

Chapter 2

Quantum Chemical Calculations on Small Protein Models

Imre Jákli, András Perczel, Béla Viskolcz and Imre G. Csizmadia

2.1 Ab Initio Quantum Chemistry of Peptides

During the past 50 years (1963–2013) many thousands ab initio computations were published on small peptides. Many molecules contain the acid amide or peptide bond ($-\text{CO}-\text{NH}-$) but the smallest molecule is formamide ($\text{HCO}-\text{NH}_2$). It might be expected that the first ab initio computations were to be carried out on formamide. However, in 1963, when digital computers, such as IBM 709, were not transistorized, therefore, only the iso-electronic formyl fluoride (HCOF) was possible to be subjected to ab initio Molecular Orbital (MO) computations. The results of this first computation were reported in 1963 in the Quarterly Progress

I. Jákli (✉) · A. Perczel
MTA-ELTE Protein Modelling Research Group Pázmány, Péter sétány,
1/A H-1117 Budapest, Hungary
e-mail: jimre@chem.elte.hu

A. Perczel
Laboratory of Structural Chemistry and Biology, Institute of Chemistry, ELTE,
Pázmány Péter sétány, 1/A H-1117 Budapest, Hungary
e-mail: perczel@chem.elte.hu

B. Viskolcz · I.G. Csizmadia
Department of Chemical Informatics, University of Szeged, Boldogasszony sgt. 6,
H-6725 Szeged, Hungary

I.G. Csizmadia
Department of Chemistry, University of Toronto, M5S 3H6, Toronto, ON, Canada
e-mail: icsizmad@jgypk.u-szeged.hu

Report of MIT [1]. The full research was published [2] in 1966, after the IBM Research Centre generously offered some time on their, transistorized, IBM 7090 computer to finish the largest basis set computation of that time. The original MIT report was kindly reproduced by an ACS journal in the Supplementary Material of the paper published, on the 50th anniversary, in the Journal of Physical Chemistry B [3]. Eventually, the original aim, the formamide molecule (HCO-NH_2) was computed afterward when the University Toronto acquired an IBM 7094 transistorized computer and a paper was published [4] in 1968.

Soon after the gate has opened, an unbelievable volume of computation appeared in the literature which created the feeling of a *Molecular Revolution*. From that point onward force-field softwares are continuously redesigned on the basis of ab initio computational data. Without the aspiration for completeness, a brief historic summary is given, in this chapter, about the historic development of Quantum Chemical Calculations on small peptides which may be treated as a prototype model of the protein folding problem. However, more emphasis will be given to the current trends that will effectively influence future directions.

2.2 Relative Stability of Peptide Bonds Formed

The thermodynamic stability of an acid amide or peptide bond may be measured by the reaction heat ($\Delta_r H$) or free energy change ($\Delta_r G$ of its format) with respect to the sum of the appropriate measure(s) of acid and amine residues involved.

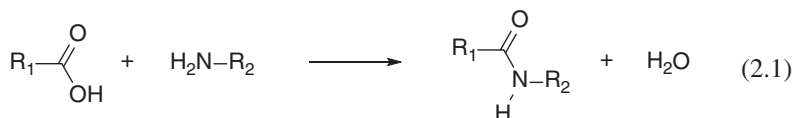
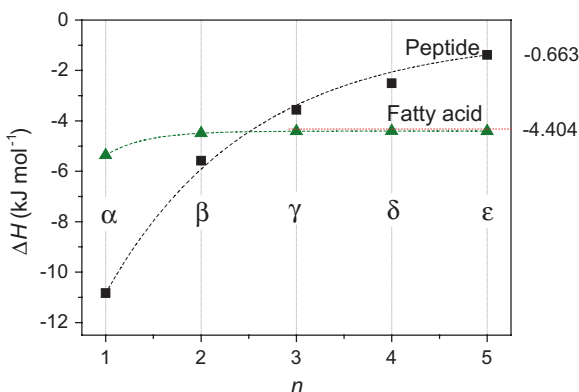


Figure 2.1 shows that the $\Delta_r H$ formation of amide from α -amino acid ($n = 1$) is the most exothermic process and thus, the most special of the whole series.

Fig. 2.1 Variation of $\Delta_r H$ for the formation of acid amide linkage, according to Eq. (2.1), with the chain length (n), for fatty acids: $\text{H}-(\text{CH}_2)_n-\text{COOH}$ and for N-acetyl amino acids: $\text{AcNH}-(\text{CH}_2)_n-\text{COOH}$. The length of the chain is also indicated in a symbolic fashion by a sequence of letters of the Greek alphabet

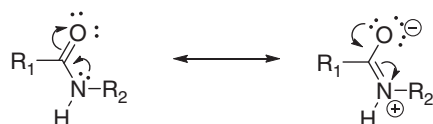


This figure also indicates that the convergence of the two curves at $n = 5$ is less than 4 kJ mol^{-1} . The convergence limits in Figs. 2.1 and 2.2 are shown on their right hand side. The convergence implies that beyond a certain length the chain makes no noticeable influence on the energetics.

Variation of $\Delta_r H$ for cyclic peptide formation is quite different as shown in Fig. 2.2.

Interestingly enough, both γ - and δ -cyclic peptides show extensive stabilization with respect to their linear form as a 5 or a 6 member ring is formed, respectively.

The peptide bond is a 4π electron containing functional group, as an allyl molecular system would do, the nitrogen lone pair forms a partial double bond with the carbonyl carbon:



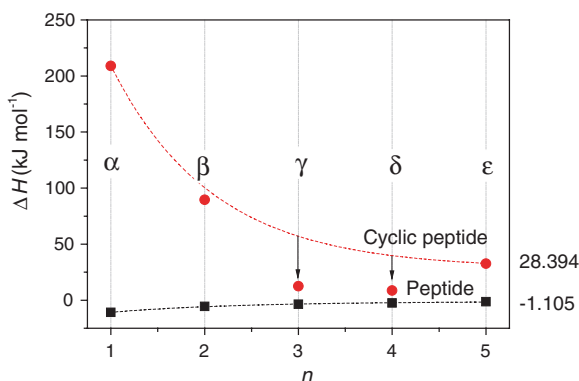
The fact that the acid amide ($-\text{CONH}_2$) is close to planarity is frequently rationalized by the above pair of resonance structures. The extent of this conjugation may be characterized numerically by a *conjugativity* value which is frequently referred to as *amidicity*. As shown below, the removal of the conjugation, by the hydrogenation of the $\text{C}=\text{O}$ double bond, may be used to define the *amidicity* (AM) through the calculation of the enthalpy change (ΔH_{H_2}). A relative value of ΔH_{H_2} may be used as a percentage value of amidicity defined in terms of the following compounds shown in Fig. 2.3. Calculations of the amidicity scale [5] is given by the following equations:

$$\Delta H_{\text{H}_2}[I] = H_B - H_A \quad (2.2)$$

$$[\text{Amidity \%}] = m \Delta H_{\text{H}_2}[I] + [\text{Amidity \%}]_0 \quad (2.3)$$

Figure 2.4 shows the variation of amidicity both for acyclic (open chain) and cyclic peptides.

Fig. 2.2 Variation of $\Delta_r H$ for the formation of acyclic and cyclic peptides as function of the chain length (n), for amino acids: $\text{H}_2\text{N}-(\text{CH}_2)_n-\text{COOH}$



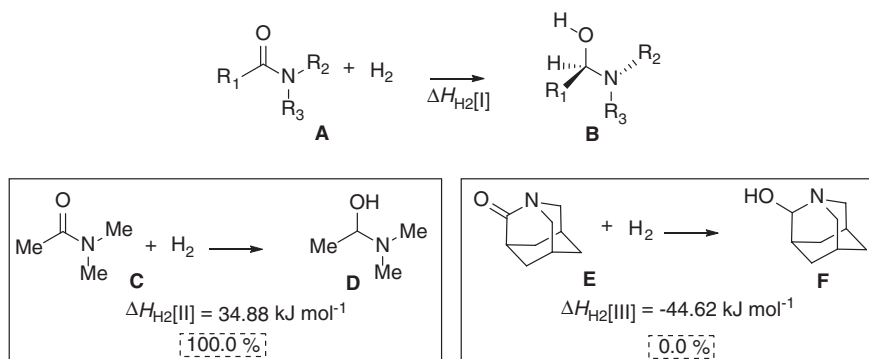
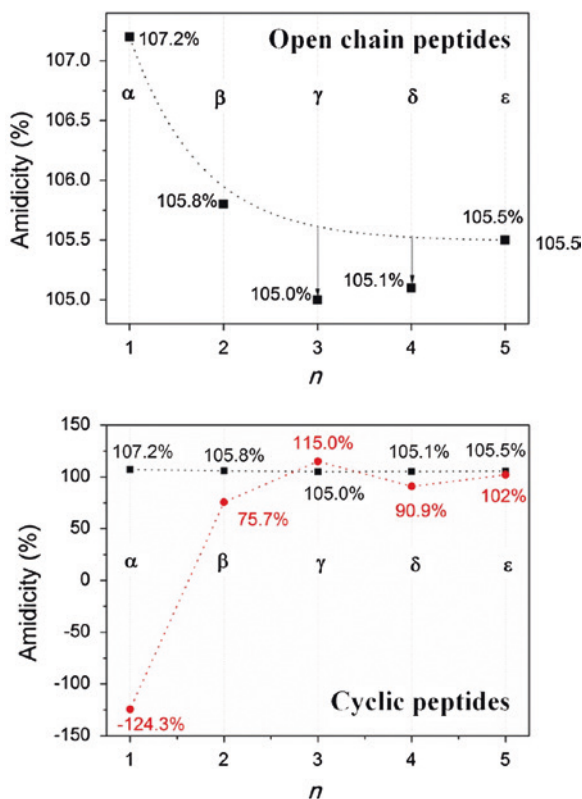


Fig. 2.3 The definition of amidicity

Fig. 2.4 Variation of amidicity (%) both for acyclic (*squares points*) and cyclic (*circular points*) peptides



Both the 5- and 6-membered ring, formed by γ - or δ -amino acids respectively show extra stability even though the difference is within 0.5 %. In view of this, it is understandable that glutamic acid can preferably form cyclic amide. In addition, asparagine residue via deamination process can form succinimide (Fig. 2.5).

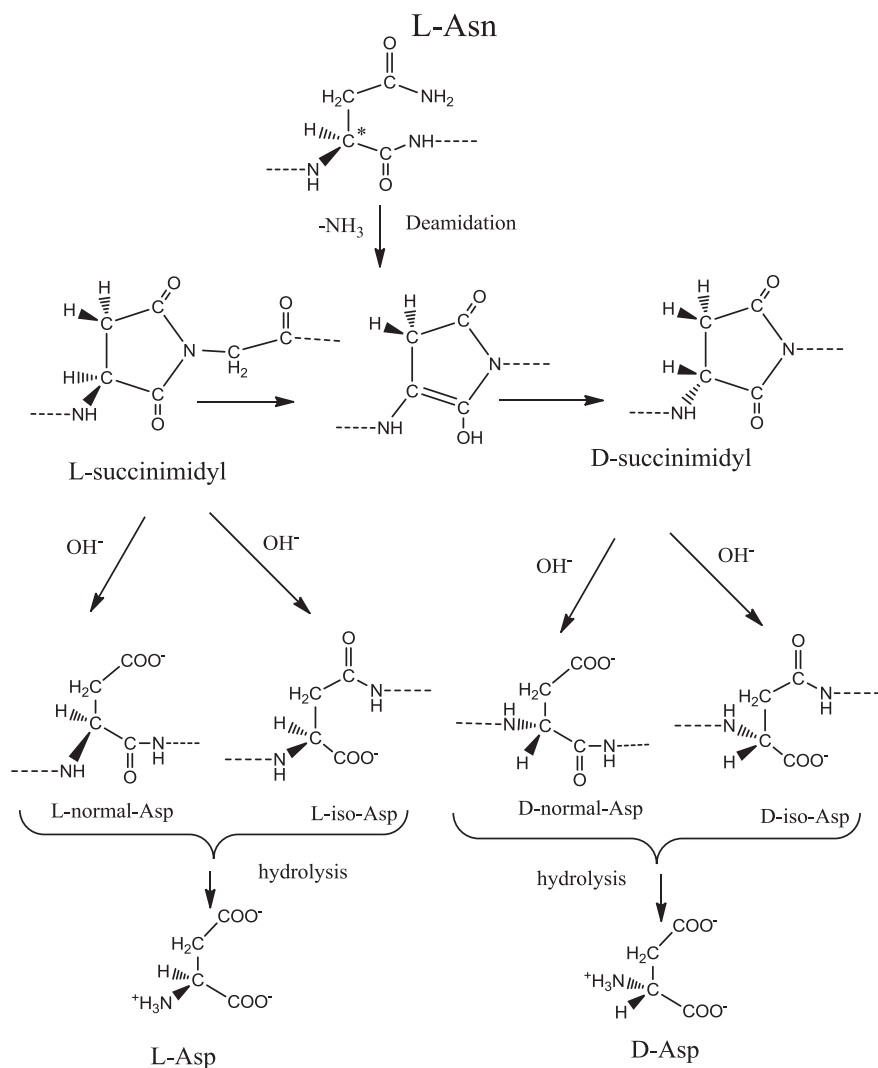


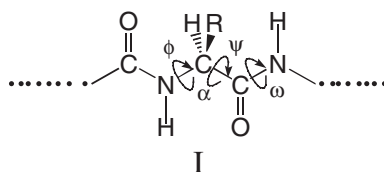
Fig. 2.5 Closed electronic shell mechanism of D-Asp formation from Asn through deamidation and enolization

The asparagine deamidation ($\text{Asn} \rightarrow \text{Asp}$) is one of the most important protein degradation pathways and residues of Asp serve as “molecular timers” that can have effects on protein turnover and aging [6–10]. Recently, Trout et al. [11] and subsequently Catak et al. [12] studied, theoretically, the deamidation process of the Asn. The racemization was subsequently explained by enolization of the $\text{H}-\text{C}=\text{O}$ moiety [13], as shown in some details in Fig. 2.5.

