Intrinsically Stable Secondary Structure Elements of Proteins: A Comprehensive Study of Folding Units of Proteins by Computation and by Analysis of Data Determined by X-ray Crystallography[+]

András Perczel,*[a] Imre Jákli,[a] and Imre G. Csizmadia[a, b]

Abstract: Different protein architectures show strong similarities regardless of their amino acid composition: the backbone folds of the different secondary structural elements exhibit nearly identical geometries. To investigate the principles of folding and stability properties, oligopeptide models (that is, HCO-(NH-L-CHR-CO)$_n$-NH$_2$) have been studied. Previously, ab initio structure determinations have provided a small amount of information on the conformational building units of di- and tripeptides. A maximum of nine differently folded backbone types is available for any natural α-amino acid residue, with the exception of proline. All of these conformers have different relative energies. The present study compiles an ab inito database of optimized HCO-(L-Xxx)$_n$-NH$_2$ structures, where $1 \leq n \leq 8$ and Xxx = Ala or Gly. All homoconformers (α helix, β sheet, collagen helix, etc.) of the different backbone folds were optimized, along with additional β-turn-type heteroconformers. The comprehensive analysis of more than 150 fully optimized polyalanine and polyglycine structures reveals the same energy-preference profile of major secondary structures as is found in globular proteins. The analysis of relative energies at three different levels of theory (RHF/3-21G, RHF/6-311++G(d,p)//RHF/3-21G, and RHF/6-311++G(d,p)) for the above-mentioned achiral (Xxx = Gly) and chiral (Xxx = Ala) molecular structures shows how these common secondary structure elements gradually become more and more stable folds in the oligopeptides as the length of the peptide chain increases. This indicates that stability (local energy preference) of conformational building units seems to be a major driving force in peptide and protein folding. Furthermore, the preferred conformers of the gas phase are rather similar to those observed in proteins crystallized from aqueous media. Indeed, the relative energies for the different computed conformers show remarkable agreement with the frequency of occurrence of the same structural motifs retrieved from a nonhomologous X-ray crystallography database.

Keywords: ab initio calculations · peptides · protein structures · secondary structure elements · structure – energy correlations

Introduction

Proteins exhibit certain conformational structures due to their numerous internal torsional modes. Such structural patterns may be determined experimentally by NMR spectroscopy or X-ray crystallography. One cannot be certain what fraction of the stability of proteins originates from internal (or intrinsic) effects, which are also present in the gas phase, and what fraction is due to environmental features (for example, solvation or structural water of crystallization). One efficient method to answer such questions is quantum molecular computations. When such computations are carried out on isolated molecules, thus ignoring the environmental influence, they can provide reliable stability data on selected molecular conformers. Such theoretical studies can help determine and classify the driving force, as well as its components, operative in the folding processes of proteins.

One important problem is to determine whether typical secondary structure elements of proteins (α helices, β sheets, collagen helices, different types of β turns, etc.) are intrinsically stable or whether they are more or less stable folds that are further stabilized by favorable external influences. The present study uses ab initio molecular computational results on a series of chiral and achiral peptide models of increasing chain length (see Scheme 1) to answer such questions.

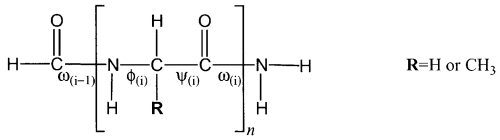
[a] Prof. A. Perczel, I. Jákli, I. G. Csizmadia
Department of Organic Chemistry
Eötvös University
P.O. Box 32, 1518 Budapest 112 (Hungary)
Fax: (+36)-1-209-0602
E-mail: perczel@para.chem.elte.hu

[b] Prof. I. G. Csizmadia
Department of Chemistry
University of Toronto
Toronto, ON M5S 3H6 (Canada)

[+] Part XXIX; for Part XXXVIII see: J. Hudáky, H. A. Baldoni, A. Perczel, *J. Mol. Struct. (Theochem)* **2002**, *582*, 233 – 249.

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

DOI: 10.1002/chem.200304843



Scheme 1. Oligopeptide models of increasing chain length ($1 \le n \le 8$) composed from chiral ($R = CH_3$) or achiral ($R = H$) natural α-amino acid residues.

Methods

Computational details: Computations were carried out by using the Gaussian 98 program[1] at the RHF level of theory and applying two different basis sets, 3-21G and 6-311++G(d,p). For oligopeptide models geometries were optimized at two levels of theory, RHF/3-21G and RHF/6-311++G(d,p). However, energies were also computed at a third level of theory, RHF/6-311++G(d,p)//RHF/3-21G, as single-point calculations to show their similarity to the geometry optimizations demanding 20–40 times more computer processing. Minima were optimized by using the recent implementation[2] of the GDIIS algorithm,[3–5] built into the Gaussian 98 program. The conformational and stability properties of nine basic oligopeptide conformers (see Figure 1 for Ala_6 models) were computed, as a function of their chain length, both for HCO-(L-Ala)$_n$-NH$_2$ and HCO-(Gly)$_n$-NH$_2$ model peptides.

Computational accuracy and reliability: Conventional wisdom, acquired in molecular computations, dictates that the larger the basis set and the higher the level of theory used the more reliable are the results obtained. This implies that restricted Hartree–Fock (RHF) results are not as "valuable" as those obtained by post-Hartree–Fock methods, such as MP2 or DFT, which include electron correlation. The validation of such an overwhelming generalization was suspect as early as the mid-1990s,[6] when MP2 relative energies obtained with large basis sets were compared to those obtained by RHF calculations with smaller basis sets. In 1998 Mohle and Hoffman[7] pointed out in their study of Ac-L-Ala-NHCH$_3$ that "…at the Gibbs free energy level the conformers are again of comparable stability as already predicted at the Hartree–Fock level…" due to "…considerable compensation of correlation and entropy effects…and a small destabilizing effect of the thermal energies." Subsequently, this conclusion was confirmed in the case of diglycine triamide (Ac-Gly-Gly-NHCH$_3$).[7]. Most recently, it has been

