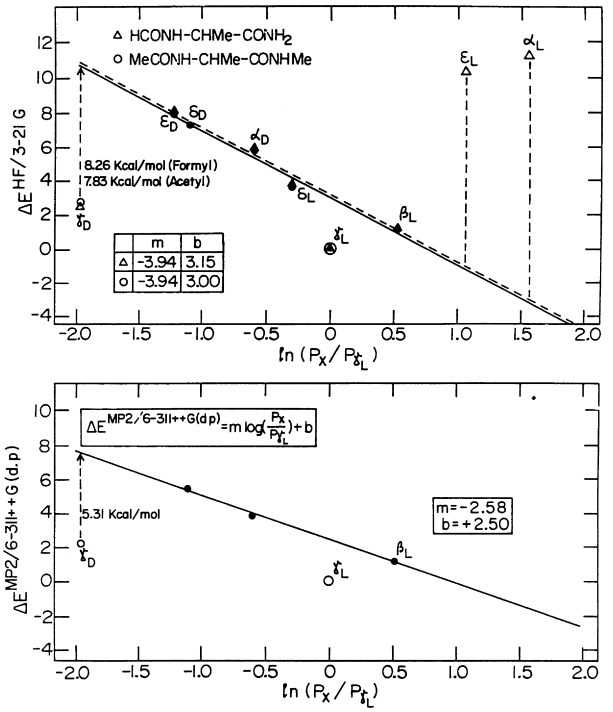


Fig. 11. Correlation of relative energies (ΔE) with $\ln(p_x/p_{\gamma_L})$. The relative stabilities were computed at two ab initio levels of theory [HF/3-21G (upper portion A) and MP2/6-311++G(d,p) (lower portion B)] including HCONH-CHMe- $CONH_2$ and MeCONH-CHMe-$CONH_2$.

The approximate nature of $\ln[(p_x/p_{\gamma_L})(p_y/p_{\gamma_L})]$, as defined in Eq. (4), must be emphasized since it is the same for two amino acids irrespective of their sequence. For example the $\ln(p_{xy}/p_{\gamma_L\gamma_L})$ for $\alpha_D\delta_D$ is –2.534 and for $\delta_D\alpha_D$ it is –2.351 (see Table 9). The approximate value for both of them is –0.60–1.10 = –1.70 as calculated according to Eq. (6). The fact that –1.70 is notably different from both –2.53 and –2.35 is not really an issue since it represents only a shift in the scale. The point here is that on the real population scale $\alpha_D\delta_D$ and $\delta_D\alpha_D$ are different (the difference is 0.2) yet on the approximate scale as defined by Eq. (6) these two values are identical.

Fig. 13 shows that although the correlation between $\ln(p_{xy}/p_{\gamma_L\gamma_L})$ and $\ln[(p_x/p_{\gamma_L})(p_y/p_{\gamma_L})]$ is far from being perfect, the two scales appear to be analogous. Plotting the RHF/3-21G energies [24] of HCONH–CHMe–$CONH_2$ against $\ln(p_{xy}/p_{\gamma_L\gamma_L})$ as well as $\ln[(p_x/p_{\gamma_L})(p_y/p_{\gamma_L})]$ taken from Table 7 and Table 8 respectively, as shown in Fig. 14A and Fig. 14B respectively, clearly reveals that the two population scales are more or less equally useful. Taking a

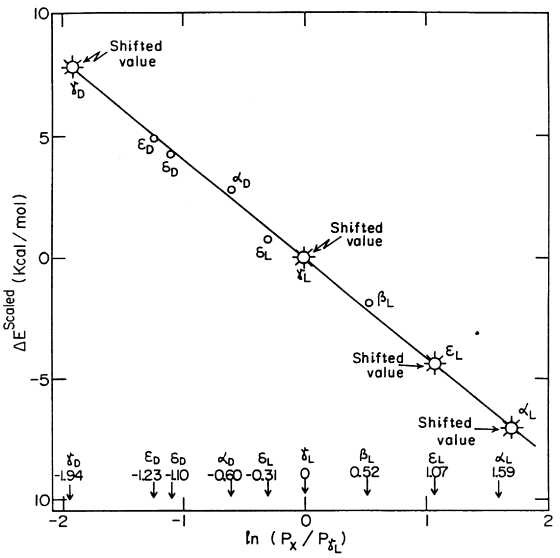


Fig. 12. Correlation of ΔE^{scaled} (derived from HF/3-21G molecular computations) with $\ln(p_x/p_{\gamma_L})$.