2.3 Topology of Peptide Conformations

2.3.1 The Concept of the Ramachandran Map

The question of protein folding is a century old problem. About a half a century ago, Ramachandran [14] in India tried to establish what conformations of a single residue diamide were disallowed for steric reason. In the absence of digital computers at that time only ball and stick molecular models were available. The conformational change is measure by two dihedral angles (ϕ , ψ) associated with the rotation about two bonds connected to the α -carbon as shown below:



The potential energy turns out to be a 2D potential energy surface (PES) [15] that is a mathematical function of two independent variables.

$$E = f(\phi, \psi) \quad (2.4)$$

The shape of the 2D PES can be investigated from its 1D-crossections (Fig. 2.6).

The idealized topology of minimum energy points is illustrated by Fig. 2.7.

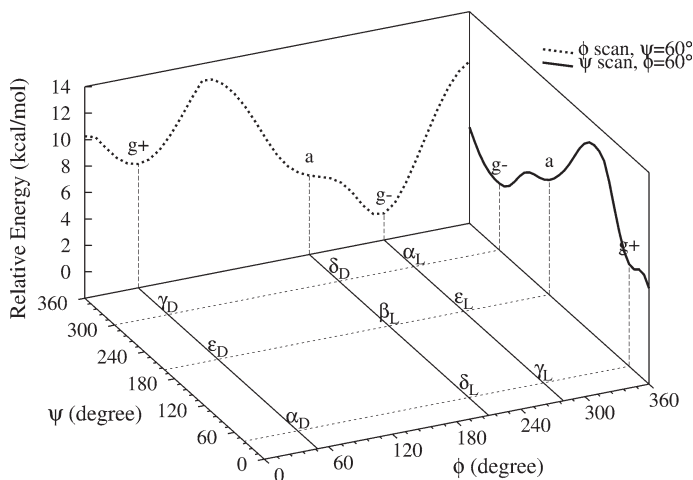


Fig. 2.6 Schematic illustration of how a chiral (PES) may be related to two component chiral potential energy curves

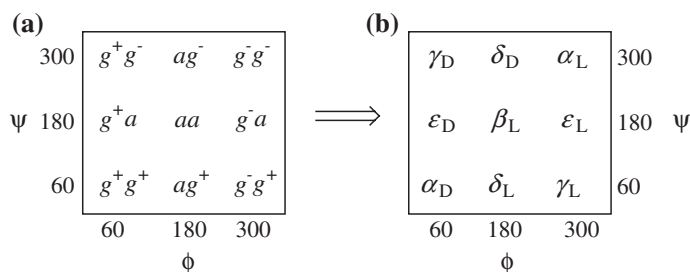


Fig. 2.7 Conformational assignments (a) and names (b) of peptide conformers on the conformational (PES) of a peptide (P-CONH-CHR-CONH-Q) [16]

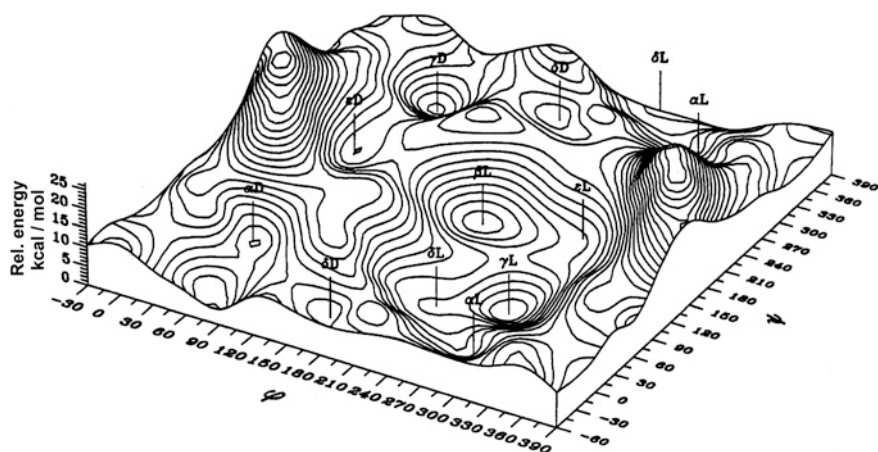


Fig. 2.8 Pseudo-three-dimensional Ramachandran potential energy surface of HCONH-CHCH₃-CONH₂, presented in the $0^\circ \leq \phi \leq 360^\circ$ and $0^\circ \leq \psi \leq 360^\circ$ range of two independent variables (2D)

2.3.2 Conformational Potential Energy Surfaces (PES)

It may be important to amplify that every point on the above surface (Fig. 2.8) is a particular conformation, but as a distinction only the minima are called conformers. The ab initio conformation potential energy surface as any similar surfaces calculated at different levels of theory or by using different methods (e.g. MM, QM/MM, coarse grain), looks like a landscape (Fig. 2.8).

Just like in cartography the landscape may also be shown as a contour diagram. The above surface (Fig. 2.8) represents one of the four equivalent quadrants in Fig. 2.9.

These conformers (Table 2.1 and Fig. 2.10) could be regarded as typical building blocks of folded proteins as the analysis of PDB data [18] have shown it. Most

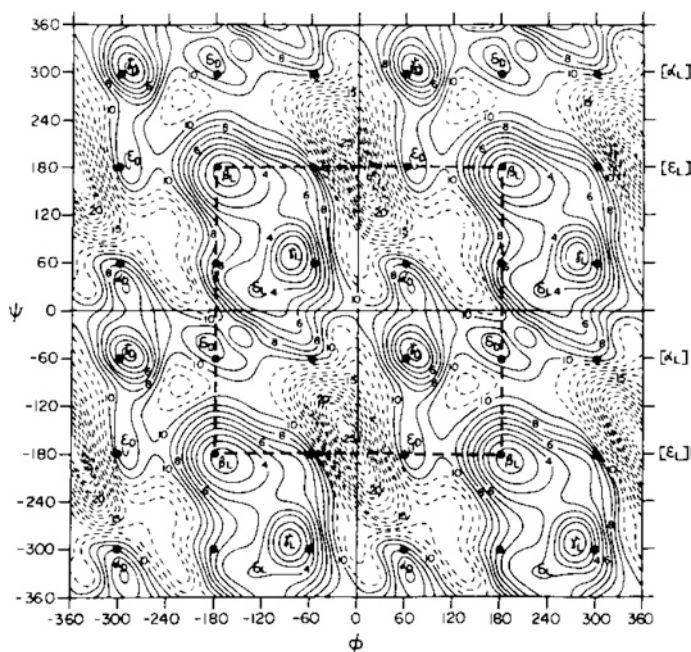


Fig. 2.9 Contour diagram of the 2D Ramachandran potential energy surface of HCONH-CHCH₃-CONH₂, presented in the $-360^\circ \leq \phi \leq 360^\circ$ and $-360^\circ \leq \psi \leq 360^\circ$ range of independent variables. The central square (*broken lines*) is the IUPAC conventional cut, while the four quadrants are the traditional cuts

Table 2.1 Optimized ϕ , ψ torsional angle pairs for alanine diamide (HCONH-(L)-CHMe-CONH₂)

| Conformer | Optimized values | | Idealized values [16] | | Conformational Classification |
|--------------|------------------|--------|-----------------------|--------|-------------------------------|
| | ϕ | ψ | ϕ | ψ | |
| α_L | -66.6 | -17.5 | -60 | -60 | g^-g^- |
| α_D | +61.8 | +31.9 | +60 | +60 | g^+g^+ |
| β_L | -167.6 | +169.9 | -180 | +180 | aa |
| γ_L | -84.5 | +68.7 | -60 | +60 | g^-g^+ |
| γ_D | +74.3 | -59.5 | +60 | -60 | g^+g^- |
| δ_L | -126.2 | +26.5 | -180 | +60 | ag^+ |
| δ_D | -179.6 | -43.7 | +180 | -60 | ag^- |
| ϵ_L | -74.7 | +167.8 | -60 | +180 | g^-a |
| ϵ_D | +64.7 | -178.6 | +60 | -180 | g^+a |

The idealized torsional angle pairs, together with their conformational classification, are also shown for the sake of comparison

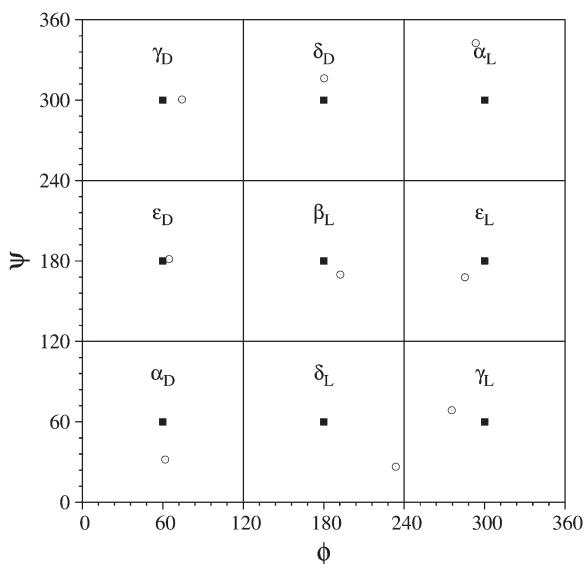


Fig. 2.10 a Schematic illustration of the PES of an amino acid diamide. The idealized positions of the main conformers are marked by *shaded squares*, while the computationally determined positions (actual location depending on methods, type of approach, level of theory etc.) are shown as *open circles*. Shifts but even elimination of some points was observed. The names of the conformers are given as subscripted Greek letters [16, 17]

frequently β_L and its neighboring conformers (e.g. ϵ_L, γ_L) is a component of the extended β -pleated sheets while α_L is a typical building unit of both 3_{10} - and/or α -helix.

2.3.3 Mathematical Representation of Conformational PES

Peptide folding, just like protein folding, has a conformational aspect. The mathematical representation of the conformational potential energy surface is a tool that may be used to decipher problems related to peptide folding. Consequently, fitting mathematical functions for computed grid points will lead to a mathematical representation of such a conformational problem.

One dimensional trigonometric fit to simple internal rotation energies has been fitted by Pople et al. [19, 20]. For potential energy surfaces (2D) and hypersurfaces (3D or nD) functions of two, three or n independent variable are necessary. These were achieved for relatively simple surfaces [21–23]. More recently the problem

has been reinvestigated to see how does the mathematical complexity of Fourier expansion function groups with the complexity of the appearance of the PEC or PES [24]. The three different torsional modes of C–C, N–N and O–O exhibit three different scenarios. The N–N rotation is the most complex because the central anti conformer at 180° becomes a higher energy minimum with respect to the two gauche minima (at around 90° and 270°). In contrast to N–N, rotation about the O–O bond in H₂O₂ the anti conformer becomes a TS and two equivalent conformers were assigned at about 120° and 240°, respectively. Thus, this PEC exhibited intermediate complexity. Ethane has 3 identical minima hydrogen peroxide has 2 identical minima. In contrast to these hydrazine has 3 minima of which 2 are identical lower minima and the 3rd minimum is a higher minimum. All of these apparent complexities are summarized in Table 2.2, in which it is indicated that C–C can be fitted by a single ($m = 1$) cosine term, while peroxide requires already two cosine terms ($m = 2$) and hydrazine must have at least three terms ($m = 3$) in the Fourier expansion (2.5).

$$E(\varphi) = a_0 + \sum_{m=1}^n \left(a_j \cos \frac{m2\pi k\varphi}{360} \right) \quad (2.5)$$

The three conformational PECs are shown in Fig. 2.11.