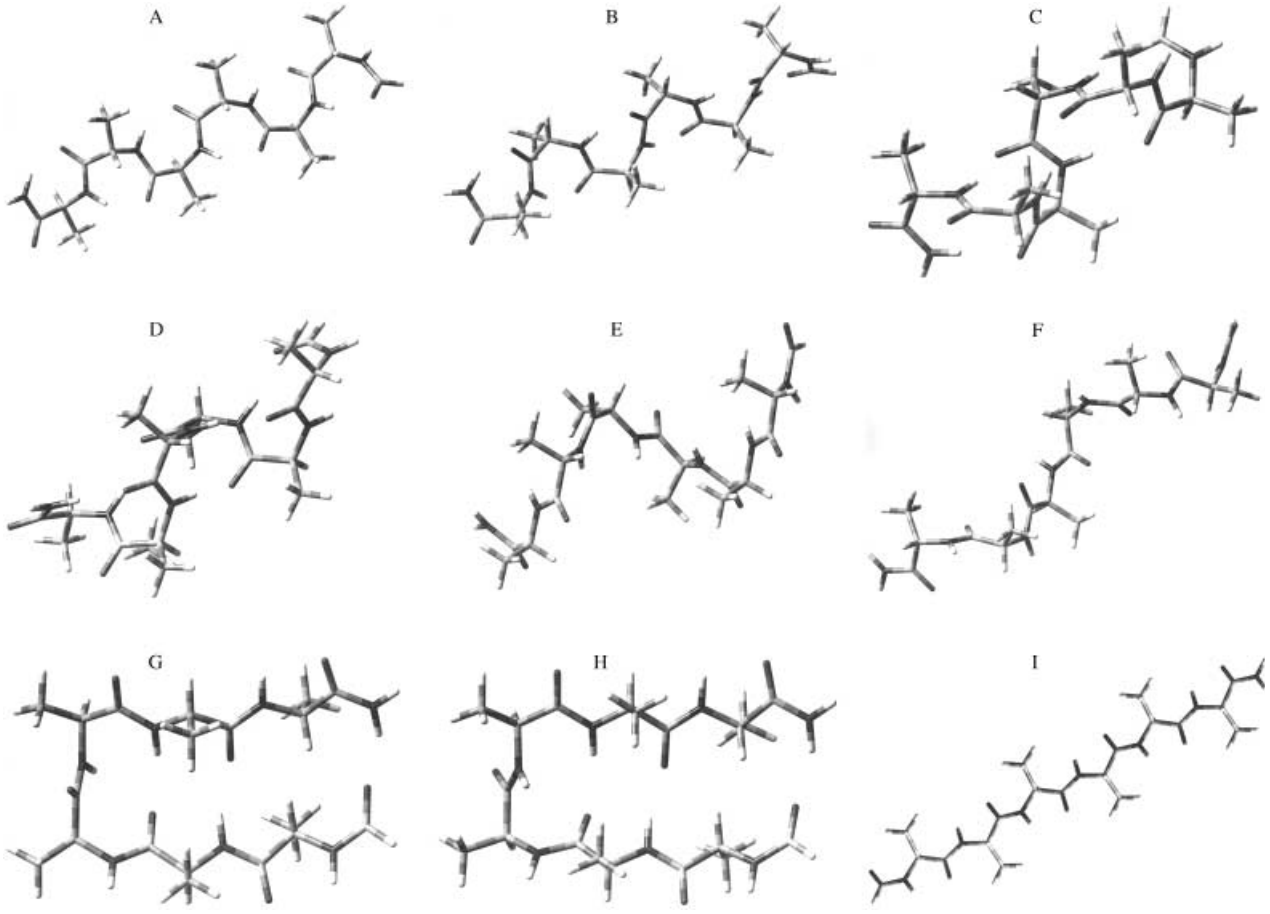


Figure 1. The different homo- and heteroconformers of a hexaalanine oligopeptide (HCO-(L-Ala)$_6$-NH$_2$): A) Repeated inverse γ turn ($[\gamma_L]_6$); B) repeated "normal" γ turn ($[\gamma_D]_6$); C) right-handed α helix ($[\alpha_L]_6$); D) left-handed α helix ($[\alpha_D]_6$); E) $[\delta_D]_6$; F) left-handed collagen helix ($[\varepsilon_D]_6$); G) type II β turn preceded and followed by antiparallel β sheets ($\beta_L\beta_L\varepsilon_L\alpha_D\beta_L\beta_L$); H) type I β turn preceded and followed by antiparallel β sheets ($\beta_L\beta_L\alpha_L\delta_L\beta_L\beta_L$); I) single-stranded β sheet ($[\beta_L]_6$).

pointed out[8] that higher-level CCSD(T) energies of different conformers of small peptide models show surprisingly good correlation with RHF/3-21G data. In view of these results it is safe to say that RHF relative energies (ΔE) obtained by using small basis sets are not fundamentally different from the relative Gibbs free energies (ΔG) computed at higher levels of theory with the inclusion of electron correlation.

Currently, geometric and energy properties of only small molecules can be computed at state-of-the-art accuracy. Since accuracy and economy form opposing requirements, there arises a competition between the level of theory applied in the calculation and the size of the studied molecule. Thus, the possibilities for a calculation are clearly limited by the available computer processing time. However, due to a fortuitous cancellation of errors, it becomes possible, at least in the case of oligopeptides, to carry out quantum chemical computations at a modest level of theory (for example, RHF/3-21G) and obtain acceptable structures with relative energies (ΔE) almost as accurate as thermodynamic stabilities (ΔG) determined at higher levels of theory (such as DFT, MP2, CCSD(T)).

Databases: A total of 1211 proteins, with a homology level equal to or lower than 25 %[9, 10] and containing 206 889 amino acid residues, were grouped in this database. These proteins were analyzed for all sequence units $(Xxx)_n$, where $1 \leq n \leq 8$, without any preference for particular amino acid sequences. All entries correspond to high-resolution X-ray crystal structures (no NMR spectroscopy structures) taken from the 1998 issue of the Protein Databank. Each typical conformation was defined by its backbone values, a sequence of consecutive ϕ and ψ units with an a priori determined tolerance value denoted by k.

Nomenclature for backbone and side-chain conformers: As with any other surface with periodic nature, the Ramachandran potential energy surface (PES),[11] $E = E(\phi,\psi)$, can be divided into conformational subsets according to numerous concepts. For the torsional angle pair ϕ and ψ, multidimensional conformation analysis[12] predicts nine catchment regions. According to IUPAC/IUB recommendations[13] the *gauche*+ $(g+)$, *anti* (a), and *gauche*− $(g-)$ descriptors can be used for notation of the basic backbone conformers of peptides. Incorporating traditional elements and the aforementioned recommendation, the following shorthand notation of typical backbone folds of peptides and proteins was introduced:[12] $\alpha_L = (g-,g-)$, $\alpha_D = (g+,g+)$, $\beta_L = (a,a)$, $\gamma_L = (g-,g+)$, $\gamma_D = (g+,g-)$, $\delta_L = (a,g+)$, $\delta_D = (a,g-)$, $\varepsilon_L = (g-,a)$, and $\varepsilon_D = (g+,a)$. Peptides composed of α-amino acid residue(s) with S absolute configuration (typically L relative configuration), prefer backbone conformers from the L regions of the PES. Their mirror-image configurational counterparts (R absolute and typically D relative configuration) favor the mirror-image backbone folds (conformers from the D region). For glycine no relative energy preference is operative. Thus, for labeling typical backbone conformers of glycine either both indices are used or the Greek symbol is given without any specification (for example, γ_{LD} or γ).

Peptide folding as a model for protein folding: In general, it is true for complex systems that the whole is more than the sum of its components. Nevertheless, we cannot formulate a comprehensive description of a complex system, for example, a full-length linear polymer such as a protein, without understanding its components. This implies that understanding details of peptide folding (such as intrinsic stability of the different secondary structure elements) could have a significant impact on our views of how proteins are folded and stabilized.[14–19].

The assumption that quantum-mechanically computed structures of peptides will reveal some basic principles important in the understanding of peptide and protein folding has a considerably long history. In the early stages (1965–1980) semiempirical methods were introduced.[20–24] Subsequently, in the early 1980s ab initio HF computations were carried out without geometry optimizations. The next phase was a few years thereafter, when selected fully relaxed conformers of a few amino acid derivatives were obtained first by Schäfer and co-workers[25–28] and subsequently by others.[29] Systematic studies first of dipeptides and later on tripeptides became feasible in the early 1990s (for reviews, see refs. [30–32]).

Results and Discussion

Out of the nine different, topologically meaningful backbone structures, only seven conformers are typically distinguishable on the ab initio Ramachandran surfaces of most amino acid diamides. In the case of HCO-L-Ala-NH_2 the α_D, β_L, γ_L, γ_D, δ_L, δ_D, and ε_D backbone orientations turn out to be minima at the RHF/3-21G level of theory, while neither the α_L nor the ε_L structures are found to be minima. Thus, it is reasonable to consider first the homoconformers of the former structural building units (for example, $[\beta_L]_n$) as conformational motifs of oligopeptides (for example, $(Ala)_n$). Some of these homoconformers are well-known secondary structure elements in proteins, while others rarely occur or are not observed at all. For example $[\beta_L]_n$ stands for the extended conformation, also called a β sheet, and occurs very frequently. In contrast to this, left-handed helices, $[\alpha_D]_n$, are occasionally observed in proteins as short segments ($n \leq 3$) but never as fully evolved left-handed helical structures. Furthermore, the conformational building unit of inverse γ turns, γ_L, is an intrinsically stable and low-energy structure, however, it forms homoconformers of only short length ($[\gamma_L]_n$ where $n \leq 4$). On the other hand, according to ab initio calculations, the structural building unit of right-handed helical conformations, α_L, is not a stable structure for most amino acid dipeptides. However, $[\alpha_L]_n$ gains stability as the chain length increases. Interestingly, δ_L is a stable conformer of single amino acid diamides but with increasing chain length the $[\delta_L]_n$ conformation constantly shifts towards $[\alpha_L]_n$, a fact suggesting that the right-handed helical structure becomes an important backbone structure of peptides.