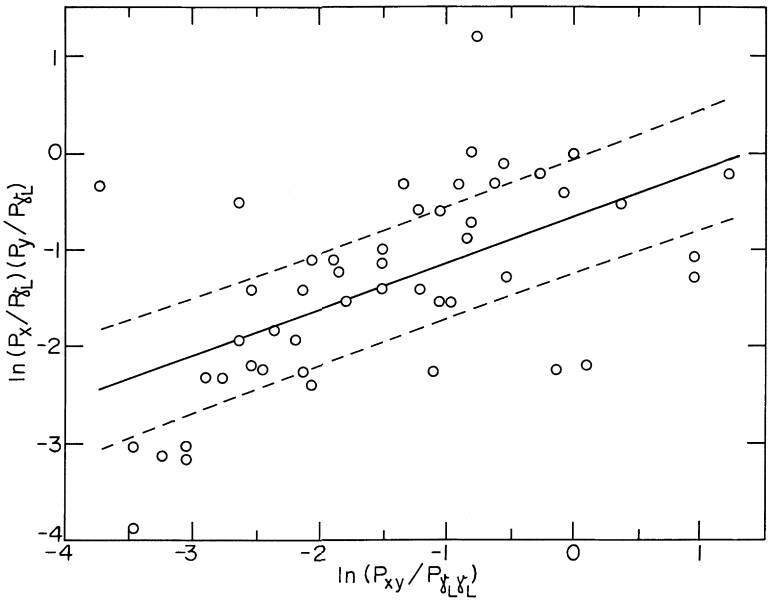


Fig. 13. Interdependence of diamino acid (dipeptide) segment populations in proteins $\ln(p_{xy}/p_{\gamma_L\gamma_L})$ with that generated from single amino acid populations in proteins $\ln[(p_x/p_{\gamma_L})(p_y/p_{\gamma_L})]$ (see Eq. (7)).

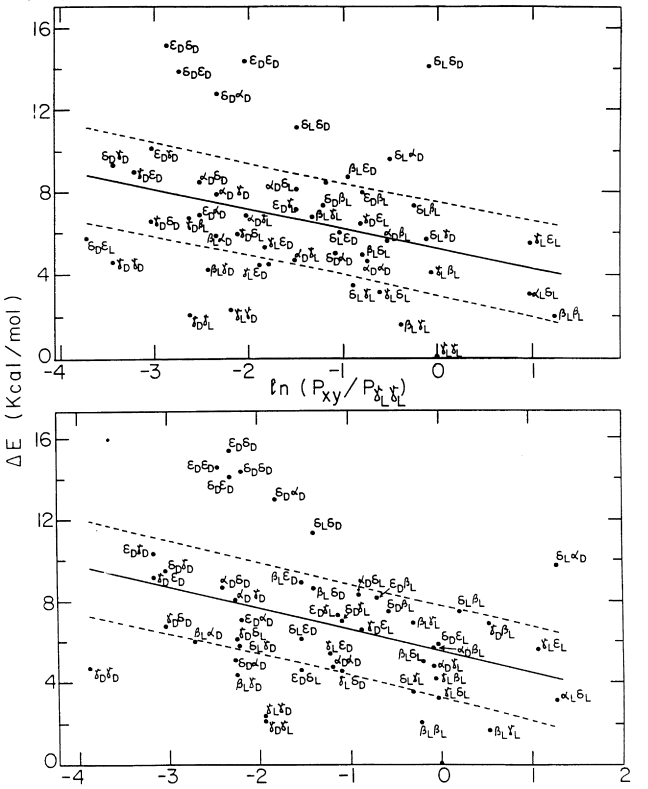


Fig. 14. Correlation of ΔE values computed at the HF/3-21G level of theory for HCONH–Ala–Ala–NH_2 with A (upper) $\ln(p_{xy}/p_{\gamma_L\gamma_L})$ and B (lower) $\ln[(p_x/p_{\gamma_L})(p_y/p_{\gamma_L})]$.

Table 9
Summary of relative populations and stabilities (ΔE) of dipeptides conformations