Table 2.2 Accuracy (R^2) of fitted functions with increasing number of terms (m) for a C–C, N–N and O–O bonds

| Molecule | R^2 | | |
|-----------|---------|---------|---------|
| | $m = 1$ | $m = 2$ | $m = 3$ |
| Ethane | 0.9936 | 1.0000 | 1.000 |
| Peroxide | 0.9427 | 0.9994 | 1.000 |
| Hydrazine | 0.7739 | 0.9802 | 0.9977 |

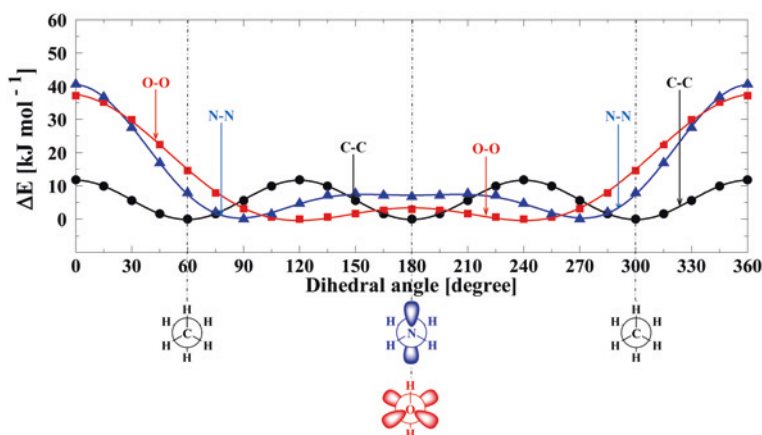


Fig. 2.11 Potential energy curves for rotation about the C–C, N–N and O–O bonds [24]

Thus, with increasing complexity (increasing topological differences of critical points) of the PEC the complexity of the fitted explicit mathematical function must also increase. For a conformation PES, pentane is a classical example since it may be treated as 1,3-dimethyl propane and therefore it could act as a 2D example (Fig. 2.12).

The surface is quite symmetric so a cosine and sine compilation of trigonometric functions (6) with $m, n = 3$ yielded on $R^2 = 0.9347$. With the increase m, n the R^2 value improved slightly. The R^2 value improved with $m, n > 3$ as shown in Table 2.3.

$$\begin{aligned}
 E(\varphi, \psi) = a_0 + \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (a_1 \cos mk_{\varphi} \varphi \cos nk_{\psi} \psi \\
 + a_2 \cos mk_{\varphi} \varphi \sin nk_{\psi} \psi + a_3 \sin mk_{\varphi} \varphi \cos nk_{\psi} \psi \\
 + a_4 \sin mk_{\varphi} \varphi \sin nk_{\psi} \psi)
 \end{aligned}
 \quad (2.6)$$

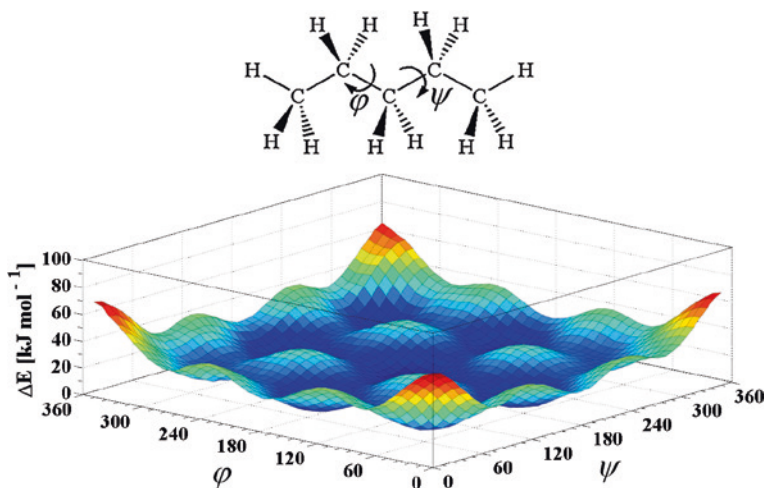


Fig. 2.12 Conformational PES of pentane treated as 1,3 dimethyl-propane

Table 2.3 Slight variation of accuracy (R^2) of fitted functions with increasing number of terms (m, n) for pentane (1,3 dimethyl propane)

| Molecule | R^2 | | | |
|----------|------------|------------|------------|------------|
| | $m, n = 3$ | $m, n = 4$ | $m, n = 5$ | $m, n = 6$ |
| Pentane | 0.9347 | 0.9367 | 0.9371 | 0.9383 |

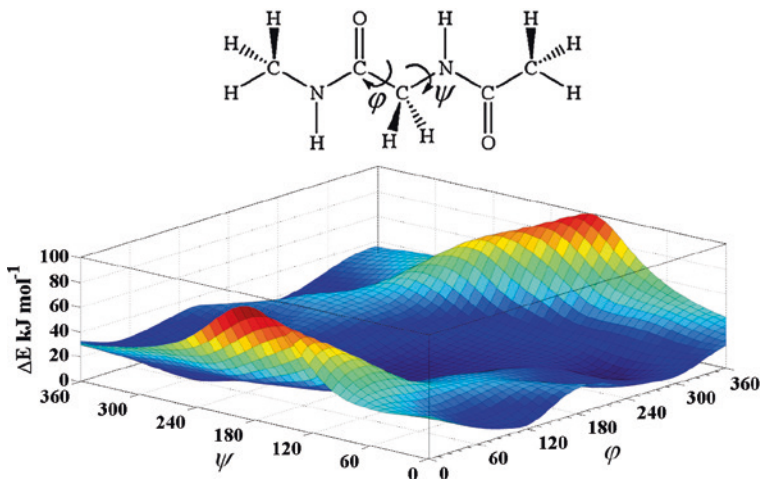


Fig. 2.13 Ramachandram type conformational PES of N-acetyl glycine-N-methylamide

A similar type of fitting for the conformational PES of glycine diamide is already a more complex task: surface shown in Fig. 2.13.

$$\begin{aligned}
 E(\varphi, \psi) = & a_0 + \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{m,n} \epsilon \left(- \left(\frac{(b_m \varphi - \varphi_0)^2}{2\sigma_\varphi^2} + \frac{(b_n \psi - \psi_0)^2}{2\sigma_\psi^2} \right) \right) \\
 & + a_1 \cos mk_\varphi \varphi \cos nk_\psi \psi + a_2 \cos mk_\varphi \varphi \sin nk_\psi \psi \\
 & + a_3 \sin mk_\varphi \varphi \cos nk_\psi \psi + a_4 \sin mk_\varphi \varphi \sin nk_\psi \psi \\
 & + \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} d_1 \cos(jk_\varphi \varphi - lk_\psi \psi) d_2 \cos(jk_\varphi \varphi - lk_\psi \psi) \\
 & + \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} d_3 \cos(jk_\varphi \varphi - lk_\psi \psi) d_4 \sin(jk_\varphi \varphi - lk_\psi \psi) \\
 & + \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} d_5 \sin(jk_\varphi \varphi - lk_\psi \psi) d_6 \cos(jk_\varphi \varphi - lk_\psi \psi) \\
 & + \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} d_7 \sin(jk_\varphi \varphi - lk_\psi \psi) d_8 \sin(jk_\varphi \varphi - lk_\psi \psi)
 \end{aligned} \tag{2.7}$$

Even though the fitted analytical Eq. (2.7) is far more complicated than that used for pentane (Eq. 2.6) the R^2 value is about the same: $R_{Gly}^2 = 0.9287$ as was in the former case. However, the main problem of such an approach is not that the analytical function gets more and more complicated as the polypeptide growth in size and complexity, but that the PES varies a lot as the level of theory is changed [25, 26]. Thus, as both shape and topology of the appropriate PEHS changes for

the very same chemical entity by varying QM method and/or the level of theory applied, the concept of using these PEHS as markers of protein building blocks seems to miss the point. Although very informative, these PEHS can hardly be used as generalized descriptors of folding properties of polypeptides.

2.3.4 Toroidal Representation of Conformational PES

The two independent periodic variables (ϕ , ψ), which characterize the Ramachandran map, denote circular motions as illustrated by Fig. 2.14. Therefore, to make any arbitrary cut off introduces artificial edges and thus, separates otherwise neighboring conformer types from each other.

To obtain the Ramachandran surface in a toroidal coordinate system we may first roll up the 2D PES along one variable, (e.g., ψ) and obtain a cylinder (Fig. 2.15, top). Subsequently the same can be done along the second variable ϕ ,

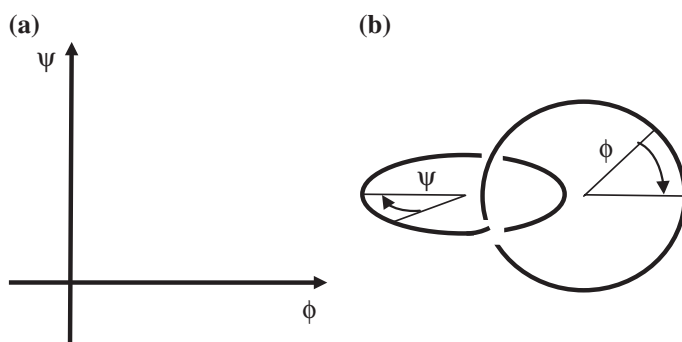


Fig. 2.14 Traditional Cartesian-coordinate system (a) and a topologically equivalent circular or toroidal coordinate system (b)

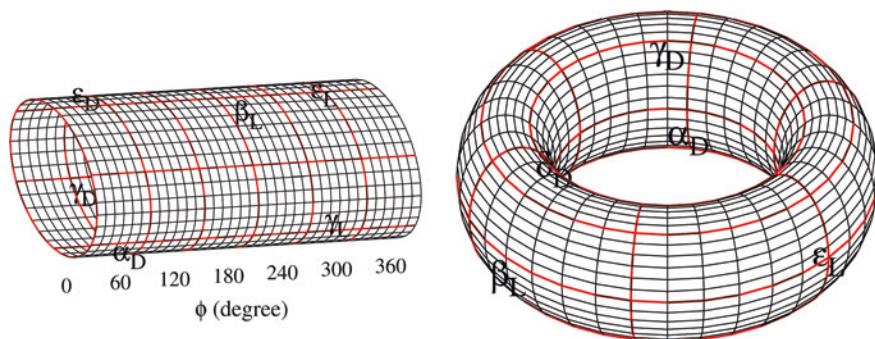


Fig. 2.15 Folding of the Ramachandran-surface coordinates first along ϕ (top) and second along ψ (bottom), with minima (α_L , β_L , etc.) highlighted

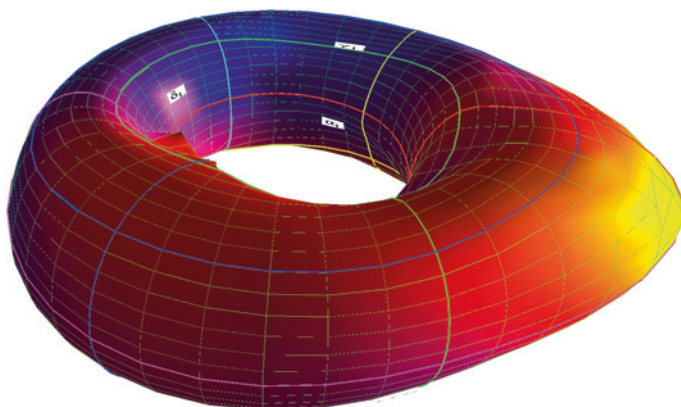


Fig. 2.16 Toroidal representation of the Ramachandran PES for alanine residue in embedded in the AcNH–(L)–CHMe–CONHMe diamide model

to obtain a doughnut-shaped object, such as an O-ring called a 2D-torus (Fig. 2.15, bottom). Location of the nine basic conformers [17], which is visible from the given perspective (Fig. 2.8) are indicated by their subscripted Greek-letter codes occur now on the surface of the torus.

If we wish to illustrate the energy depths or heights, shown in Fig. 2.8, on these topographically equivalent geometrical objects, then we may use an appropriate color code. Also, it is easy to squeeze these geometrical objects at their energy minima and blow them up at points related to energy maxima [27] (Fig. 2.16).