What is the driving force of such a selection? In order to answer this question, seven homoconformers (Figure 2), both for chiral alanine and for achiral glycine, were optimized and

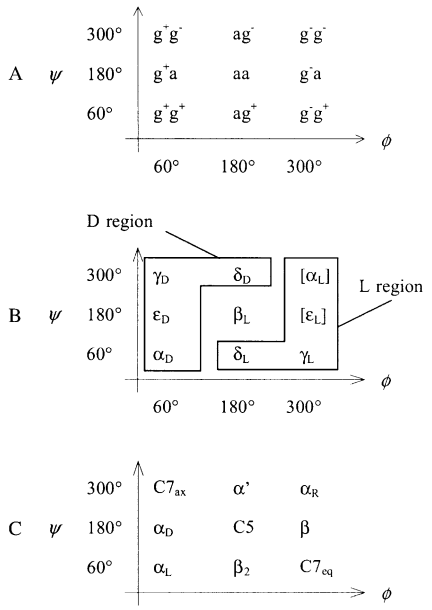


Figure 2. Ideal location of selected conformational building units of polypeptide homoconformers on a Ramachandran surface.

analyzed in terms of structure and energy. For this purpose, all seven homoconformations of both HCO-(L-Ala)$_n$-NH$_2$ and HCO-(Gly)$_n$-NH$_2$ were fully optimized with two different basis sets.

The shorthand notations stand for the following backbone structures: $[\beta_L]_n$ = single-stranded β-pleated sheet, $[\gamma_L]_n$ = repeated inverse γ turns, $[\gamma_D]_n$ = repeated γturns, $[\varepsilon_L]_n$ = polyproline II or collagen helix, $[\varepsilon_D]_n$ = mirror-image structure of polyproline II or collagen helix, $[\alpha_L]_n$ = right-handed α helix, and $[\alpha_D]_n$ = left-handed α helix. No obvious name is known for the backbone structure of $[\delta_D]_n$. In addition to homoconformers, two important secondary structure elements were also considered, namely the two major classes of β turns (types I and II). Both structures were optimized and incorporated into the present comprehensive analysis. If the polypeptide chain was long enough, β turns were embedded in the central part of two antiparallel β-pleated sheets. Thus, if $(n = k+l+2)$ amino acid residues form two extended regions of polypeptide separated by a type I β turn in the central part, the name of the polypeptide is abbreviated as $[\beta_L]_k$-$\alpha_L\delta_L$-$[\beta_L]_l$. The same structure with a type II β turn is denoted as $[\beta_L]_k$-$\varepsilon_L\alpha_D$-$[\beta_L]_l$. For labeling schemes and structural data, see refs. [33–36], in which selected turn conformers of diamino acid dipeptides were previously computed and analyzed.

Molecular geometry: The backbone conformational parameters (ϕ and ψ) in the fully relaxed peptide models show some variation as the length of the main chain increases. In our analysis, our first goal was to distinguish local and insignificant perturbation from perhaps small but significant structural shifts. As an example, the backbone conformational parameters, optimized with small (3-21G) and larger (6-311++G(d,p)) basis sets, of alanine and glycine octapeptides with β-pleated-sheet and α-helix structures are presented in the Supporting Information (Table S1).

For β-pleated-sheet conformers, standard deviation of the average dihedral angles is small (see the last line of Table S1 in the Supporting Information) at both levels of theory for both models (Gly_8 and Ala_8). Although the average values for the achiral (Gly_8) and for the chiral (Ala_8) must be different, the repetitive backbone values show insignificant fluctuation (values always smaller than 0.7°). The larger basis set (6-311++G(d,p)) results in average ϕ and ψ data ($\phi = -154.7°$, $\psi = +160.4°$) close to the typical text book values of $-150°$ and $+150°$, however, the values obtained with the smaller basis set are different from the frequently observed values in proteins.

In contrast to the homoconformer β-pleated sheet ($[\beta_L]_n$), most theoretical studies of the right-handed helical conformation have resulted in a type I β-turn subunit ($\alpha_L\delta_L$) at the C terminus of the peptide chain.[37, 38]. Rather than using the "obvious" notation of this homoconformer ($[\alpha_L]_n$), the backbone conformation should be denoted as $[\alpha_L]_{(n-1)}\delta_L$ to signal the significant structural shift at the C terminus of this fold. (Instead of the α_L catchment region, the last ϕ and ψ values belong to δ_L catchment region.) This structural shift also results in a change of the ψ value of residue $(n-1)$ from that of a typical right-handed helical value. Nevertheless, residues 1 to $(n-1)$ show insignificant fluctuation, thereby providing ϕ and ψ values rather typical of a 3_{10} helix at both levels of theory.

All homo- and heteroconformers incorporating alanine(s) and glycine(s) were analyzed in a similar manner at both levels of theory. As for the two major secondary structural elements above, the average ϕ and ψ values for all seven homoconformers, as well as for both types of β turns (types I and II), were computed and are summarized in Table 1. As the length (n) of the peptide chain increases from 1 to 8 these average ϕ and ψ values, typical for any type of backbone conformation, change on a very small scale, and therefore, such variation is insignificant. For example, when comparing the backbone torsional values of repeated inverse γ turns with $n = 2$ or 8 the shift of average backbone values (ϕ and ψ) is clearly marginal (Ala_2: $\phi = -84.6°$, $\psi = +66.7°$; Ala_8: $\phi = -84.3°$, $\psi = +65.3°$). Although in the case of the left-handed α helix, $[\alpha_D]_n$, backbone torsional values undergo a slightly more systematic shift (Ala_1: $\bar{\phi} = +63.8°$, $\bar{\psi} = +32.7°$; Ala_8: $\bar{\phi} = +56.7°$, $\bar{\psi} = +27.1°$), the magnitude of this change varies to a small degree; it falls in the range of $\pm 1°$, on average, per residue length increase.

A number of interesting observations can be made for β turns, the central parts of the antiparallel β sheets in our oligopeptide model (Figure 1g and h). Table 2 summarizes the optimized dihedral angles both for positions $(i+1)$ and $(i+2)$ with the numbering convention typically applied for β turns. Both shorter and longer peptide models, all having either type I or II β turns at their center, incorporate an almost ideal β turn. Furthermore, these rather typical backbone values show almost no conformational change as a function of the length of the oligopeptide (see Figure S1 in the Supporting Information). Thus, these "longer" peptides with β turns (either type I or type II) can be regarded as realistic models of this type of secondary structure element. Systems of β turns embedded in shorter or longer β-pleated sheet

Table 1. Average backbone torsional values [°] of ab initio optimized homoconformers of HCO-(L-Ala)$_n$-NH$_2$ of increasing main-chain length ($1 \le n \le 8$) optimized at two levels of theory.[a]

Level of theory	n	$[\varepsilon_D]_n$ right-handed collagen helix		$[\delta_D]_n$		$[\gamma_D]_n$ γ turn		$[\alpha_D]_n$ left-handed α helix		$[\gamma_L]_n$ inverse γ turn		$[\alpha_L]_n$ right-handed α helix		$[\beta_L]_n$ extended	
		ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ
RHF/3-21G	1	67.2	−177.3	−178.4	−44.2	74.1	−57.4	63.8	32.7	−84.5	67.3	–[b]	–[b]	−168.3	170.6
	2	65.7	−174.3	−177.5	−47.6	74.1	−56.2	61.3	26.6	−84.6	66.7	−68.5	−17.5	−168.1	170.7
	3	65.7	−175.7	−175.7	−49.4	74.0	−55.4	59.5	27.0	−84.5	66.2	−67.4	−15.7	−168.1	170.8
	4	65.7	−176.1	−174.8	−50.9	74.0	−54.9	58.6	26.7	−84.5	65.9	−65.5	−17.5	−168.1	170.8
	5	65.7	−176.4	−174.1	−51.9	73.9	−54.6	57.9	26.8	−84.4	65.7	−64.3	−18.4	−168.1	170.9
	6	65.7	−176.7	−173.4	−52.6	73.9	−54.3	57.4	26.9	−84.4	65.6	−63.5	−19.1	−168.1	170.9
	7	65.7	−177.0	−172.8	−53.2	73.8	−54.1	57.0	27.0	−84.3	65.4	−62.8	−19.6	−168.1	171.0
	8	65.6	−177.1	−172.3	−53.6	73.8	−54.0	56.7	27.1	−84.3	65.3	−62.3	−20.1	−168.2	171.0
	average	65.9	−176.3	−174.9	−50.4	73.9	−55.1	59.0	27.6	−84.4	66.0	−64.9	−18.3	−168.1	170.8
	deviation	0.5	1.0	2.2	3.2	0.1	1.2	2.5	2.1	0.1	0.7	2.3	1.5	0.1	0.1
RHF/6-311++G(d,p)	1	–[b]	–[b]	165.2	42.1	75.3	−55.4	69.0	26.9	−86.2	78.8	–[b]	–[b]	−151.1	161.0
	2	60.3	−143.9	–[b]	–[b]	76.0	−54.0	63.3	28.7	−87.9	76.9	−72.2	−17.3	−154.8	160.4
	3	60.7	−146.3	–[b]	–[b]	76.1	−54.3	60.9	28.6	−88.3	75.4	−68.9	−17.9	−154.8	160.3
	4	61.0	−148.1	–[b]	–[b]	76.2	−54.1	59.7	28.7	−88.5	74.6	−67.4	−18.9	−154.7	160.3
	8	61.7	−152.8	–[b]	–[b]	76.2	−53.9	57.2	29.4	−88.7	73.0	−65.0	−20.5	−154.7	160.4
	average	60.9	−147.8	165.2	42.1	76.0	−54.3	62.0	28.5	−87.9	75.7	−68.4	−18.7	−154.0	160.5
	deviation	0.5	3.3	–[b]	–[b]	0.1	0.2	2.3	0.4	0.4	1.4	2.6	1.2	0.3	0.1