xy	p_{xy}[a]	$(p_{xy}/p_{\gamma_L\gamma_L})$	$\ln(p_{xy}/p_{\gamma_L\gamma_L})$[b]	ΔE[a]	$\ln[(p_x/p_{\gamma_L})(p_y/p_{\gamma_L})]$[c]
$\alpha_L\,\delta_L$	330	2.6190	0.9628	3.13	1.28
$\alpha_D\,\alpha_D$	59	0.4683	–0.7587	4.73	–1.20
$\alpha_D\,\beta_L$	73	0.5794	–0.5458	5.69	–0.08
$\alpha_D\,\delta_D$	10	0.0794	–2.5337	8.64	–1.70
$\alpha_D\,\delta_L$	28	0.2222	–1.5041	8.29	–0.91
$\alpha_D\,\gamma_D$	15	0.1190	–2.1282	8.04	–2.54
$\alpha_D\,\gamma_L$	44	0.3492	–1.0521	4.78	–0.60
$\beta_L\,\alpha_L$	33	0.2619	–1.3398	6.91	2.11
$\beta_L\,\beta_L$	432	3.4286	1.2321	2.05	1.04
$\beta_L\,\delta_D$	38	0.3016	–1.1987	8.59	–0.58
$\beta_L\,\delta_L$	56	0.4444	–0.8109	5.03	0.21
$\beta_L\,\varepsilon_D$	48	0.3810	–0.9651	8.87	–0.71
$\beta_L\,\gamma_D$	11	0.0873	–2.4384	4.34	–1.42
$\beta_L\,\gamma_L$	186	1.4762	0.3895	1.65	0.52
$\varepsilon_D\,\alpha_D$	10	0.0794	–2.5337	7.04	–1.83
$\varepsilon_D\,\beta_L$	56	0.4444	–0.8109	8.41	–0.71
$\varepsilon_D\,\delta_D$	7	0.0555	–2.8904	15.37	–2.33
$\varepsilon_D\,\delta_L$	21	0.1666	–1.7918	4.85	–1.54
$\varepsilon_D\,\varepsilon_D$	16	0.1270	–2.0637	14.57	–2.46
$\varepsilon_D\,\gamma_D$	6	0.0476	–3.0445	10.31	–3.17
$\varepsilon_D\,\gamma_L$	28	0.2222	–1.5041	7.29	–1.23
$\delta_L\,\alpha_D$	75	0.5952	–0.5188	9.72	–0.91
$\delta_L\,\beta_L$	97	0.7698	–0.2616	4.46	0.21
$\delta_L\,\delta_D$	28	0.2222	–1.5041	11.28	–1.41
$\delta_L\,\varepsilon_D$	44	0.3492	–1.0521	6.11	–1.54
$\delta_L\,\gamma_D$	15	0.1190	–2.1282	5.78	–2.25
$\delta_L\,\gamma_L$	51	0.4048	–0.9045	3.52	–0.31
$\delta_D\,\gamma_D$	12	0.0952	–2.3514	12.97	–1.70
$\delta_D\,\beta_L$	37	0.2937	–1.2254	7.47	–0.58
$\delta_D\,\delta_D$	141	1.1190	0.1125	14.33	–2.20
$\delta_D\,\varepsilon_D$	8	0.0635	–2.7568	14.09	–2.33
$\delta_D\,\varepsilon_L$	31	0.2460	–1.4023	5.88	–0.33
$\delta_D\,\gamma_D$	4	0.0317	–3.4499	9.84	–3.04
$\delta_D\,\gamma_L$	16	0.1270	–2.0637	7.00	–1.10
$\gamma_L\,\alpha_D$	42	0.3333	–1.0986	5.09	–0.60
$\gamma_L\,\beta_L$	117	0.9286	–0.0741	4.17	0.52
$\gamma_L\,\delta_D$	19	0.1508	–1.8918	4.53	–1.10
$\gamma_L\,\delta_L$	68	0.5397	–0.6168	3.20	–0.31
$\gamma_L\,\varepsilon_D$	20	0.1587	1.8405	5.40	–1.23
$\gamma_L\,\varepsilon_L$	331	2.6270	0.9658	5.61	1.07
$\gamma_L\,\gamma_D$	14	0.1111	–2.1972	2.34	–1.94
$\gamma_L\,\gamma_L$	126	1.00	0.00	0.00	0.00
$\gamma_D\,\beta_L$	9	0.0714	–2.6391	6.87	–1.42
$\gamma_D\,\delta_D$	6	0.0476	–3.0445	8.73	–3.04
$\gamma_D\,\delta_L$	15	0.1190	–2.1282	6.11	–2.25
$\gamma_D\,\varepsilon_D$	5	0.0397	–3.2268	9.15	–3.17
$\gamma_D\,\varepsilon_L$	55	0.4365	–0.8289	6.57	–0.87
$\gamma_D\,\gamma_D$	4	0.0317	–3.4499	4.69	–3.88
$\gamma_D\,\gamma_L$	9	0.0714	–2.639	2.11	–1.94

[a]Taken from [24].
[b]Numbers correspond to the numbers given in Table 7.
[c]Numbers correspond to the numbers given in Table 8.

Table 10
Success rate for the five MM methods used, in terms of number of conformations, to score within ± 3 kcal mol^{-1} limits of the ab initio established scaled relative stabilities (ΔE^{scaled})

Force field	ΔE^{scaled} by HF/3-21G		ΔE^{scaled} by MP2/6-311 + + G(d,p)	
	Within ± 3 kcal mol^{-1}	Border line	Within ± 3 kcal mol^{-1}	Border line
MM +	1	0	2	0
AMBER	5	0	6	1
BIO +	3	0	3	0
OPLS	3	1	2	1
ECEPP/2	7	0	7	0
Total	19	1	20	2
(slope)	0.81		0.60	
(intercept)	–0.21		–0.35	

± 3 kcal mol^{-1} deviation from the perfect straight line, one gets a fairly large number of points falling within these limits. The numbers, on the basis which Fig. 14 is constructed, are based on the acceptance of the ΔE^{scaled} values of a single amino acid diamide (i.e. HCONH–CHMe–CONH$_2$). In turn, the ΔE^{scaled} values are based on the $\ln(p_x/p_{\gamma_L})$ data, the population of various conformations of single amino acid