The situation is analogous to view the Earth as a 3D-globe rather than on a 2D map. It can be seen that points that are located on two sides of a 2D-PES map may in fact adjacent on the 3D-toroid. In other words the periodicity of dihedral rotation are explicitly observable.

To summarize both the traditional as well as any more recent representation of the backbone PES of any amino acid diamide, called as Ramachandran surface, are adequate toll for depicting the different conformational building blocks, lego elements, of peptides and proteins. Moreover, conformational interconversions (paths) via TSs as well as avoided regions (mountains) are easy to visualize, using any sensible cuts and representations. On the other hand, these PESs are very sensitive and changes qualitatively to the method, level of theory, concept etc. used to determine them. Thus, just because these PESs are very method dependent, they can hardly be used as numerical and/or analytical representation of the actual peptide model. In other words for the very same peptide model, multiple PESs can be calculated. In conclusion, these PESs are often useful tool to qualitatively interpret conformational properties but are inadequate to stand on their own for a polypeptide model.

2.4 Backbone Conformations of Small Peptides

2.4.1 Single Amino Acid Diamides

There are 21 DNA coded natural amino acids occurring in proteins. A total of 20 of these amino acids have DNA codons and the 21st of them, Selenocystein (Sec), has only RNA codon. For the description of backbone conformations all of these amino acids need the appropriate 2D-Ramachandran potential energy surfaces (PES). Two of these 21 amino acids have exceptional structures: (i) Glycine (Gly) which has no side chain and thus achiral and (ii) Proline (Pro) in which the side chain is connected back to the amino nitrogen atom fixing the ϕ torsional angle at around 70° . For the remaining 19 proteogenic residues side chains are of varying length and therefore have different degree of conformational freedom. The 1, 2, 3 and 4 dihedral angles add 1, 2, 3 and 4 extra dimensions to the basic 2D Ramachandran PES of the residue. In this way we may end up as $2 + 1 = 3D$, $2 + 2 = 4D$, $2 + 3 = 5D$ and $2 + 4 = 6D$ PEHS per residue. Side chains are classified as being, apolar, polar or explicitly charged as function of the pH. Characteristics of the various amino acids are summarized in Table 2.4. Hetero atom(s) of selected side chains are of importance as they can “self” interact and form intramolecular side chain-backbone interaction(s). The nature of these interactions can be observed even in smaller peptides. However, in proteins there are many more type of interactions (e.g. side/chain-side/chain). These inter-residue interactions require larger peptides to be used as model systems as shown below.

The classification of amino acid residues is also possible via their side-chain topologies. In 19 out of the 20 proteinogenic amino acids the first substituting atom are always carbon. In 16 of the 19 cases the C atom is a methylene (CH_2) group, with one substituent on it, while in 3 cases it is a methine (CH) group, equipped with two substituents. These structural features are shown below.

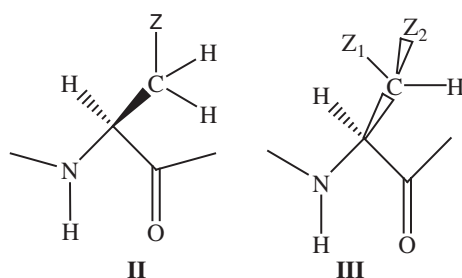


Table 2.4 list of 21 amino acid residues showing their side chains and their dihedral angles of rotation as well as the dimensionality (D) of the full Ramachandran PES.

Table 2.4 List of the 21 proteinogenic α -amino acids, of variable side-chains lengths

| Amino Acids with Electrically Charged Side Chains | | | | | |
|---|----------------------|----------------------|----------------------|----------------------|--|
| 4D Arg | 2D His | 4D Lys | 2D Asp | 3D Glu | |
| Positive | | | Negative | | |

| Amino Acids with Uncharged Side Chains | | | | | |
|--|----------------------|----------------------|----------------------|----------------------|----------------------|
| 2D Ser | 2D Sec | 2D Cys | 2D Thr | 2D Asn | 3D Gln |

| Amino Acids with Hydrophobic Side Chain | | | |
|---|----------------------|----------------------|----------------------|
| 2D Tyr | 3D Met | 2D Ile | 1D Val |
| 1D Ala | 2D Leu | 2D Phe | 2D Trp |

| |
|---------|
| Pro |
| Gly |

These amino acids can be coupled to each other by forming polypeptides via condensation reaction (Eq. 2.1).

Initially, glycine (Gly) and alanine (Ala) diamide backbone conformers were studied in details [17, 28, 29]. Basis set study for the assessment of the reliability

of the used methods of computation were carried out [30]. For selected number of amino acids certain conformers have been identified by matrix isolation infrared spectroscopy [25, 26]. Valine diamide was also studied quite early [31]. Since the α -helical (g^-g^-) conformer was not a minimum on the initial PES, therefore various hydrogen bonded structures were generated via the direct solvation method with the inclusion of a single H_2O molecule [32]. This set up was satisfactory to obtain the α -helix minimum on the PES. Serine (Ser) became very popular [33–38] during the years. Cystein (Cys), the sulfur analogue of Ser was investigated considerably later. Actually Cys was computed on its own [39] as well in connection with disulfide bridge formation providing cysteine [40, 41]. The selenium analogue, selenocystein (Sec) was also in focus on its own merit [42] as well for its antioxidant role [43]. Two more sulfur containing amino acids were also investigated, one of them was methionine (Met) [44, 45] and the other one was its demethylated form: homocystein [46]. In addition in their side-chain oxygen containing amino acid diamides were also studied by various QM methods, for example those included threonine (Thr) [47, 48] and hydroxy-proline [49]. In terms of other polar side chains, asparagine (Asn) [50, 51] and glutamine (Gln) [52] as well as their corresponding acids, Aspartic acid (Asp) [53–56] and glutamic acid (Glu) [52] were also studied.

Conformational properties of aromatic side chain containing amino acids, like phenylalanine (Phe) [57–60] and tyrosine (Tyr) [61] and special apolar amino acid residue like proline (Pro) [62] were also elaborated. As a special case of apolar side chain dehydro-alanine [63] in which a $C=C$ double bond exist between the α and β carbon atoms has also been investigated. Also special attention was given to the question of *trans* \rightarrow *cis* isomerization of the peptide bond [64, 65].

To discuss all the conclusions of the above individual publication goes beyond the limit of this chapter. The readers are asked to explore them in individual publications. However at least one common conclusion can be drawn from all these conformational stability studies completed on these different side-chain equipped backbone units of peptides and proteins, namely that neither the side chain chemical composition, nor their conformational properties will fundamentally distort the backbone topology of a $-\text{CONH}-(L)-\text{Xxx}-\text{NH}-$ molecular building unit. Thus, the fact that all these proteinogenic residues are α - and L-amino acid residues provides for them a common backbone conformational characteristic.

2.4.2 Dipeptide Diamides

In the early 1990th amino acid dimers concentrated on the simplest dialanine diamide models [66–68], such as -Ala-Ala- and were used to study the first folded systems (e.g. β -turns) by ab initio methods [69]. QM approaches combined with the conformational preferences of the constituent building blocks, known β -turn structures were approved and new forms of turns were revealed. Although found less frequently in proteins, these new β -turns (e.g. $\epsilon_D\delta_L$, $\epsilon_D\alpha_L$, $\delta_L\beta_L$) do exists and

were successfully assigned in polypeptides and proteins. Thus, multidimensional conformational analysis driven *ab initio* calculations concluded, that 1 \rightarrow 4-type intramolecular H-bond of β -turns is more related to the preferred α_L -type substructure of the first residue within a β -turn, than to the true nature of the turn-like folded backbone structure. These early QM calculations resulted in already total at least 18 different turn-like foldamers for the For-L-Ala-L-Ala-NH₂ model system. Additional β -turn structure analysis revealed [70] the conformational preferences of more complex turn models, such as Pro-Thr [71] and Pro-Pro [72]. These studies revealed how fixing the first torsional angle of the backbone of a dipeptide initiates β -turn formation.

2.4.3 Oligo- and Polypeptide Diamides

Oligopeptides containing 1–5 amino acids were also investigated during the past couple of decades. These studies included either solely alanine residues [73, 74] or Ala in combination with other proteinogenic residues [75]. Antifungal Phe-Arg-Trp [76] was perhaps considered and its conformational preference was established. In addition, the question of beta sheet [77] and α -helix [78–81] stability has also been addressed. Point mutations at the central position of the Pro-Pro-Pro [72] sequence was in focus as the hinge region of immunoglobulin has such sequential properties (e.g. Pro-Xxx-Pro) (Fig. 2.17).

These tripeptide units are already nano-structures in terms of size and they may assume a number of conformations as illustrated by the next figures (Figs. 2.18 and 2.19).

If in the figure below the substituent R is either H or Ph the tripeptide could adopt a relatively long and rather extended backbone structure. However if the central amino acid is Thr, then the internal hydrogen bonds could fold the molecule resulting in a more compact backbone arrangement (Fig. 2.20).

Special attention was given to glutathion [82–85] (Fig. 2.21), as it can form both α - and γ -peptide bond.

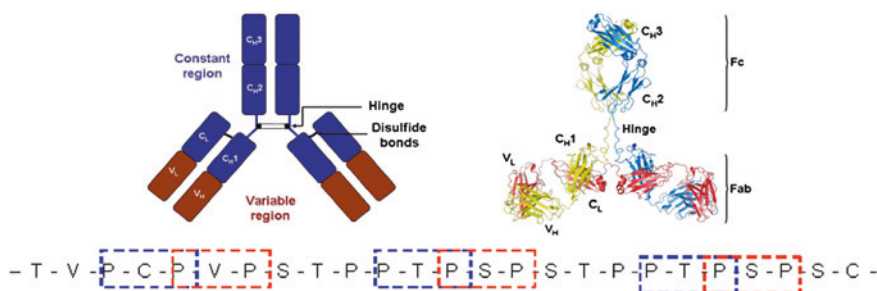


Fig. 2.17 Hinge region location and PXP peptide sequence of immunoglobulin

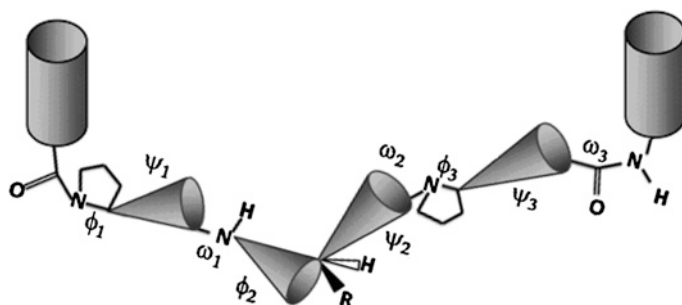


Fig. 2.18 A schematic representation of conformational variations of a PXP tripeptide

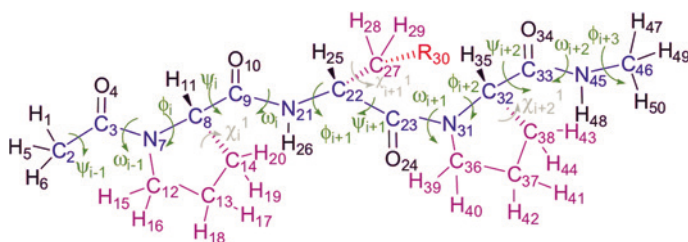


Fig. 2.19 Structural representations of a PXP tripeptide

Fig. 2.20 Side chain/backbone hydrogen-bonding networks available in the different conformers of HCO-Pro-Thr-Pro-NH₂

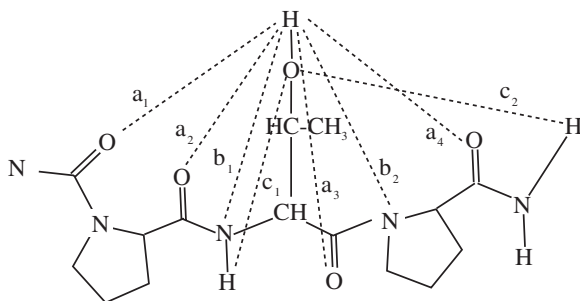
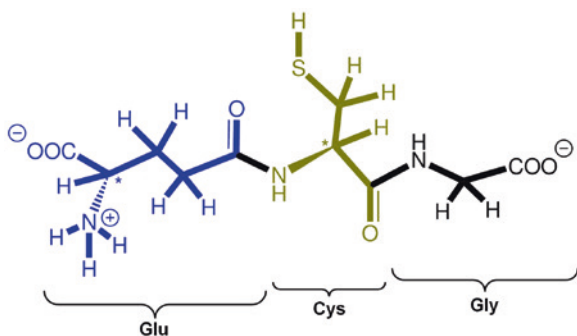


Fig. 2.21 A γ -peptide bond in Glutathione



The γ -glutamyl-cystein present in glutathione (γ -Glu-Cys-Gly) is an ancient sequence bit, which has a special role in virtually all living organism as an important antioxidant [86]. It prevents damage to important cellular particles caused by reactive oxygen such as free radicals. In some cases this dipeptide (γ -Glu-Cys) occurs alone in halobacteria [87, 88]. In plants certain mutants occur where the C-terminal glycine is replaced by β -alanine [89], serine [90] or glutamate [91]. All of these suggest that the γ -Gly-Cys moiety was first formed in the ancient period of molecular evolution during which the more stable γ -peptide bond was formed under thermodynamic control.