[a] Polyalanine adopts neither a left-handed collagen helix $[\varepsilon_L]_n$ nor a $[\delta_L]_n$ type set of homoconformers in vacuum at these levels of theory. [b] No minima found at this level of theory.

segments are also useful to investigate the structural properties of antiparallel β-pleated sheets. A peptide model containing (k+l+2) amino acid residues can adopt a hairpin conformation where a β turn is embedded in a β-pleated sheet. If $k = l$ the hairpin part can be positioned in the central region in a symmetric manner, but if $k \neq l$ the model becomes asymmetric. With a type I β turn ($\alpha_L\delta_L$) in its hairpin part the symmetric or asymmetric model will have the conformational code $[\beta_L]_k$-$\alpha_L\delta_L$-$[\beta_L]_l$. Similarly, a type II β turn ($\varepsilon_L\alpha_D$) located at the central part of the molecule can be denoted as $[\beta_L]_k$-$\varepsilon_L\alpha_D$-$[\beta_L]_l$. (The average ϕ and ψ values of the extended parts of the antiparallel β sheets containing type I and type II β turns are summarized in the Supporting Information (Table S3).) The average value of ϕ ($\bar{\phi}$) is approximately $-140°$ both for Ala_4 and Gly_4; this is different from that measured in longer peptide models (approximately $-170°$; see Figure S2A in the Supporting Information). The latter value stabilizes in model systems composed of more than six amino acid residues

Table 2. Backbone torsional values [°] of ab initio optimized type I and type II β turns of HCO-(L-Ala)$_n$-NH$_2$ and HCO-(Gly)$_n$-NH$_2$ embedded in the central part of the polypeptide.

Level of theory	Polypeptide	n	Type I β turn			Type II β turn			
			$\phi(i+1)$	$\psi(i+1)$	$\phi(i+2)$	$\phi(i+2)$	$\phi(i+1)$	$\psi(i+1)$	$\phi(i+2)$	$\phi(i+2)$
RHF/3-21G	$(Ala)_n$	2	−68.5	−17.5	−113.1	21.3	−70.0	101.9	66.2	32.3
		4	−60.7	−29.8	−107.5	25.0	−70.0	95.3	67.0	32.5
		6	−56.6	−32.3	−129.4	30.5	−63.6	108.0	67.1	33.1
		8	−56.5	−32.4	−127.6	28.69	−62.89	108.9	66.8	32.3
		average	−60.6	−28.0	−119.4	26.4	−66.6	103.5	66.8	32.6
		deviation	5.6	7.1	10.8	4.1	3.9	6.3	0.4	0.4
	$(Gly)_n$	2	−66.5	−20.7	−109.1	19.8	−61.1	134.0	106.3	−22.1
		4	−64.9	−21.7	−105.5	19.8	−78.4	77.7	123.4	−0.7
		6	−64.3	−20.2	−103.9	17.2	−76.0	83.8	145.7	−16.8
		8	−64.21	−20.3	−103.5	16.85	−75.87	83.79	145.68	−16.7
		average	−65.0	−20.7	−105.5	18.4	−72.9	94.8	130.3	−14.1
		deviation	1.1	0.7	2.6	1.6	7.9	26.3	19.1	9.3
RHF/6-311++G(d,p)	$(Ala)_n$	2	−72.2	−17.3	−92.3	−2.1	−60.9	132.7	67.4	22.8
		4	−65.2	−28.2	−98.9	2.8	−76.8	98.7	64.6	29.8
		8	−70.8	−22.5	−137.7	16.2	−72.3	112.2	64.3	30.8
		average	−69.4	−22.7	−109.6	5.6	−70.0	114.5	65.4	27.8
		deviation	3.7	5.4	24.5	9.5	8.2	17.1	1.7	4.4
	$(Gly)_n$	2	−73.1	−15.4	−99.6	7.1	−61.3	134.2	95.3	−8.4
		4	−68.3	−24.7	−109.9	28.6	−83.0	81.4	99.5	5.4
		8	−71.1	−16.8	−108.1	20.9	−82.9	84.2	117.8	−0.9
		average	−70.8	−19.0	−105.8	18.9	−75.7	99.9	104.2	−1.3
		deviation	2.4	5.0	5.5	10.9	12.5	29.7	12.0	6.9

($k+l+2 \geq 6$). On the other hand, the average value of the other backbone conformational value ψ (ψ) shows practically no such change (see Figure S2B in the Supporting Information).

Molecular stability: With the increase of the length of the peptide chain of selected conformers of alanine and glycine oligopeptides, relative energies were obtained at three different levels of theory (Tables 3, 4, Figure 3). The conformation of a fully extended single strand of β-pleated sheet approximates the unfolded structure. Thus, the corresponding energy of this conformer, $E([\beta_L]_n)$, is the reference energy which may also be regarded as the energy of the unfolded structure. The relative energies of all other optimized structures mentioned above (folded conformers) can be calculated against $E([\beta_L]_n)$.

Table 3. Relative energies [kcal mol^{-1}] of selected conformers of HCO-(L-Ala)$_n$-NH$_2$ at three levels of theory relative to the unfolded or fully extended β-pleated sheet conformer $[\beta_L]_n$.

Level of theory	Conformer	$(Ala)_1$	$(Ala)_2$	$(Ala)_3$	$(Ala)_4$	$(Ala)_5$	$(Ala)_6$	$(Ala)_7$	$(Ala)_8$
RHF/3-21G	γ_L	−1.25	−2.05	−3.31	−4.66	−6.14	−7.67	−9.23	−10.83
	β_L	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	α_L	2.57	1.08	−1.18	−4.66	−8.72	−13.07	−17.81	−22.83
	α_D	4.70	2.69	0.22	−3.36	−7.58	−12.18	−17.13	−22.32
	δ_D	6.05	12.28	18.93	25.79	32.63	39.45	46.30	53.16
	ε_D	6.91	12.52	18.50	24.51	30.51	36.49	42.50	48.50
	γ_D	1.28	2.64	3.40	3.97	4.37	4.69	4.95	5.17
	β turn type II[a]	–[b]	2.39	3.66, 6.25	−0.49	−8.74, 1.08	−13.33	−10.7, −10.59	−16.31
	β turn type I[a]	–[b]	1.08	1.84, 3.81	−2.02	−7.04, −1.82	−12.89	−9.34, −10.28	−15.84
RHF/6-311++G(d,p)//RHF/3-21G	γ_L	−0.01	0.49	0.76	0.97	1.15	1.31	1.45	1.57
	β_L	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	α_L	1.94	1.74	1.72	0.86	−0.33	−1.80	−3.49	−5.28
	α_D	4.84	6.17	7.61	8.31	8.70	8.83	8.74	8.48
	δ_D	5.65	11.71	17.88	24.24	30.50	36.66	42.78	48.91
	ε_D	6.20	11.08	16.74	22.27	27.79	33.32	38.86	44.38
	γ_D	2.54	5.65	8.42	11.11	13.74	16.33	18.89	21.43
	β turn type II[a]	–[b]	4.51	5.64, 6.33	5.56	2.14, 6.26	−1.05	0.88, 3.09	−0.74
	β turn type I[a]	–[b]	1.74	3.67, 6.64	4.83	3.22, 5.16	−0.54	2.9, 1.39	−0.11
RHF/6-311++G(d,p)	γ_L	−0.11	0.37	0.61	0.83	n.d.	n.d.	n.d.	1.32
	β_L	0.00	0.00	0.00	0.00	n.d.	n.d.	n.d.	0.00
	α_L	2.12	1.81	1.47	0.40	n.d.	n.d.	n.d.	−6.82
	α_D	4.45	5.89	7.08	7.53	n.d.	n.d.	n.d.	6.39
	δ_D	5.29	–	–	–	n.d.	n.d.	n.d.	–
	ε_D	–	9.23	14.25	19.21	n.d.	n.d.	n.d.	39.21
	γ_D	2.43	5.31	7.90	10.40	n.d.	n.d.	n.d.	19.81
	β turn type II/A[a]	–[b]	3.10	3.59, 3.66	4.54	n.d.	n.d.	n.d.	−3.22
	β turn type I/A[a]	–[b]	1.81	2.64, 2.23	3.76	n.d.	n.d.	n.d.	−2.76