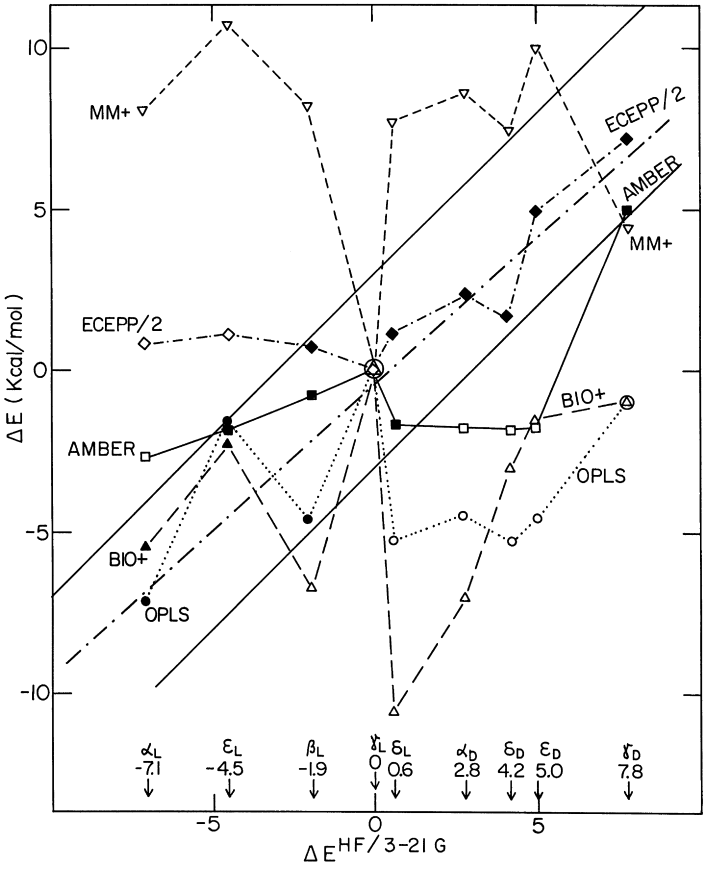


Fig. 15. Correlation of relative energies (ΔE) computed with the use of five different force fields [MM+, AMBER, BIO+(CHARMM), OPLS and ECEPP/2] for HCONH–CHMe–CONH$_2$ with ΔE^{scaled} generated at the HF/3-21G molecular level of theory.

Table 11
Thermodynamic properties of HCONH–CHMe–$CONH_2$ computed at the HF/3-21G level of theory for the seven stable conformations

	E^a	ΔE^b	S^c	G^a	ΔG^b
α_D	–412.465294	5.95	87.7	–412.357189	5.35
β_L	–412.472783	1.25	87.7	–412.364536	0.73
δ_D	–412.463135	7.31	88.6	–412.355356	6.50
δ_L	–412.468680	3.83	90.3	–412.361753	2.48
ε_D	–412.461774	8.16	88.9	–412.354334	7.14
γ_D	–412.470750	2.53	86.2	–412.361099	2.89
γ_L	–412.474780	0.00	86.7	–412.365707	0.00

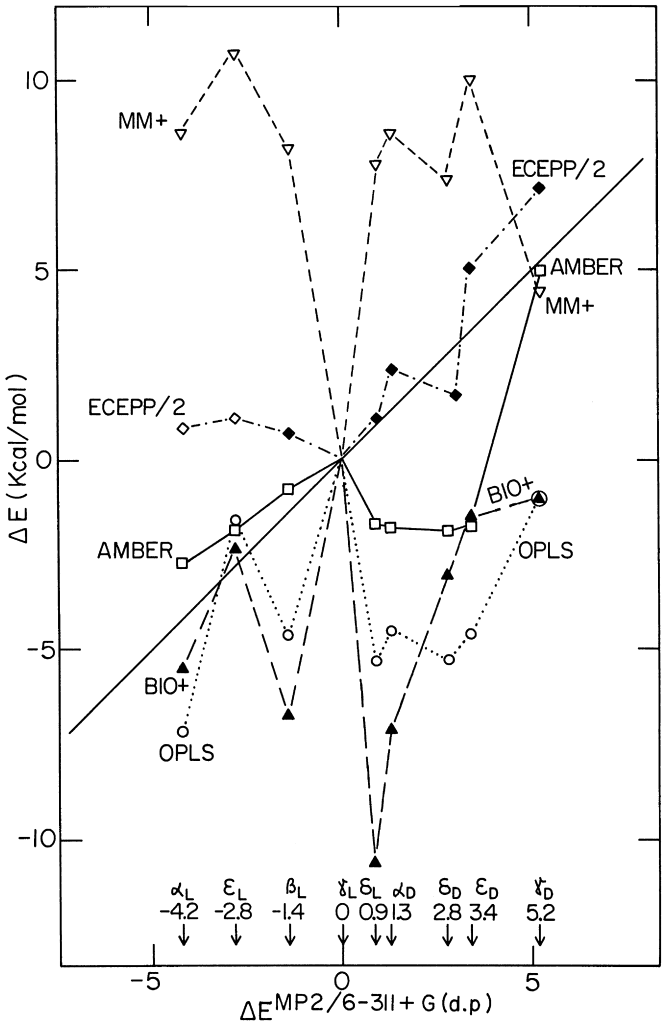


Fig. 16. Correlation of relative energies (ΔE) computed with the use of five different force fields [MM+, AMBER, BIO+(CHARMM), OPLS and ECEPP/2] for HCONH–CHMe–$CONH_2$ with ΔE^{scaled} generated at the MP2/6-311++G(dp) level of theory.