Collagen build up from -Pro-Gly-HyPro- or POG “triplets” in a repetitive manner has a well characterized triple helical secondary structure, where the chains are stabilized by interchain H-bonds, using the HN of Gly. The relative stability of fully optimized collagen triple helices was determined with respect to a three stranded β -pleated sheet structure by using DFT calculations. In addition the secondary structure preference of Pro, Gly, Ala and HyPro residues was established [92]. De novo calculated collagen structures show a great resemblance to those determined by X-ray crystallography. Interestingly enough the calculated triple helix formation affinities correlate well with the experimentally determined stabilities retrieved from melting point data. The very abundant collagen is not only special by presenting a triple helical structure, but also it is specifically and intensively hydrated [93]. Bella and Berman reported experimentally the structure of the first hydration layer and found that H_2O s form bridges of different length and type around the POG repeats of collagen. Stability and helicity of these hydration layers were computationally determined via 8–12 explicit placed water molecules. Although the stability order of these waters varies from binding places, but they do it in line with the X-ray data. In conclusion, these water binding places on the surface of the triple helix can provide explanation on how an almost liquid-like hydration environment exists between the closely packed tropocollagens [94]. Using the ab initio data it was speculated that these water molecules could serve as reservoirs or buffers providing space for “hole conduction” of water molecules and thus, contribute to the elasticity of collagen known macroscopically for quite some time [95].

Amyloid-like aggregates made of extended like backbone folds of simple polypeptides were also studied by ab initio methods [96]. Accumulated evidences on conformational diseases (e.g. Alzheimer’s disease) show the presence of amyloid aggregates, found independent of the primary sequence of the polypeptide chain [97]. Thus the driving force of the conversion from the original to amyloid type foldamer of a primary sequence unit is most probably driven by favorable backbone backbone interactions [98, 99]. In this way most polypeptides and proteins gravitate to an unexpected and highly irreversible thermodynamic minima. They assemble in form of supramolecular nano-systems, within the component macromolecules adopt a common form called as “dead-end” structure(s). Using MDC driven QM calculations on large enough, but still computable polypeptides, it was found that β -pleated sheet structure(s) dominate the “dead-end” molecular foldamer. Several very different building block forms were probed and found both

in vacuum and in aqueous environments that their di-, oligo- and polymers make amyloid like fibers. Even in a crystalline state (periodical, tight peptide attachment), the β -pleated sheet assembly remains the most stable superstructure. This theoretical study provides a quantum-level explanation for why proteins can take the amyloid state when local structural preferences jeopardize the functional native and often global folds. The ribbon form of such a nano-assembly is shown in Fig. 2.22.

While the β -pleated sheet structure (Fig. 2.22) made of β -layer is typical of an oligopeptide made up of α -amino acid residues, for oligo and polymers of β -amino acids, a self-rapping and thus a spontaneous (energetically favored) nanotubes formation was observed [100]. Octapeptide structure segments of the longer oligopeptide, Penetratin, have been studied computationally as well as by NMR spectroscopy. The computed and the NMR observed structural results agreed well [101].

Certain oligopeptides that have two cysteine residues may be cyclized via disulfide bond formation. In nature, oxytocin is an important example for such cyclic structures (Fig. 2.23a). The relative stability of disulfide bridges that may lead to cyclization or to dimerization (Fig. 2.23b) has been also studied recently [41]. Oligopeptides such as the mammalian neurohypophyseal hormone, oxytocin

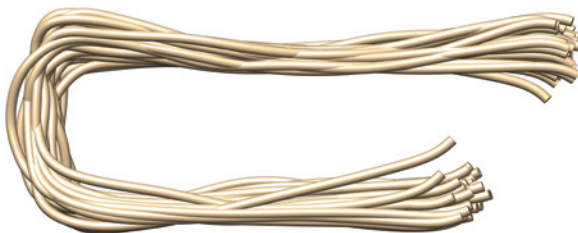


Fig. 2.22 A schematic “side” view of an amyloid aggregate formed from dozens of polypeptide chains

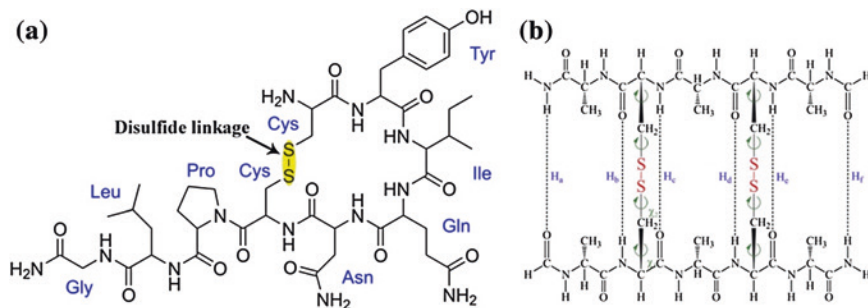


Fig. 2.23 a The structure of oxytocin. b Inter-chain disulfide linkage

and vasopressin (Fig. 2.23b) contain two cysteines and by forming an intramolecular cystine they result in a peptidic macrocycle; composed of six residues of 20 atoms in total within the ring. This poses some conformational limitations to them found ideal for binding to their receptors. The relative stability of disulfide bridges that may lead to cyclization or dimerization (Fig. 2.23b) has been also studied recently and found that [41].

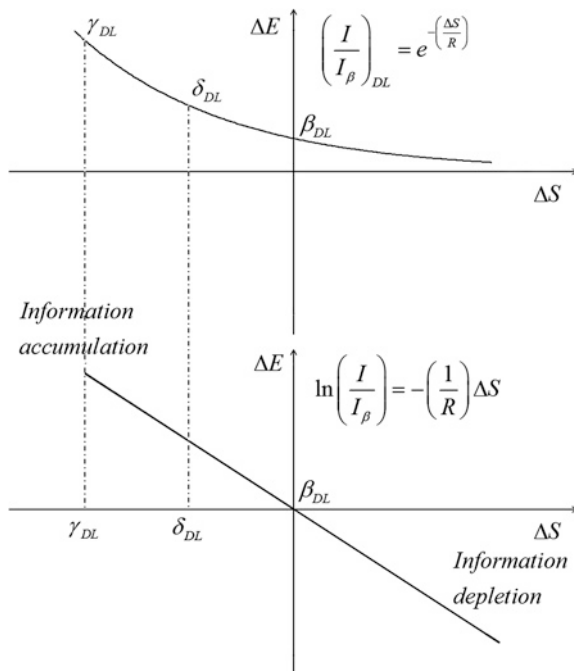
2.4.4 Information Accumulation During Polypeptide Folding

The information content change associated with peptide folding has been studied by computing the associated entropy change (ΔS) [80, 102–105] (Figs. 2.24 and 2.25).

The variations of ΔS with extent of polymerization is shown in Fig. 2.26 and the variations of thermodynamic functions in Fig. 2.27. Both figures indicate that the entropy change of poly glycine is faster with increasing degree of polymerization than that of poly alanine.

All these entropy focused studies suggest that in folded conformers there is net information accumulation with respect to the unfolded structure.

Fig. 2.24 Variation of relative information content (I/I_0) with ΔS



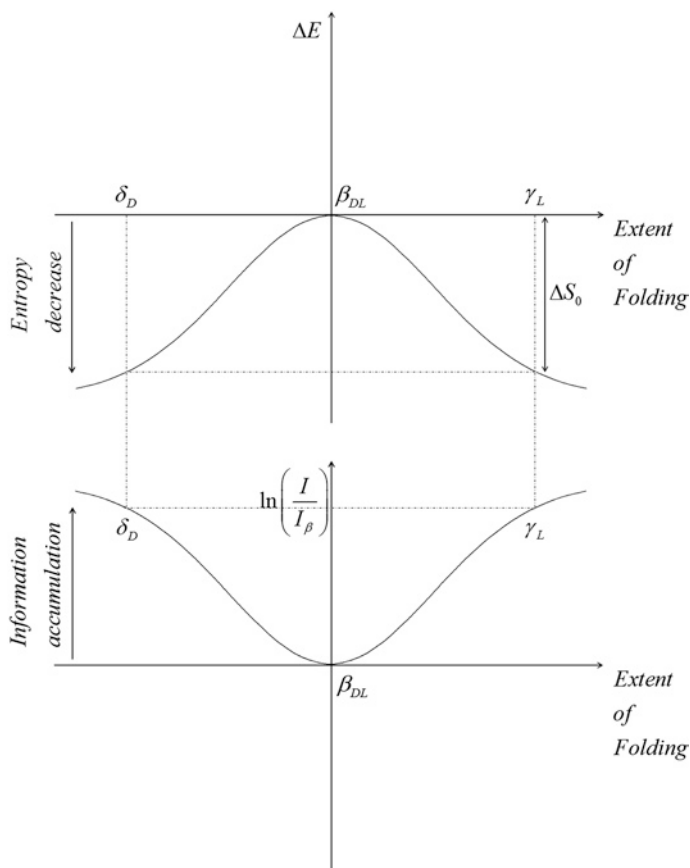
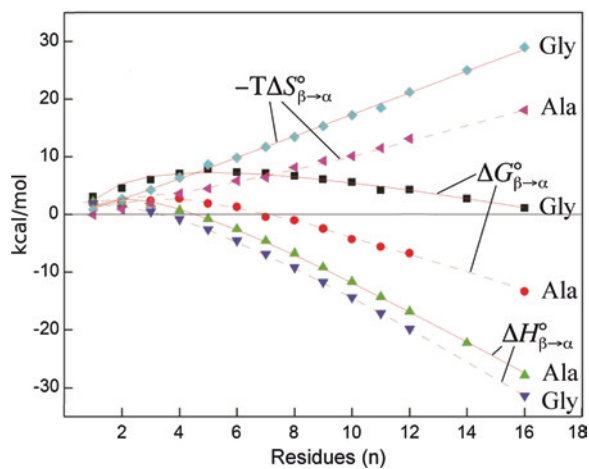


Fig. 2.25 A schematic illustration how ΔS and $\ln(I/I_0)$ varies with peptide folding

Fig. 2.26 Variation of thermodynamic functions, $\Delta H^\circ_{\beta \rightarrow \alpha}$, $-T\Delta S^\circ_{\beta \rightarrow \alpha}$, $G^\circ_{\beta \rightarrow \alpha}$ of $\beta_L \rightarrow \alpha_L$ folding with degree of polymerization



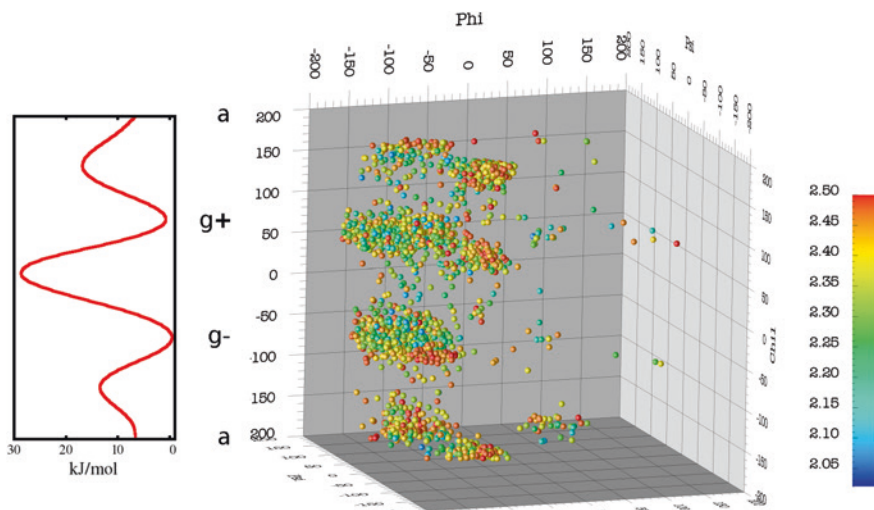


Fig. 2.27 PDB statistics of the distance between the oxygen and the hydrogen in the C-H...O sidechain-backbone hydrogen bond in valine residues of proteins. On the *left hand side* the approximate potential energy curve is shown indicating the heavier population of the gauche conformers

2.5 Side Chain Conformations of Small Peptides

Any conformational interconversion of For-L-Ser-NH₂ takes place on its 4D hypersurface: $E(\varphi, \psi, \chi_1, \chi_2)$. Both visualizing and making any graphical representation of a 4D-hyperspace are complex. Thus, appropriate 2D-crosssections of the parent 4D hyperspace were calculated and pasted “together” to decipher the topology associated with -Ser- conformers and conformational changes. The analysis of these ab initio conformation energy maps made the tracing of some relaxation paths possible, revealing characteristic side-chain-induced backbone conformational shifts [81]. Interestingly, the partly relaxed conformational hypersurfaces revealed alternative relaxation paths.