[a] Model peptides formed by an odd number of alanine residues ($(Ala)_3$, $(Ala)_5$, and $(Ala)_7$) have at least two different type I and two different type II β turn conformers. (Figure 5 reports their arithmetical averages.) [b] This type of conformer requires at least a dipeptide (triamide model system).

Table 4. Relative energies [kcal mol^{-1}] of selected conformers of HCO-(Gly)$_n$-NH$_2$ at three levels of theory relative to the unfolded or fully extended β-pleated sheet conformer $[\beta_L]_n$.

Level of theory	Conformer	$(Gly)_1$	$(Gly)_2$	$(Gly)_3$	$(Gly)_4$	$(Gly)_5$	$(Gly)_6$	$(Gly)_7$	$(Gly)_8$
RHF/3-21G	γ_L	−0.65	−1.28	−2.40	−3.65	−5.03	−6.47	−7.96	−9.48
	β_L	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	α_L	2.62	0.56	−1.68	−5.65	−10.09	−14.94	−20.19	−25.67
	β turn type II[a]	–[b]	0.37	1.59, −3.53	−3.53	−12.13, −3.02	−17.25	−14.43, −14.03	−20.30
	β turn type I[a]	–[b]	0.56	2.08, −2.66	−4.71	−10.09, −4.2	−17.71	−15.14, −14.37	−20.85
RHF/6-311++G(d,p)//RHF/3-21G	γ_L	1.23	2.62	3.69	4.72	5.69	6.63	7.56	8.45
	β_L	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	α_L	2.66	3.44	3.95	3.84	3.09	2.25	1.20	−0.01
	β turn type II[a]	–[b]	2.38	3.29, 2.99	3.31	0.31, 3.87	−1.63	0.38, 1.8	−1.18
	β turn type I[a]	–[b]	3.44	4.47, 4.19	2.83	2.72, 3.41	−1.94	−0.13, 1.54	−1.48
RHF/6-311++G(d,p)	γ_L	0.75	1.65	2.28	2.85	n.d.	n.d.	n.d.	4.69
	β_L	0.00	0.00	0.00	0.00	n.d.	n.d.	n.d.	0.00
	α_L	–	2.50	2.68	1.96	n.d.	n.d.	n.d.	−3.79
	β turn type II/A[a]	–[b]	1.38	2.23, 0.91	1.19	n.d.	n.d.	n.d.	−4.77
	β turn type I/A[a]	–[b]	2.50	3.50, 2.21	0.26	n.d.	n.d.	n.d.	−4.68

[a] Model peptides formed by an odd number of glycine residues ($(Gly)_3$, $(Gly)_5$, and $(Gly)_7$) have at least two different type I and two different type II β turn conformers. (Figure 5 reports their arithmetical averages.) [b] This type of conformer requires at least a dipeptide (triamide model system).

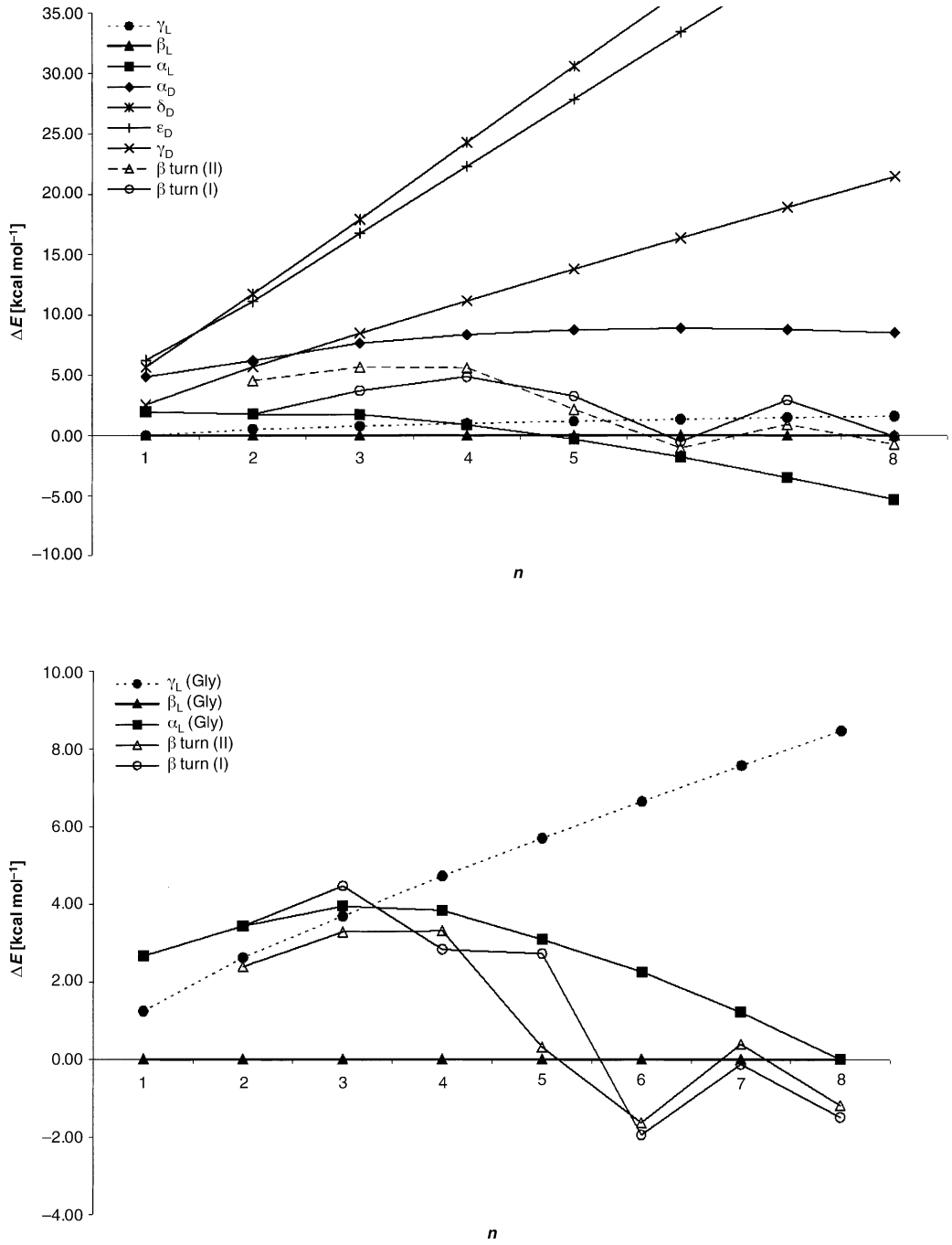


Figure 3. Relative energies of selected conformers computed at the RHF/6-311++G(d,p)//RHF/3-21G level of theory for: Top: HCO-(L-Ala)$_n$-NH$_2$ and Bottom: HCO-(Gly)$_n$-NH$_2$ as a function of the length of the oligopeptide chain. All energies are relative to the single-stranded extended conformer $[\beta_L]_n$. Glycine doesn't contain a stereocenter, thus its oligopeptides form enantiomeric-type homoconformers; for example, the left-handed and right-handed helices of HCO-(Gly)$_n$-NH$_2$ have identical geometric end energetic properties.

Variation of these relative energies as function of the number of amino acid residues (n) in the peptide chain is reported in Figure 3.