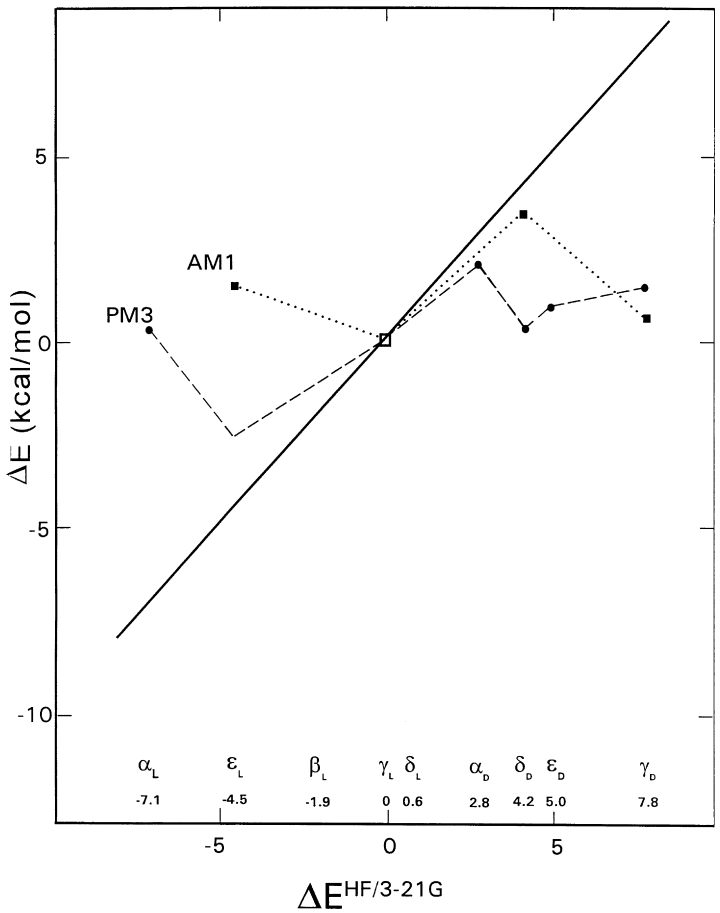


Fig. 17. Correlation of relative energies (ΔE) computed with the use of two different force fields [AM1 and PM3] for: HCONH–CHMe–$CONH_2$ with ΔE^{scaled} generated at the HF/3-21G molecular level of theory.

residues in proteins. This is how the ΔE^{scaled} was arrived at as the primary standard. In view of that we need to compare the ab initio ΔE^{scaled} values with the computed MM relative stabilities: ΔE^{MM}.

Plotting the relative stabilities computed by MM (ΔE_{MM}), using five different force fields, against the "primary standard", ΔE^{scaled}, as defined at the HF/3-21G and at the MP2/6-311++G(dp) levels of theory (Figs. 15 and 16 respectively) one gets a much stronger hint of linearity than it was apparent in Fig. 7. Taking a

Table 12
Scaled Gibbs free energy of stabilities (ΔG^{scaled}) for the conformations of HCONH–CHMe–$CONH_2$ computed at the HF/3-21G level of theory

Conformation *x*	ΔG^{HF}	$\Delta\Delta G^{shift}$	$\Delta G^{HF} + \Delta\Delta G^{shift}$	ΔG^{scaled}
β_L	0.73		0.73	–1.87
γ_L	0.00	2.60	2.60	0.00
δ_L	2.48		2.48	–0.12
α_D	5.35		5.35	2.75
δ_D	6.50		6.50	3.90
ε_D	7.14		7.14	4.54
γ_D	2.89	6.91	9.80	7.20

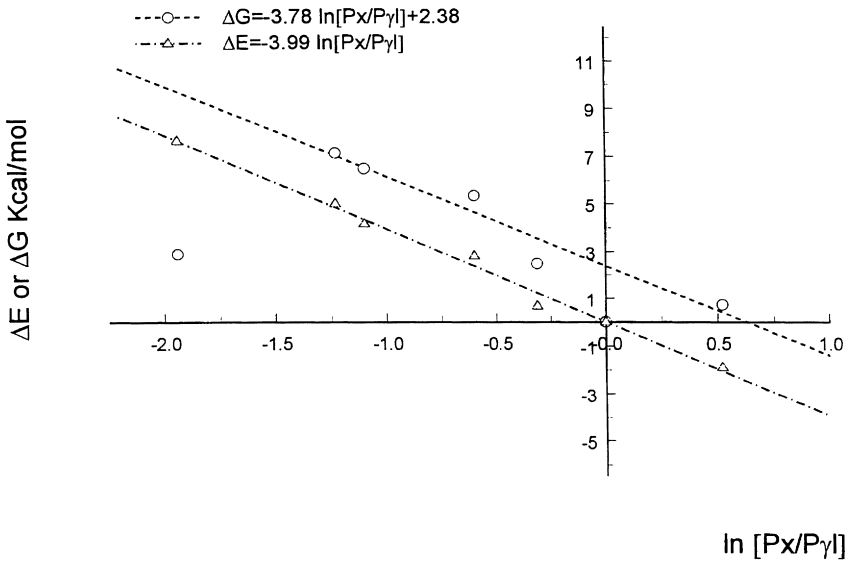


Fig. 18. Correlation of ΔG values, computed at the HF/3-21G level of theory for HCONH–CHMe–$CONH_2$, with relative populations: $\ln(p_x/p_{\gamma_L})$. The correlation for ΔE^{scaled} is shown for comparison (see Fig. 12).