2.5.1 A Case Study for Apolar Side Chain (Val)

Beside Alanine the simplest apolar amino acid is Valine which residue requires a 3D PEHS as it has single side-chain dihedral angle beside the backbone φ and ψ variables. Figure 2.27 shows how the energetically more stable g^+ and g^- side chain rotamers are frequent in proteins, with respect to the less stable and thus less frequent a side chain form.

For-L-Val-NH₂ could have in total 27 legitimate minima on its 3D Ramachandran map, $E = E(\varphi, \psi, \chi_1)$. By QM calculations as many as 20 conformers were optimized [31]. A new method was developed for energy partitioning in order to quantify the magnitude of the side chain/backbone interaction probed for the present L-Val model system. Such a side chain/backbone interaction was established by calculating the iPr group for the various backbone conformers of For-L-Val-NH₂ relative to that of hydrogen in the corresponding backbone folds of For-Gly-NH₂. The comprehensive analysis showed that even an apolar side chain is able to interact with the peptide backbone so “strongly” that it could annihilate otherwise legitimate backbone minima.

Both L-Val and L-Phe residues, prototypes of hydrophobic aliphatic and hydrophobic aromatic amino acid residues, were studied at several basis sets by using different methods (e.g. B3LYP/6-311 ++G**), resulting in a larger dataset compiling results of different levels of theory [106]. Both conformational and energetic properties of these “libraries” were analyzed as a function of the method applied. In addition comparisons of calculated populations of peptide foldamers of these hydrophobic residues were matched with their natural abundance derived from proteins. Analysis concluded that at least for the hydrophobic core of proteins, the conformations of Val (Ile, Leu) and Phe (Tyr, Trp) are controlled by the local energetic preferences of the respective amino acid residues.

2.5.2 A Case Study for Polar Side Chain (Asn and Asp)

Both in the case of Asn as well as Asp residues there are hydrogen bonds between side chains and backbone (Fig. 2.28). Since the H atom of the –COOH group is more protic than that of the –CONH₂ moiety, the hydrogen bond is shorter in the case of Asp than in the case Asn. For assessing all H-bonded structures a full 4D potential energy hypersurface would need to be analysed.

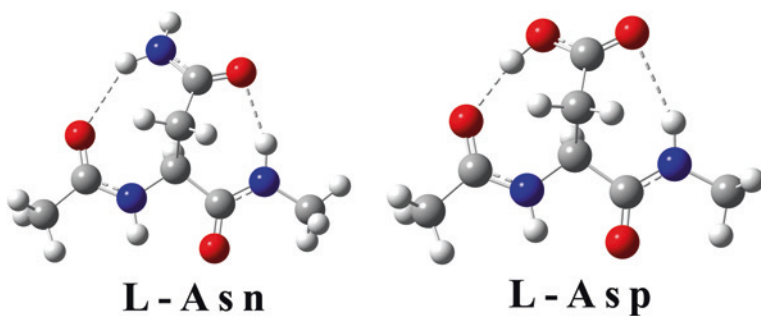


Fig. 2.28 Sidechain-backbone N–H...O hydrogen bond in Asn and OH...O hydrogen bond in Asp residues

2.5.3 A Case Study for Protonated Side Chain (His)

Proton affinity and pKa values of N-formyl-L-His-NH₂ are found to vary as a function of its backbone and/or side-chain orientation [107]. Examples were presented and confirmed by ab initio calculations, where proteins were crystallized under various pH conditions, resulting in the same histidine residue to adopt different conformers. Furthermore, a hypothesis is given for a protonation-induced conformational modification of the histidine residue in the catalytic triad of chymotrypsin during catalysis, which lowers the pKa value of the catalytic histidine by 1.2 units. Both the experimental and theoretical results support that proton affinity as well as that pKa values of histidine residues are strongly conformationally dependent.

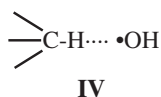
2.6 Peptide Radicals

Accumulated evidence indicates that oxidative stress plays a significant role in a number of diseases such as Parkinson's disease (PD), Alzheimer's disease (AD), Diabetes II and atherosclerosis, just to list a few of the more than 50 examples [108–110]. The mechanisms leading to cellular oxidative stress has been shown to be the result of the excessive production of reactive oxygen species (ROS), that includes non radicals (i.e. H₂O₂) and free radicals (i.e. OH, O₂^{•−}, and NO) as well. Consequently, ROS can interact with different bioactive molecules, to initiate a cascade of events that can lead to cell death [111]. In this way, protein oxidation is a result of hydrogen abstraction by hydroxyl radicals [111]. Under normal conditions, ROS are generated are eliminated by the cell's antioxidant capacity [112]. When the antioxidant capacity is no longer reduce capable to the reduce ROS excess they can accumulate in the cell. This can cause the oxidation of the amino acid residues within the protein backbone, which can result in (a) protein fragmentation (b) a change from the L-configuration to the D-configuration (c) protein aggregation and (d) protein misfolding [113]. Altered protein structure have been observed in a neurodegenerative disease such as Alzheimer's disease, which is generally found in the elderly [114].

It was once thought that all living organisms are composed of only L-amino acids [115]. The discovery of D-aspartic acids (D-Asp) in various human tissues of the elderly people indicates that oxidative stress related to ageing is a main factor in the production of D amino acids. Such a configurational change can result in the accumulation of the D-amino acid and decrease the “original” enantionerically pure protein concentration [116–119]. In addition, the accumulation of D-amino acids in the brain is affiliated with Alzheimer's disease [116].

2.6.1 Radical Structures and Reactivity

One of the damages that oxidative stress causes at a molecular level is the hydrogen abstraction of the α -hydrogen atom of an amino acid either by hydroxy radicals (IV) or by other reactive oxygen species (ROS).



It has been shown [120] that the hydrogen atom attached to the α -carbon is the most vulnerable part of proteinogenic residues for such a damaging attack. It has also been suspected that the radical center of a polypeptide chain would react differently as function of the main chain folding. Thus thermodynamic measures of the very same reaction would be different for a folded, unfolded and aggregated nano-system. Penta-glycine [121] and penta-alanine were studied in this respect. The hydrogen radical, ($\text{H}\bullet$), recapturing by the $\text{C}\alpha$ -radical, has been investigated [109]. The following Eq. (2.8) summarizes

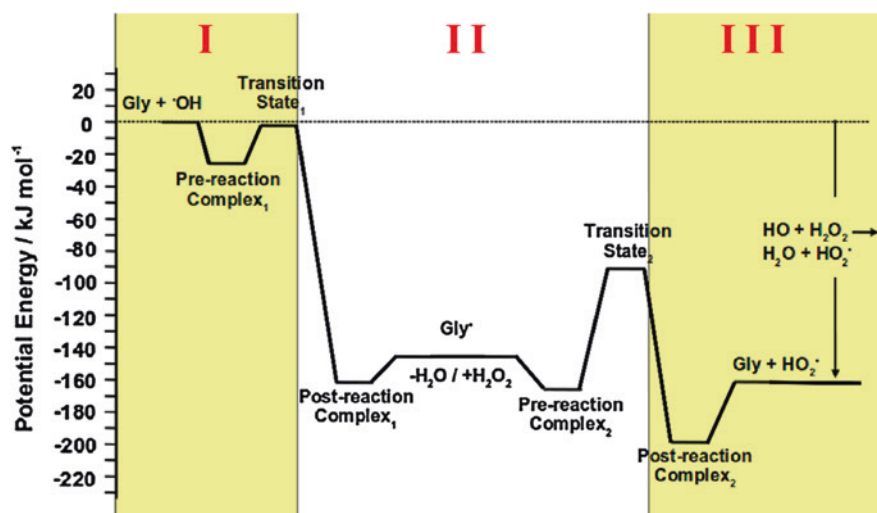


Fig. 2.29 H atoms are less likely to form α -radical followed radical unfolding and ended with H-atom capture of glycine-diamide. Potential Energy variation of the process is which the α -radical of glycine was formed by OH radical and the H-atom was donated the α -radical by H_2O_2 . The roman numerals on the top indicating the initial, intermediate and final states

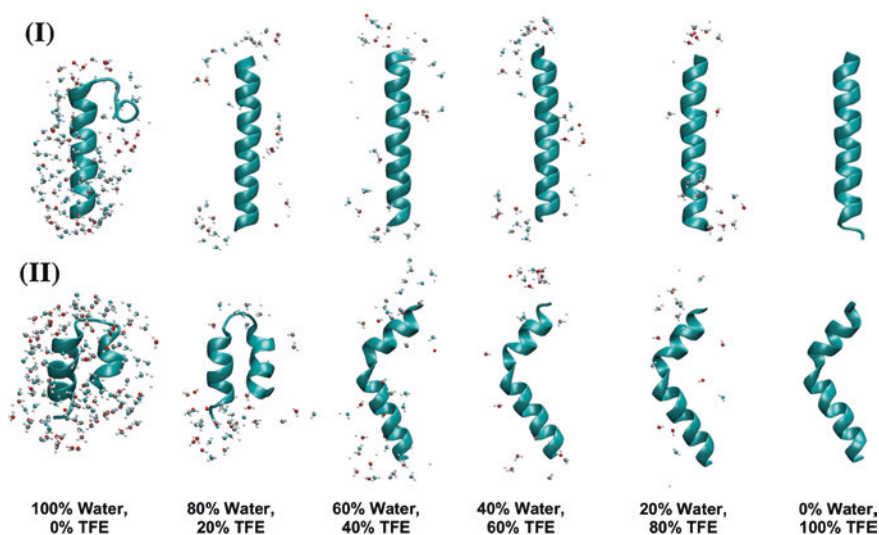


Fig. 2.30 **a** Change of polypeptide conformation from water to TFE solution. **b** Change of polypeptide radical conformation from water to TFE

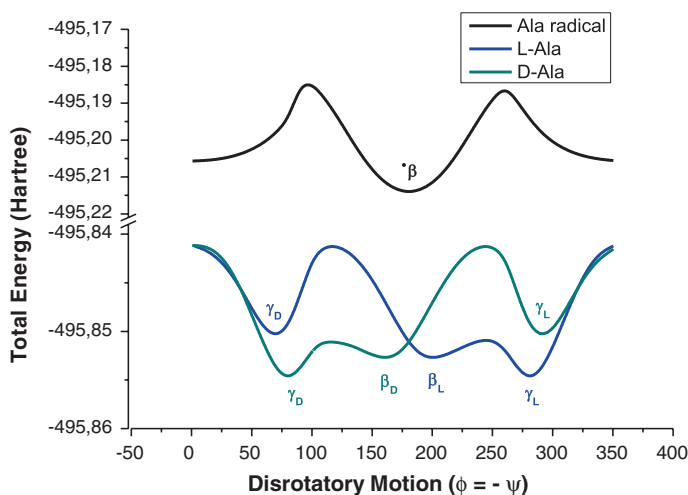
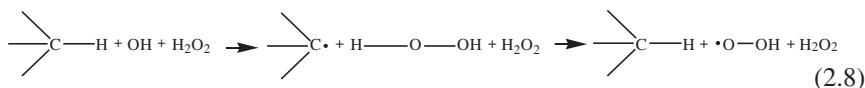


Fig. 2.31 Potential energy curves along the disrotatory ($\phi = -\psi$) cross section of the 2D Ramachandran potential energy surface, $E = f(\phi, \psi)$, for N- and C-protected Ala enantiomeric pairs (in color) and their “common” achiral α -radical (in black)

the H-abstraction followed by H-recapture and Fig. 2.29 illustrates the details of this process.