Before discussing the observed stability trends, a few general cases are to be established. The three simplest cases are as follows for all values of n investigated:

A) $E([x]_n) > E([\beta_L]_n)$
B) $E([x]_n) \approx E([\beta_L]_n)$
C) $E([x]_n) < E([\beta_L]_n)$

Case A represents a conformer which turns out to be less stable, at any length of the peptide chain, than the appropriate extended structure. Case B is where the energy difference between conformer x and that of β_L is nearly zero, a situation indicating a more or less equal stability. Finally, if case C holds, the relative energy of the given conformer is consistently lower than that of the extended structure; this suggests greater stability of such conformers than of a single strand of β-pleated sheet. Besides these three, there are additional cases to be expected. If there is a crossing point between the

functions $E = E([x]_n)$ and $E([\beta_L]_n)$, it is denoted as case D; this describes the situation where, compared to β_L, the relative stability of conformer x changes as function of n. In other words, a conformer can be more stable (or less stable) for a given length of peptide chain than its fully extended counterpart. However, at a critical value of n ($n = m$), for some reason, the relative energy between these two conformers switches over and the less-stable conformer becomes more (or less) stable than $(\beta_L)_n$. Case D is considered if only a single crossing point is observed. If, with the further increase of the length of the peptide chain, additional crossing points are detected then energy oscillation is observed. Such a situation is labeled as case E. All these cases can be examined in Tables 3 and 4, where results are presented at three levels of theory. However, for practical reasons, only one level of theory is presented in Figure 3.

In the case of the alanine series, $(Ala)_n$, the following points should be emphasized on the basis of Figure 3 and Table 3.

All structures built up and optimized from D-type homoconformers, $[x_D]_n$, such as the mirror image of the collagen helix ($[\varepsilon_D]_n$), repeated γ turns ($[\gamma_D]_n$), or the left-handed α helix ($[\alpha_D]_n$), have higher relative energy for all values of n investigated ($1 \le n \le 8$). In fact, the relative energy of these structures constantly increases as the length of the peptide chain increases. Therefore, at any level of theory, the situation of $E = E([\varepsilon_D]_n)$, $E = E([\gamma_D]_n)$, and $E = E([\delta_D]_n)$ relative to $E = E([\beta_L]_n)$ is that of case A.

No case B was observed for any of the tabulated results.

When the repeated forms of inverse γ turns ($[\gamma_L]_n$) are investigated at the RHF/3-21G level of theory their stability, $E = E([\gamma_L]_n)$, can be classified as case A, but at higher levels of theory (RHF/6-311++G(d,p)//RHF/3-21G and RHF/6-311++G(d,p)) the tendency is that of case C (Table 3). In other words, at a higher level of theory inverse γ-turn structures are expected to be less stable than β-pleated sheet structures.

The relative stability of the right-handed α helix, $[\alpha_L]_n$, falls into the category of case D. Compared to the fully extended form of the β-pleated sheet, $[\beta_L]_n$, there is a crossing point at $m = 3$ at the lower level and $m = 5$ at the higher level of theory. This means that if the peptide chain is too short ($n \le m$) then the α-helical structure is less stable then the extended structure. However, at about the length where the first intramolecular hydrogen bond can be formed in a helix, this structure becomes more stable, and this stability increases as n becomes greater and greater than m.

The left-handed α helix shows a somewhat different energy profile. At the lower level of theory (RHF/3-21G) $E = E([\alpha_D]_n)$ behaves similarly to $E = E([\alpha_L]_n)$, thereby providing an additional example of case D but with a crossing point at $m = 4$. However, at the higher level of theory case A appears to be a more appropriate classification. At all three levels of theory investigated, the left-handed α helix is always less stable then the right-handed helical conformation. Furthermore, at the higher level of theory, $[\alpha_D]_n$ is computed to be as unstable as any other x_D-type homoconformer.

Finally, the stability of type I and II β turns embedded in an antiparallel β-sheet structure is to be discussed. Clearly, there are some minor differences in stability between type I and type II β turns. It seems that for shorter peptides type I is the more stable structure, while for longer ones the type II hairpin conformer is more stable. These differences vary over a range of 1–2 kcal mol^{-1}. On the other hand, when stability values of these β turns are compared with that of a single stranded β sheet (unfolded structure), the hairpin structures are clearly more stable if the peptide chain is long enough. At the lower level of theory even a tetrapeptide can be more stable then the unfolded structure, while at a higher level of theory a hexapeptide is more stable than a β sheet, regardless of whether a type I or type II β-turn conformation is at the reversal of the peptide chain. Although, peptides composed of more than eight residues still need to be investigated to obtain final proof, both types of major β turns look like examples of case D. In other words, turns embedded in antiparallel β sheets become a stable structure.

Simple glycine diamide models (for example, HCO-Gly-NH_2 or Ac-Gly-$NHCH_3$) contain achiral structural building units for which the conformational mirror-image structures have the same relative energy: $E(\gamma_L) = E(\gamma_D)$, $E(\delta_L) = E(\delta_D)$, etc.[12] The only exception to this rule is that of the fully extended β-pleated sheet conformation located at the very center of the Ramachandran surface; it is an unpaired structure. This symmetry also holds also for oligopeptides composed from glycine only: $E[(\gamma_L)_n] = E[(\gamma_D)_n]$, $E[(\delta_L)_n] = E[(\delta_D)_n]$, etc. In conclusion, the left-handed and right-handed α-helix structures are isoenergetic: $E[(\alpha_L)_n] = E[(\alpha_D)_n]$. Therefore, in terms of stability the glycine series, $(Gly)_n$, is much more simple than the alanine one, $(Ala)_n$. The energy curve computed for the helical $[\alpha]_n$ conformation of glycine, relative to the unfolded $[\beta]_n$ structure, is representative of case D. The crossing point of these two curves varies with the level of theory applied (Table 4). When comparing the stability of oligoglycine in its repeated γ-turns form, $[\gamma]_n$, case A is found: $E([\gamma]_n) > E([\beta_L]_n)$ at any value of n (Table 4 and Figure 3).

The clear difference between the $(Ala)_n$ and $(Gly)_n$ series is simply due to the lack of point chirality of oligoglycines. There is no possibility to distinguish the stability of mirror-image conformers because stereocenters are not present in the latter type of polymer. On the other hand, in oligoalanine, or in any other amino acid residue with a chiral α carbon atom, such energetic degeneracy of mirror-image backbone folds disappears. Thus, energy discrimination of mirror-image structural pairs of oligopeptides is due to local asymmetry or axis chirality in the presence of point chirality. In order to demonstrate this, the arithmetical average of selected mirror-image structural pairs of $(Ala)_n$ were plotted together with those computed for $(Gly)_n$ and great similarity was found (Figure 4).

In the case of oligopeptides, the question of computational accuracy is even more important than it was in the case of dipeptides such as Gly_2 or Ala_2. The haunting question is whether the inaccuracies are accumulative with the increase of the length of the polypeptide chain or whether they remain more or less constant. A practical way to assess the accuracy or reliability is to correlate results obtained at different levels of theory. This can be carried out both for geometrical parameters, like ϕ and ψ torsional angles, and relative