$\pm$ 3 kcal mol^{-1} deviation from the perfect straight line, which corresponds to the hypothetical case

$$\Delta E_{\text{MM}} = \Delta E^{\text{scaled}} \tag{8}$$

one gets a number of points falling within these limits. Just how many points of each of the five MM fall within these limits may be regarded as a measure of success of the individual force fields.

Such an assessment is summarized in Table 10.

A linear least-square fit to the points that fall within the above limits yields a straight line, Eq. (9), which gives an indication how the MM methods, on average, correlate with the primary standard chosen above.

$$\Delta E_{\text{MM}} = (\text{slope})\Delta E^{\text{scaled}} + (\text{intercept}) \tag{9}$$

If the (slope) is unity and the (intercept) vanishes then one has a perfect fit. However, the (slope) and (intercept) are far from such ideal values, as may be seen in Table 10. The fitted straight lines are also shown in Figs. 15 and 16 for ΔE^{scaled} established for RHF/3-21G and MP2/6-31G levels of theory respectively.

Table 13
Selected optimized geometrical parameters of HCO–NH–CH–Me–$CONH_2$ as computed by various methods

Method	β_L			γ_L		
	N1C1 (Å)	C1C2 (Å)	C2C1N1 (°)	N1C1 (Å)	C1C2 (Å)	C2C1N1 (°)
Empirical						
MM +	1.4503	1.5308	109.81	1.4528	1.5292	111.00
AMBER	1.4578	1.5324	109.96	1.4573	1.5326	110.47
BIO +	1.4788	1.5506	110.55	1.4776	1.5512	110.24
OPLS	1.4560	1.5284	109.69	1.4556	1.5302	110.39
Semi-empirical						
AM1				1.4429	1.5509	111.90
PM3	1.4854	1.5331	108.19	1.4899	1.5387	111.34
Non-empirical						
HF/3-21G	1.4512	1.5222	106.39	1.4723	1.5350	109.81
HF/6-311 + + G(d,p)	1.4441	1.5256	107.75	1.4584	1.5364	109.56
MP2/6-311 + + G(d,p)	1.4443	1.5245	107.53	1.4597	1.5371	109.13

Fig. 17 shows analogous representations for the two semiempirical methods (AM1 and PM3) used.

3.5. The overstabilization of γ-turn (γ_D) and inverse γ-turn (γ_L)

We may now return to the question of overstabilization of the γ conformations as proposed by the definition of the primary standard for stability, ΔE^{scaled}. The principle of such a definition is irrespective of whether it is based on the HF/3-21G or MP2/6-311++G(d, p) computations.

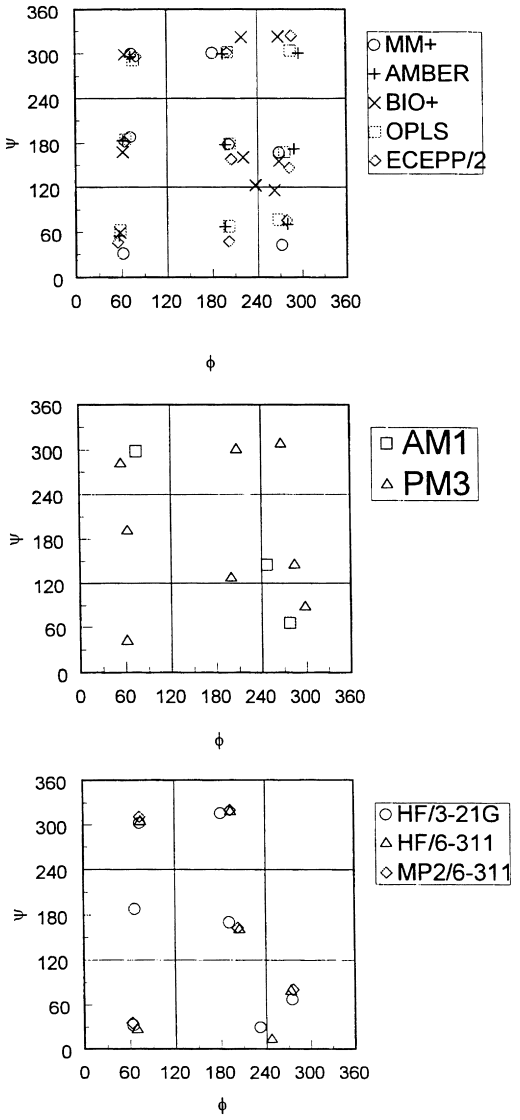


Fig. 19. Distribution of molecular conformations in the Ramachandran framework computed by (A) empirical, (B) semi-empirical and (C) non-empirical methods.

One may wonder how much of this overstabilization is due to entropy factors. When a cyclic structure is formed the system is more ordered, therefore its entropy is expected to increase the value ΔG with respect to ΔH or ΔE. For this reason, the frequencies were computed and the ΔG values were generated for HCONH–CHMe–$CONH_2$ at the HF/3-21G level of theory.

Clearly, this level of theory is not recommended for reliable frequency calculations but it is thought to be at least indicative of the shift in relative stability. The results are shown in Table 11. Plotting the ΔG values against $\ln(p_x/p_{\gamma_L})$ values, taken from Table 3, we obtain a fairly good correlation as shown in Fig. 18.