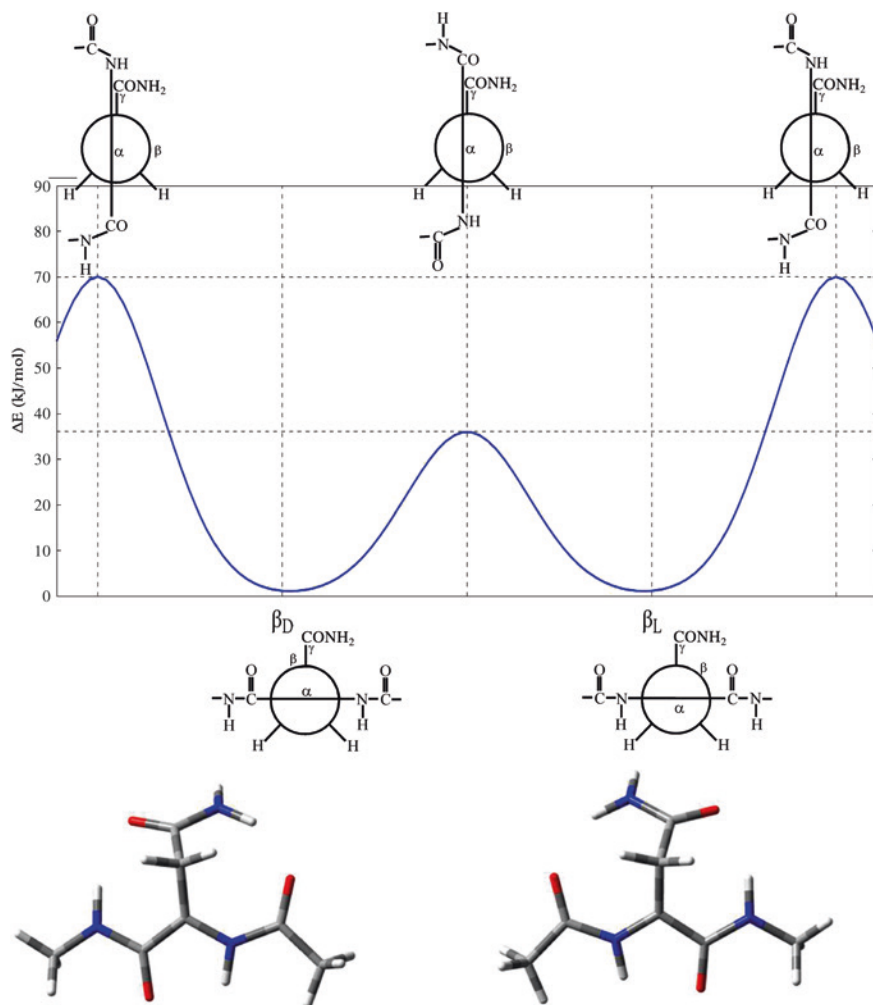


Fig. 2.32 Cross section of the full PES of Ac-Asn•-NHCH₃ and the schematic stereoisomers of Asn β_D and •Asn β_L , detailed path is show in Fig. 2.34

Unlike Gly, all the other amino acid residues are chiral and thus, the possibility of racemization when H. is recaptured is significant even for folded macromolecules, where the attack of the H. could have spatial preference.

The fact that an α -helix is not very stable as a foldamer is known from the literature. In fact if a helical subset of a protein is cut off and separated from the parent protein, it will almost always unfold in H₂O. The affinity of such a sequence to present nascent helices can be enhanced by adding TFE or other fluorinated cosolvents to water as shown for instance for penetratine (see below) [101, 122].

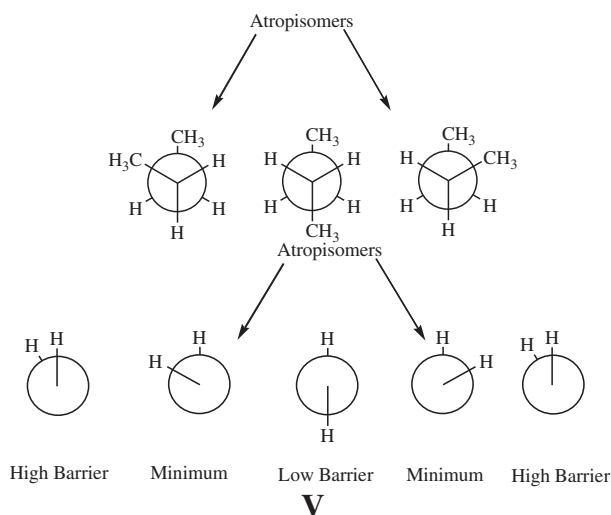
However, as the helical backbone gets more and more structured beside the equilibrium state of a folded and of an unfolded set of molecules the helix can also turn and/or bend (Fig. 2.30). The conformational change of a helix in water as well as in apolar environment, like it is in the case of trans-membrane proteins [109] is illustrated in Fig. 2.30.

Clearly, at the radical center (Fig. 2.30b) the folding of the helix is more pronounced in water than in a less polar environment (like in a membrane).

2.6.2 Atropisomerism of Radicals

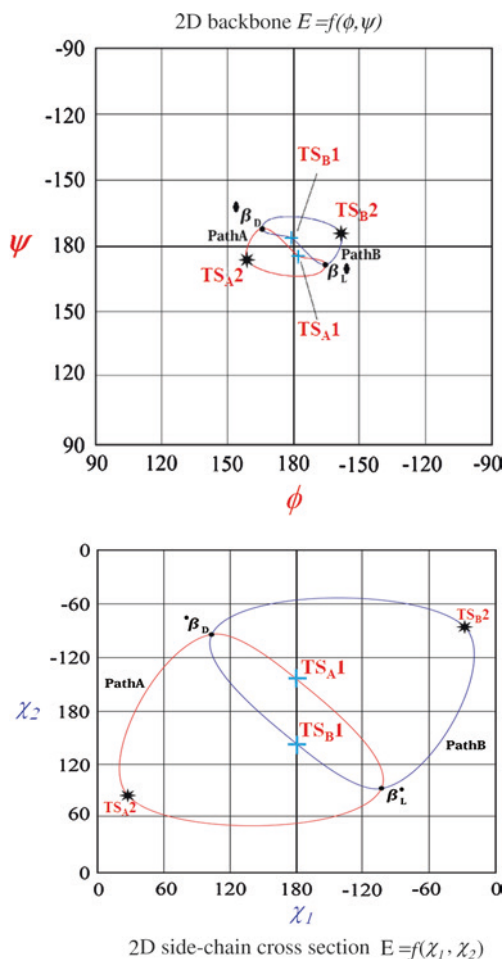
The potential energy surface (PES) associated with Ac-Ala-NHMe clearly shows that regardless of whether the C_α -radical was generated from the L- or D-alanine, it has a single, fully extended β -like conformation, with ($\phi = \psi = 180^\circ$). (Center of the black line in Fig. 2.31.)

However, if the side-chain is larger than that of a CH_3 as it is in Ala, then the PES gets more complex. For Asn which has a longer side-chain, the C_α -radical itself will present a pair of degenerate β -conformers, labeled as β_L and β_D as shown in Fig. 2.32. This type of potential energy curve is typical of axis chirality which leads to atropisomerisation. Atropisomerism occurs even in simple compounds, the rotation about C_2-C_3 of n-butane is a typical example; the g^+ and g^- conformers are atropisomers.



One of the simplest examples for axis chirality is hydrogen peroxide.

Fig. 2.33 Schematic illustration of enantiomeric topomerization paths A and B on the 2D-PES cross sections associated with backbone (*top*) and side-chain (*bottom*) conformational change



The above 1D representation of Asn radical (Fig. 2.32) is an over simplification of the molecular topomerization phenomenon [3] shown in Fig. 2.33. However, it is clear that the system may proceed along two directions. One of the two directions is along the path toward the low energy maximum point (TS1), and the other one toward the high energy maximum point (TS2). Clearly, when we are moving in a 2D PES we have two orthogonal paths that may be labeled as path A and path B. This situation will lead us to four transition states (TS_{A1} , TS_{A2} , TS_{B1} and TS_{B2}). Figure 2.34 illustrates this phenomenon in terms of a pair of 2D cross-sections. Noting that Asn conformations are describable by a 4D-Ramachandran Potential Energy Surface, such a 4D problem may be represented by a pair of 2D energy

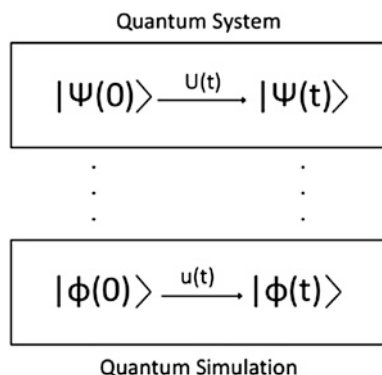


Fig. 2.34 A schematic illustration of the relationship between the quantum system to be investigated and the quantum simulation actually studied. Note that instead of the exact initial state $\psi(0)$ actually the approximate $\phi(0)$ is used as initial state

Table 2.5 Free energy changes, ΔG^0 (kJ/mol) of radical formation and enantiomeric topomerization computed at several levels of theory

| Species | DFT ^a | | | | MP2 ^b | | | |
|-----------------|------------------|-------------------|----------------|---------------------|------------------|-------------------|----------------|---------------------|
| | ΔE_1 | Corr ^d | ΔG_1^0 | E(CC ^c) | ΔE_2 | Corr ^d | ΔG_2^0 | E(CC ^c) |
| TS A1 | 38.47 | −5.9 | 32.57 | 43.29 | 42.71 | −7.34 | 35.37 | 42.29 |
| TS B1 | 38.47 | −5.88 | 32.58 | 43.28 | 42.71 | −7.34 | 35.38 | 42.29 |
| TS A2 | 70.88 | −5.46 | 65.42 | 67.5 | 72.8 | −5.11 | 67.69 | 63.72 |
| TS B2 | 70.87 | −5.46 | 65.41 | 67.49 | 72.8 | −5.11 | 67.69 | 63.72 |
| Asn β_L^e | −1676.3 | 31.03 | −1645.2 | −1604.9 | −1694.6 | 32.22 | −1662.4 | −1602.4 |
| Asn β_D^e | −1676.3 | 31.03 | −1645.2 | −1604.9 | −1694.6 | 32.22 | −1662.4 | −1602.4 |
| •Asn β_L | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| •Asn β_D | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

^aGeometry optimized at B3LYP/6-311 ++G(d,p) level of theory (calculated vibrational wave-numbers are scaled by 0.97)

^bValues are obtained by MP2/cc-pVTZ//B3LYP/6-311 ++G(d,p) level of theory

^cValues are obtained by CCSD(T)/cc-pVTZ//B3LYP/6-311 ++G(d,p) level of theory

^dCorrection for ΔE to ΔG^0

^eFor dissociation free energy the ΔG^0 must be reduced by the hydrogen energy which may takes as 0.5 hartree (1312.75 kJ/mol)

surfaces, one in the backbone subspace, $E = f(\phi, \psi)$, and the other one in the side-chain subspace, $E = f(\chi_1, \chi_2)$ as shown in Fig. 2.33 (Table 2.5).

Clearly, the enantiotopic TS A1 and TS B1 represent the lower barrier (38.47 kJ/mol) while the enantiotopic TS A2 and TS B2 correspond to the higher barrier (70.87 kJ/mol).

2.7 Future Perspectives

According to a classical and old cliché “the future is no longer what it used to be”. This poetic statement implies that we cannot predict the future by direct extrapolation from the events of the past. Nevertheless the final section of this chapter was constructed to envisage certain visions to highlight certain principles, knowing that they might not be true and thus may not turn into reality.