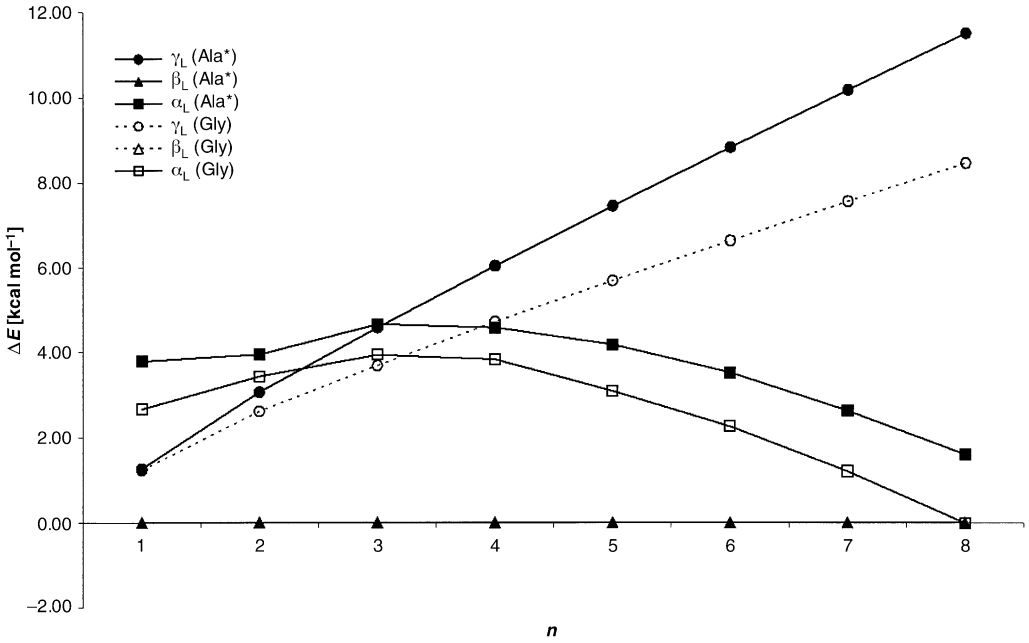


Figure 4. Relative energies of selected conformers of alanine- and glycine-containing oligopeptides computed at the RHF/6-311++G(d,p)//RHF/3-21G level of theory. Relative energies associated with left- and right-handed helical structures, $\Delta E[\alpha_D]_n$ and $\Delta E[\alpha_L]_n$, as well as those of "normal" and inverse γ turns, $\Delta E[\gamma_D]_n$ and $\Delta E[\gamma_L]_n$, of HCO-(L-Ala)$_n$-NH$_2$ were averaged ($\Delta E = 1/2[\Delta E_1 + \Delta E_2]$) and plotted with the relative energies of HCO-(Gly)$_n$-NH$_2$ against the "length" (n) of the oligopeptide ($1 \leq n \leq 8$).

energies. The correlation of the relative stabilities determined at two levels of theory for over 30 structures of oligoalanine has resulted in an R^2 value higher than 0.98.

Correlation of computed energy with X-ray structure populations: In order to perform a comprehensive analysis of the relative stability order of ab initio computed secondary structural elements and their experimental counterparts in proteins, a selected nonhomologous X-ray data set was analyzed (see Methods section for details). Table 5 summarizes the relative occurrence of typical conformational subunits (p_x) in proteins as a function of the length of the peptide chain. (see Table S4 in the Supporting Information) Figure 5 shows the raw data converted into the negative logarithm of the relative populations compared to the unfolded $[\beta_L]_n$ structure. This quantity, $-RT\ln(p_x/p_{\beta_L})$, is related to the relative energy by a Boltzmann-type distribution:

$$\Delta E = -RT\ln(p_x/p_{\beta_L})$$

Table 5. Number of the different types of backbone folds found in proteinsobserved for oligopeptides of increasing length.[a]

Type of secondary structure	Abbreviation of conformation	$(Xxx)_1$[b]	$(Xxx)_2$[b]		$(Xxx)_3$[b]	$(Xxx)_4$[b]	$(Xxx)_5$[b]
$(Xxx)_6$[b]							
inverse γ turn	$[\gamma_L]_n$	7984	546	51	6	1	–
extended	$[\beta_L]_n$	17722	3951	1120	365	127	45
	$[\delta_L]_n$	13580	970	75	8	–	–
right-handed α helix (1): repeated type I β turn	$[\alpha_L\delta_L]_n$	11178	10549	9134	6224	5519	5192
right-handed α helix (2): 3_{10} helix	$[\alpha_L{}^{3-10}]_n$	84871	65632	54874	47572	41687	36342
right-handed α helix (3): α helix	$[\alpha_L]_n$	79831	61540	51935	45212	39513	34285
right-handed collagen helix: polyproline II	$[\varepsilon_L]_n$	31414	7873	1926	520	145	42
left-handed α helix	$[\alpha_D]_n$	7080	713	43	3	–	–
	$[\delta_D]_n$	734	23	2	–	–	–
left-handed collagen helix	$[\varepsilon_D]_n$	1376	26	–	–	–	–
"normal" β turn	$[\gamma_D]_n$	868	14	–	–	–	–
type of β turn at the end of antiparallel β-sheets	β turn type II/A	–	1245	147	33	6	4
	β turn type II/B	–	1245	844	33	14	4
	β turn type I/A	–	10549	340	24	12	8
	β turn type I/B	–	10549	822	24	4	8

[a] The data set used contains 1211 nonhomologues proteins composed from a total of 206889 amino acid residues. $p_x/p_{\beta L}$ = the number of x type conformer relative to the number of extended structures (β_L) observed for the oligopeptide composed from n amino acid residues. [b] Xxx stands for any of the 20 natural amino acid residues that are found in proteins.

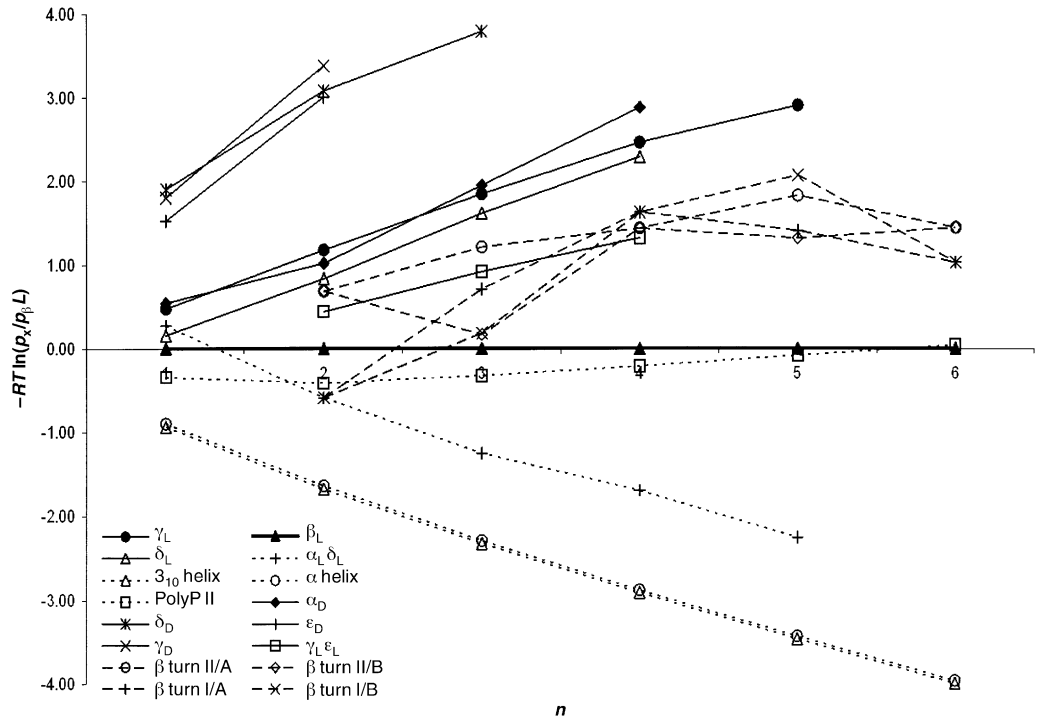


Figure 5. The $-RT\ln[p_x/p_{\beta_L}]$ values of the different types of backbone folds typical in proteins, observed as a function of the length of the oligopeptide chain (see Table 4). A total of 206889 amino acid residues were analyzed from 1211 nonhomologues proteins. The type of the amino acid was ignored when the occurrence of the conformational types were counted. p_x/p_{β_L} is the number of the x type of conformer relative to the number of extended structures (β_L) observed for the oligopeptide composed from n amino acid residues.

For selected conformers, results (see Figure S3 in the Supporting Information) demonstrate clearly that a significant correlation can be obtained between ab initio computed relative energies and experimental abundance. The size of the protein database used, although large, gives natural limitation. As mentioned above, the repeated inverse γ-turn form, $[\gamma_L]_n$, is a stable structure but its relative stability decreases with the increase of the length of the polypeptide chain. The significant correlation found between experimental and theoretical probabilities (see Figure S3A in the Supporting Information) of $[\gamma_L]_n$ for both the Gly and Ala series suggests that for longer peptides this type of conformer is less and less populated because it is less and less stable relative to the unfolded main-chain structure. Furthermore, the correlation established both for the right-handed and the left-handed α helix (see Figure S3B and C in the Supporting Information) suggests that the intrinsic stability and instability of these structures is primarily due to local energetic preferences of the backbone conformation. Undoubtedly, side-chain backbone interaction, molecular cooperativity of different secondary structure elements, and different environmental factors will fine tune the final fold of the macromolecule. However, local energetics play a crucial role when secondary structure elements are selected.