Clearly the γ_L and γ_D points are shifted towards the straight line more closely than they were on the ΔE^{scaled}. However, the change is hardly satisfactory. While MP2/6-311++G(d, p) results would be more reliable, nevertheless it is not expected that the change will be sufficiently large to override the overstabilization phenomenon. The ΔG^{scaled} values are summarized in Table 12. The $\Delta\Delta G^{shift}$ values, 2.60 (γ_L) and 6.91 (γ_D) are much smaller than the $\Delta\Delta E^{shift}$ values 3.15 (γ_L) and 8.26 (γ_D) listed in Table 5.

Another factor may be that in the protein structures the C=O....H–N hydrogen bonds, which hold the γ_L and γ_D conformations in the 7-membered ring, are not as strong as they are in their isolated form. One could only speculate what might cause the weakening of this internal hydrogen bond in the protein. One such possibility could be the involvement of structural water molecules in the protein crystal, while another possibility might be the slight distortion of the ω values from their ideal 180° (trans) configuration. These possibilities however have not been investigated. Consequently, we are not in a position, at this stage to offer, an explanation for the overstabilization of the γ_L and γ_D conformations.

3.6. Molecular conformations

The various methods yield torsional angles that are far more comparable than relative energies (Tables 1 and 2). The Ramachandran-type presentation (Fig. 19) illustrates this point.

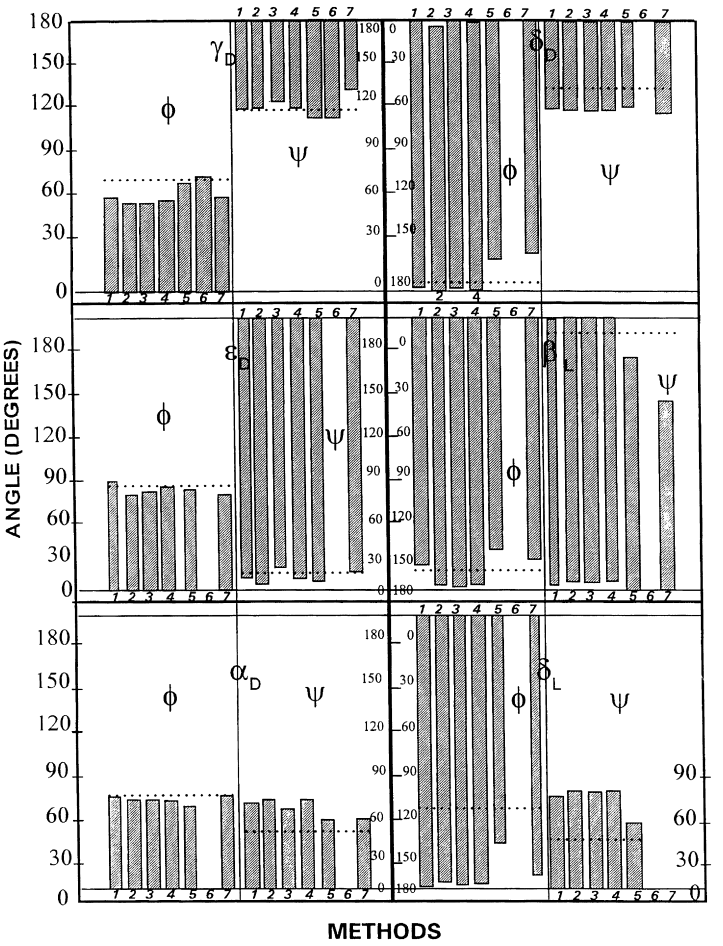


Fig. 20. Bar diagrams of ϕ and ψ variations as computed by various methods. The HF/3-21G values are shown as · · ·. The methods used are enumerated from 1 to 7. 1, MM + ; 2, AMBER; 3, BIO + ; 4, OPLS; 5, ECEPP/2; 6, AM1; 7, PM3.

Fig. 20 shows graphically the ϕ, ψ values optimized by the various methods for HCONH–CHMe–$CONH_2$ and MeCONH–CHMe–CONHMe. As far as other optimized geometrical parameters are concerned, they are also more comparable than computed relative energies. Selected bond lengths and bond angles, summarized in Table 13, show the extent of comparability.

In conclusion, it is fair to say that empirical force fields (MM) and semiempirical MO calculations are far more reliable when used for computing geometries of molecular conformers than when used for the assessment of relative stabilities of various conformers.

Acknowledgements

The continued financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada is gratefully acnowledged. The research was also supported in part by the Hungarian Scientific Research Foundation (OTKA Nº F-017188, F-013799, T-016328, and T-017129) Budapest, Hungary and by grants from Universidad Nacional de San Luis (UNSL) and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina. Leslie L. Torday would like to thank the Hungarian Eötvos Fellowship Committee for granting a research fellowship at the University of Toronto.

References

[1] I.K. Roterman, M.H. Lambert, K.D. Gibson, H.A. Scheraga, J. Biomolecular Structure and Dynamics 7 (1989) 421–453.

[2] M.A. McAllister, A. Perczel, P. Csaszar, W. Viviani, J.L. Rivail, I.G. Csizmadia, J. Mol. Struct. (THEOCHEM) 288 (1993) 161–179.

[3] R.C. Bingham, M.J.S. Dewar and D.H. Lo, J. Am. Chem. Soc., 97 (1975) 1285, 1294, 1302, 1307.

[4] M.S.J. Dewar, W. Thiel, J. Am. Chem. Soc. 99 (1977) 4899.

[5] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1986) 3902–3909.

[6] J.J.P. Stewart, J. Comp. Chem. 10 (1989) 221.

[7] A. Perczel, J.G. Angyan, M. Kajtar, W. Vivini, J.L. Rivail, J.F. Marcoccia, I.G. Csizmadia, J. Am. Chem. Soc. 113 (1991) 6256–6265.

[8] (a) J.N. Scarsdale, C. Van Alsenoy, V.J. Klimkowski, L. Schäfer and F.A. Momany, J. Am. Chem. Soc., 105 (1983) 3438-3445.(b) L. Schäfer, V.J. Klimkowski,F.A. Momany, H. Chuman and C. Van Alsenoy, Biopolymers, 23 (1984) 2335-2347. (c) S.J. Wiener, U.C. Singh, T.J. O'Donnel and P. Kollman, J. Am. Chem. Soc., 106 (1984) 6243-6245.(d) T. Haed-Gordon, M. Head-Gordon, M.J. Frisch, C.L. Brooks III and J.A. Pople, Int. J. Quantum Chem., Quantum Biol., 16 (1989) 311-322.(e) H.J. Böhn and S. Brode, J. Am. Chem. Soc., 113 (1991) 7129-7135.(f) T. Haed-Gordon, M. Head-Gordon, M.J. Frisch, C.L. Brooks III and J.A. Pople, J. Am. Chem. Soc., 113 (1991) 5989-5997.

[9] W. Viviani, J.L. Rivial, A. Perczel, I.G. Csizmadia, J. Am. Chem. Soc. 115 (1993) 8321–8329.

[10] (a) A. Perczel, R. Daudel, J.G. Angyan and I. G. Csizmadia, Can. J. Chem., 68 (1990) 1882-1888.(b) Ö. Farkas, A. Perczel, J.F. Marcoccia, M. Hollósi and I. G. Csizmadia, J. Mol. Struc. (THEOCHEM), 331 (1995) 27-36.(c) A. Perczel, Ö. Farkas and I.G. Csizmadia, J. Comp. Chem., 17 (1996) 821-834.(d) A. Perczel, Ö. Farkas and I.G. Csizmadia, J. Am. Chem. Soc., 118 (1996) 7809-7817.(e) A. Perczel, Ö. Farkas, J.F. Marcoccia and I.G. Csizmadia, International J. Quantum Chem., 61 (1997) 797-814.

[11] (a) Ö. Farkas, M.A. McAllister, J.H. Ma, A. Perczel, M. Hollósi and I.G. Csizmadia, J. Mol. Struc. (THEOCHEM), 396 (1996) 105-114.(b)A. Perczel, Ö. Farkas and I.G. Csizmadia, Can. J. Chem., 75 (1997) 1120-1130.

[12] G. Endrédi, A. Perczel, Ö. Farkas, M.A. McAllister, G.I. Csonka, J. Landik, I.G. Csizmadia, J. Mol. Struct. (THEOCHEM) 391 (1997) 15–26.

[13] R.F. Frey, J. Coffin, S.Q. Newton, M. Ramek, V.K.W. Cheng, F.A. Momany, L. Schäfer, J. Am. Chem. Soc. 114 (1992) 5369.

[14] HYPERCHEM® 4.5, (1994) Hypercube, Inc. 419 Phillip St., Waterloo, Ontario, Canada N2L 3X2.

[15] The HyperMM + force field is derived from the public domain code developed by N.L. Allinger, J. Am. Chem. Soc., 99 (1977) 8127.

[16] S.J. Wiener, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S.Profeta Jr., , P. Wiener, J. Am. Chem. Soc. 106 (1984) 765–784.

[17] The BIO + force field is an implementation of the CHARMM force field, B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, J. Comput. Chem., 4 (1983) 187-217.

[18] (a) W.L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc., 110 (1988) 1657-1666.(b) J. Pranate, S. Wierschke, W.L. Jorgensen, J. Am. Chem. Soc., 113 (1991) 2810.

[19] J.J.P. Stewart (1990) MOPAC 6.00: A Semiempirical molecular Orbital Program. 6th edn., F.J. Seiler Research Lab., US Air Force Academy, Boulder, CO, USA.

[20] A. Perczel, Ö. Farkas, I.G. Csizmadia, J. Am. Chem. Soc. 117 (1995) 1653.

[21] Gaussian 94, M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T.A. Keith, G.A. Petersson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Defrees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez and J.A. Pople, Gaussian, Inc., Pittsburgh PA, 1995.

[22] A. Perczel, I.G. Csizmadia, International Reviews in Physical Chemistry 14 (1995) 127–168.

[23] A. Perczel, M.A. McAllister, P. Csaszar, I.G. Csizmadia, J. Am. Chem. Soc. 115 (1993) 4849–4858.

[24] A. Perczel, M.A. McAllister, P. Csaszar, I.G. Csizmadia, Can. J. Chem. 72 (1994) 2050–2070.