Conclusions

HCO-(L-Ala)$_n$-NH$_2$ peptides were introduced as the simplest model systems of chiral oligo- and polypeptides. In the case of the alanine series, regardless of the level of theory applied, the most stable secondary structure element among the investigated conformers for an oligopeptide composed of more than six residues is the right-handed α helix. Furthermore, both type I and II β turns embedded in antiparallel β-pleated sheets of different length become more stable then the unfolded fully extended reference structure. Solitary polyproline II (a component of the collagen triple helix) is an unstable structure on its own in the case of the presently used polyalanine model, thus it was not part of our present analysis. Although, the repeated inverse γ-turn form contains a series of intramolecular hydrogen bonds, it is slightly less stable then the reference unfolded structure. Finally, all oligo-L-alanine homoconformers composed of "D-type" conformational subunits are significantly less stable then their "L-type" conformational counterparts, especially with the increase of the polypeptide chain length. All of these facts agree perfectly with the accumulated data and are generally accepted in structural biology. Thus, the folding of peptides on their own, as well as the selection of these well-known secondary structural elements for the folded proteins, is in fact due to the intrinsic properties of the peptide chain. Even for the simplest chiral oligopeptide and even in a vacuum these secondary structure elements are found to be folded in the same shape and geometry that are commonly observed in proteins. Furthermore, the relative stability order of these intrinsically stable structures shows significant correlation with the natural abundance of the same conformers assigned in a nonhomologous protein database.

Acknowledgement

This research was partially supported by grants from the Hungarian Scientific Research Fund (Grant: OTKA T030841). The authors thank Péter Hudáky, Anna K. Füzéry, and Árpád Kiss for helpful discussions.

[1] *Gaussian 98 (Revision A.7)*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, Ö. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komáromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc., Pittsburgh, PA, **1998**.
[2] Ö. Farkas, H. Bernhard Schlegel, *J. Chem. Phys.* **2003**, in press.
[3] P. Pulay, *Chem. Phys. Lett.* **1980**, *73*, 393 – 398.
[4] P. Pulay, *J. Comp. Chem.* **1982**, *3*, 556 – 560.
[5] P. Császár, P. Pulay, *J. Mol. Struct. (Theochem)* **1984**, *114*, 31 – 34.
[6] G. Endrédi, A. Perczel, Ö. Farkas, M. A. McAllister, G. I. Csonka, J. Ladik, I. G. Csizmadia, *J. Mol. Struct. (Theochem)* **1997**, *391*, 15 – 26.
[7] K. Möhle, H.-J. Hoffman, *J. Mol. Model.* **1998**, *4*, 53 – 60.
[8] A. Perczel, Ö. Farkas, I. Jákli, I. A. Topol, unpublished results.
[9] U. Hobohm, M. Charf, R. Schneider, C. Sander, *Prot. Sci.* **1992**, *1*, 409 – 417.
[10] U. Hobohm, C. Sander, *Prot. Sci.* **1994**, *3*, 522 – 524.
[11] C. Ramakrishnan, G. N. Ramachandran, *Biophys. J.* **1965**, *5*, 909.
[12] A. Perczel, J. G. Ángyán, M. Kajtár, W. Viviani, J.-L. Rivail, J.-F. Marcoccia, I. G. Csizmadia, *J. Am. Chem. Soc.* **1991**, *113*, 6256 – 6265.
[13] IUPAC/IUB Commission on Biochemical Nomenclature, Abbreviations and symbols for description of conformation of polypeptide chains—Tentative rules: *Biochemistry* **1969**, *9*, 3471.
[14] I. Simon, L. Glasser, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 3661 – 3665.
[15] I. L. Shamovsky, G. M. Ross, R. J. Riopelle, *J. Phys. Chem. B* **2000**, *104*, 11 296 – 11 307.
[16] C. Park, W. A. Goddard III, *J. Phys. Chem. B.* **2000**, *104*, 7784 – 7789.
[17] M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren, R. A. Friesner, *J. Am. Chem. Soc.* **1997**, *119*, 5908 – 5920.
[18] P. Derreumaux, *Theor. Chem. Acc.* **2000**, *104*, 1 – 6.
[19] F. Forcellino, P. Derreumaux, *Struct. Funct. Genet.* **2001**, *45*, 159 – 166.
[20] G. Némethy, H. A. Scheraga, *Biopolymers* **1965**, *3*, 155.
[21] F. A. Momany, R. F. McGuire, A. W. Burgess, H. A. Scheraga, *J. Phys. Chem.* **1975**, *79*, 2361 – 2381.
[22] S. J. Leach, G. Nemethy, H. A. Scheraga, *Biopolymers* **1966**, *4*, 369.
[23] J. C. Howard, A. Ali, H. A. Scheraga, F. A. Momany, *Macromolecules* **1975**, *8*, 607 – 622.
[24] S. S. Zimmerman, M. A. Pottle, G. Nemethy, H. A. Scheraga, *Macromolecules* **1977**, *10*, 1 – 9.
[25] J. N. Scarsdale, C. Van Alsenoy, V. J. Klimkowski, L. Schäfer, F. A. Momany, *J. Am. Chem. Soc.* **1983**, *105*, 3438 – 3445.
[26] L. Schäfer, C. Van Alsenoy, J. N. Scharsdale, *J. Chem. Phys.* **1982**, *76*, 1439 – 1444.
[27] L. Schäfer, V. J. Kimkowski, F. A. Momany, H. Chuman, C. Van Alsenoy, *Biopolymers* **1984**, *23*, 2335 – 2347.
[28] L. Schäfer, S. Q. Newton, M. Cao, A. Peeters, C. Van Alsenoy, K. Wolinski, F. A. Momany, *J. Am. Chem. Soc.* **1993**, *115*, 272 – 280.
[29] A. M. Sapse, B. S. Schweitzer, A. P. Dicker, J. R. Bertino, V. Frecer, *Int. J. Pept. Prot. Res.* **1992**, *39*, 18 – 23.
[30] A. Perczel, I. G. Csizmadia, *The amide linkage, structural significance in chemistry, biochemistry and materials sciences* (Eds.: A. Greenberg, C. M. Breneman, J. F. Liebman), Wiley, New York, **2000**, pp. 409 – 463.
[31] A. G. Császár, A. Perczel, *Progr. Biophys. Mol. Biol.* **1999**, *71*, 243 – 309.
[32] M. A. Berg, G. A. Chasse, E. Deretey, A. K. Füzéry, B. M. Fung, D. Y. K. Fung, H. Henry-Riyad, A. C. Lin, M. L. Mak, A. Mantas, I. V. Repyakh, M. Staikova, S. J. Salpietro, T. H. Tang, J. C. Vank, A. Perczel, G. I. Csonka, Ö. Farkas, L. L. Torday, Z. Székey, I. G. Csizmadia, *J. Mol. Struct. (THEOCHEM)* **2000**, *500*, 5 – 59.
[33] A. Perczel, M. A. McAllister, P. Császár, I. G. Csizmadia, *J. Am. Chem. Soc.* **1993**, *115*, 4849 – 4858.
[34] M. A. McAllister, A. Perczel, P. Császár, I. G. Csizmadia, *J. Mol. Struct. (THEOCHEM)* **1993**, *290*, 181 – 198.
[35] A. Perczel, M. A. McAllister, P. Császár, I. G. Csizmadia, *Can. J. Chem.* **1994**, *72*, 2050 – 2070.
[36] C.-H. Yu, L. Schafer, M. Ramek, *J. Phys. Chem. A.* **1999**, *103*, 8337 – 8345.
[37] A. Perczel, G. Endrédi, M. A. McAllister, Ö. Farkas, P. Császár, J. Ladik, I. G. Csizmadia, *J. Mol. Struct. (THEOCHEM)* **1994**, *331*, 5 – 10.
[38] I. A. Topol, S. K. Burt, E. Deretey, T.-H. Tang, A. Perczel, A. Rashin, I. G. Csizmadia, *J. Am. Chem. Soc.* **2001**, *123*, 6054 – 6060.

Received: February 12, 2003 [F4843]