2.7.1 Time Dependent Quantum Simulations

To become an exact science for chemistry is an eschatological hope. In order to avoid any misunderstanding, we first need to give a careful definition of the term “exact”. We mean by it not only something which is accurately computable, but rather a rigorous theory behind what is computed. The rigorous theory is quantum mechanics in principle and the equation to be solved is the time-dependent form of the Schrödinger equation:

$$i\hbar \frac{d}{dt}[\Psi(t)] = \hat{H}[\Psi(t)] \quad (2.9)$$

For the sake of simplicity, \hat{H} is the time independent Hamiltonian. Time dependent Schrödinger equations are needed to describe processes, and time will occur in the form of an evolutionary operator U shown below in Eq. (2.11). The time independent Schrödinger equation (at $t = 0$)

$$\hat{H}[\Psi(0)] = E[\Psi(0)] \quad (2.10)$$

where $\Psi(0)$ is the initial state. The solution of the time dependent state is

$$\Psi(t) = \hat{U}\Psi(0) = e^{-i\hat{H}t}\Psi(0) \quad (2.11)$$

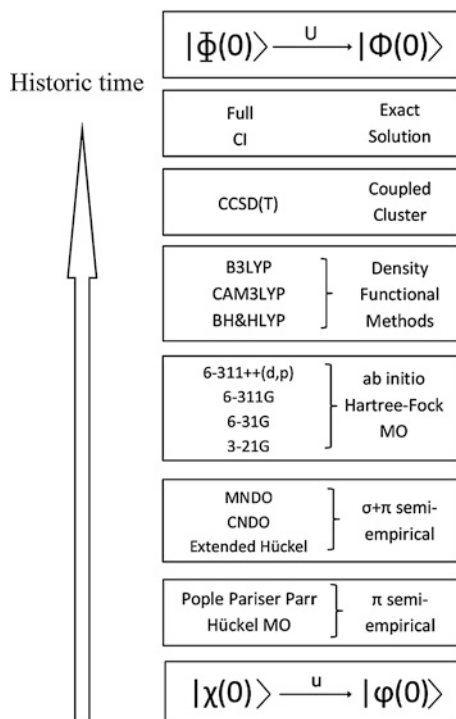
where the unitary transformation is achieved by the following operator

$$\hat{U} = \exp(-i\hat{H}t) \quad (2.12)$$

With the advent of quantum computers it was hoped that such time dependent solution will be possible. Since all processes, including chemical changes are time dependent, the solution of the time dependent Schrödinger equation is the ultimate goal [123–129].

It turned out that most of the efforts today in chemistry and particularly in protein folding studies is to overcome the computational struggle on the time dependent aspect of the latter problem, by brushing aside the determination of the initial state $\Psi(0)$. Furthermore, since the problem is not solvable rigorously, at this time, therefore a simple problem is solved instead and that is actually regarded as a simulation of the actual process. This simulation process is illustrated by Fig. 2.34.

Fig. 2.35 A schematic illustration for the computation of the initial state $\Phi(0)$ from the one electron functions AO, $\chi(0)$, and MO, $\varphi(0)$. In between the *top* and the *bottom*, the various methods developed during the 20th century are shown



2.7.2 Time Independent Quantum Simulations

In order to be ever successful in the time dependent quantum simulation we have to make a serious effort to generate a reliable initial state wave function: $\psi(0)$. Since the time of Schrödinger's publication in 1927, which is getting closer and closer to a century, we made a considerable scientific effort to generate an accurate wave function for the initial state, $\psi(0)$, of a stable molecular system which undergoes the time dependent process. Figure 2.35 shows the hierarchy of the various attempts during the 20th century to state wave functions $\Phi(0)$ from orbitals (χ or φ).

We are today in the second box from the top: The transition to the top box is occurring in the 21st century.

2.7.3 The Protein Folding Problem

This is an old problem. It is a fact that the molecular conformation problem coupled with intramolecular interactions such as hydrogen bondings.

The ultimate rationale behind all purposeful structures and behavior of living things is embodied in the sequence of residues of nascent polypeptide chains – the precursors of

the folded proteins which in biology play the role of Maxwell's demons. In a very real sense it is at this level of organization that the secret of life (if there is one) is to be found. If we could only determine these sequences but also pronounce the law by which they fold, then the secret of life would be found – the ultimate rationale discovered!

(Jaques Monod (1970))

The social impact of the discovery how proteins fold exceeds the combined social impact of

the discovery of fire
the discovery of writing
the discovery of wheel

(Unknown authors)

For a while, it was assumed that the biologically active conformer or folded structure is thermodynamically the most stable one. Protein misfolding or even denaturation studies suggest that the biologically active structure is thermodynamically not the most stable one. At that time, the existence of chaperones were postulated in which case the protein-chaperone complex would find the appropriate folded structure. Since the folding itself, like all chemical process, is time dependent therefore the process of folding could be studied by Quantum Simulation (perhaps on a quantum computer), if the initial conditions are defined [130–143]. Figure 2.36 illustrates a four-amino acid peptide either without or with chaperone assistance. This figure also shows how such computations are treating the initial and final quantum states for such a tetrapeptide.

Thus the Quantum Simulation that is jumping from the initial state to the final state is currently reducing protein folding to a lattice folding problem.

However, even if we ignore all complexity arising from the method applied we have to face the problem rising from molecular structure flexibility. For example if we start with a tetra-glycine to make the problem easy and simple (e.g. Ac-Gly₄-NHMe), the folding has to be represented on a Ramachandran conformational hypersurface of eight independent variables (if counting only the most important backbone torsional angles):

$$E = f(\phi_1, \psi_1, \phi_2, \psi_2, \phi_3, \psi_3, \phi_4, \psi_4) \quad (2.13)$$

Fig. 2.36 Lattice folding simulation of a tetrapeptide folding, **a** alone **b** with chaperone assistance

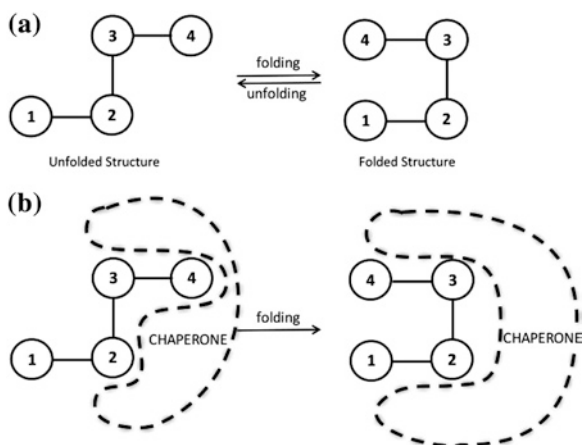


Table 2.6 N_0 (number of minima), N_1 (number of transitional states or TS) and N grid points (in case of 30° grid spacing) to be calculated for a 2D and a 4D amino acid PEHS

| n | 2nD-PEHS | | | 4nD-PEHS | | |
|----|------------------------------------|-----------------------------------|-----------------------|-----------------------------------|----------------------------------|-----------------------|
| | N_0 (min) backbone or side chain | N_1 (TS) backbone or side chain | N grid points | N_0 (min) backbone + side chain | N_1 (TS) backbone + side chain | N grid points |
| 1 | 9 | 18 | 144 | 81 | 324 | 20,736 |
| 2 | 81 | 324 | 20,736 | 6,561 | 52,488 | 4.30×10^8 |
| 4 | 6,561 | 52,488 | 4.30×10^8 | 4.30×10^7 | 6.89×10^8 | 1.85×10^{17} |
| 10 | 3.49×10^9 | 6.97×10^{10} | 3.83×10^{21} | 1.21×10^{19} | 4.86×10^{20} | 1.47×10^{43} |
| 20 | 1.21×10^{19} | 4.86×10^{20} | 1.47×10^{43} | 1.47×10^{38} | 1.18×10^{40} | 2.16×10^{86} |

Therefore such a potential energy hypersurface (PEHS) is expected to have up to the following number of minima (N_0) and the following number of transition states (N_1).

$$N_0 = 3^8 = 6561 \quad (2.14)$$

$$N_1 = 8 \cdot 3^8 = 52488 \quad (2.15)$$

So the initial and final states need to be selected from something like 6,516 stable conformers and the paths interconnecting these minima must pass through a number of transition states (TS) to be picked from the 52,488 TS structures. Keeping track of such simple folding events seems already complicated enough, not talking about how to visualize such a phenomena.

To view the situation from a more general point of view we may consider the followings: (i) Backbone of an amino acid residue has two rotors (ϕ, ψ) and (ii) on average (see Table 2.6) side chain also has two rotors (χ_1, χ_2). Thus, both backbone (BB) 2D- and side chain (SC)—on average 2D—makes for each amino acid residues at least a 4D-problem.

For a 2D amino acid problem, where n is the number of amino acid residues and $2n$ is the number of rotors we have the following relationships.

$$N_0 = 3^{2n} \quad (2.16)$$

$$N_1 = 2n \cdot N_0 = 2n \cdot 3^{2n} \quad (2.17)$$

The reason we expect to have $2n$ times N_0 transitional states, N_1 , is illustrated on Fig. 2.37, where it can be seen that each minimum, N_0 , has two nearest transitional states along each variable.

One possible way to estimate all N_0 initial minimal energy structures is to code algorithmically our knowledge of conformational analysis. These predicted foldamer structures can then be optimized. However, the coded algorithm will be quite complicated, meaning that in certain cases, perhaps several times we may need to attempt a given optimization to find the rotamer or foldamer. Then the algorithm can be modified to look for transitional states. Another alternative would be to generate grid points and subject the fitted functions to the process of mathematical analysis.

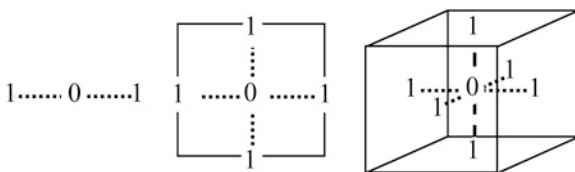


Fig. 2.37 Location of the transitional states (1) around a minima (0), in a one, two and three dimensional space. Critical points labelled by their index (λ) meaning the number of imaginary frequencies. Minima have 0, first order TS have 1 etc

If we generate a grid with 30° intervals i.e. 12 points in the range $0-360^\circ$ then we need N grid points

$$N = 12^{2n} \quad (2.18)$$

for a 4D problem where n is again the numbers of the amino acid residues we have

$$N_0 = 3^{4n} \quad (2.19)$$

$$N_1 = 4n \cdot N_0 = 4n \cdot 3^{4n} \quad (2.20)$$

$$N = 12^{4n} \quad (2.21)$$

Figure 2.38 shows the variation of N_0 and N as a function of n on a logarithmic scale.

At this time of human history, we can handle 20,736 grid points, what is a 4D representation, either a single amino acid diamide with a 2D side chain, like Asn, or diglycine–diamide or dialanine–diamide backbone conformations. However, for a peptide of 20 amino acids, a 10^{82} times increase in the computing capabilities would be demanded.

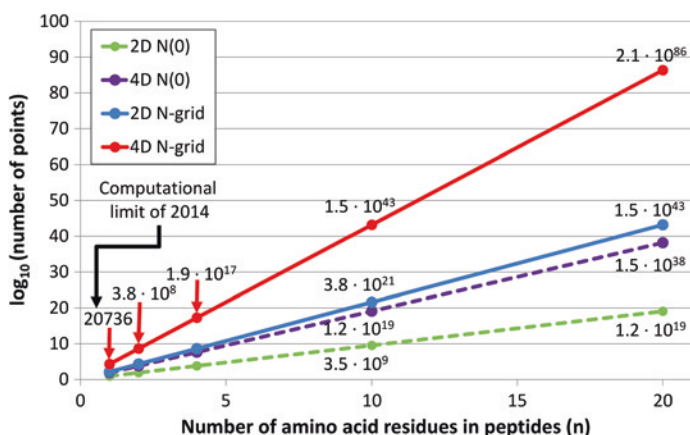


Fig. 2.38 Logarithmic scale of N_0 (number of minima) for a 2D and a 4D case (broken lines) and N (number of grid points) for a 2D and a 4D case (solid lines). Data is summarized in Table 2.6

Further development represents the combination of QM and MM methods in which the chemically more important regions, like active sites of a proteins are computed quantum mechanically, and the less important regions are represented by molecular mechanics. This QM/MM method is a ‘hybrid offspring’, like a mule is, of the QM and MM methods, yet remarkable achievements can be made by this more recent development.

Nevertheless, force field potential energy hyper-surfaces are simulations of the ab initio PEHS. However, the time has arrived to aim for such a gigantic effort that could produce the ab initio surfaces. It has taken 50 years for about 20 researchers to develop Gaussian and Schrödinger from the initial POLYATOM software. That is a 1,000 man–years. However, it could take considerably less to develop a rigorous mathematical equation that encompasses all possible structures of a given protein. Certainly this is doable in the 21st Century if we are really dedicated to achieve that.

Acknowledgments This work was supported by grants from the Hungarian Scientific Research Fund (grant numbers: OTKA NK101072 and TÁMOP-4.2.2.A-11/1/KONV-2012-0047. TÁMOP-4.2.1.B-09/1/KMR). This work was also supported by the following NDA grants: TÁMOP-4.2.2.C-11/1/KONV-2012-0010, TÁMOP-4.2.2.A-11/1/KONV-2012-0047. The authors thank M. Labádi and L. Müller for the administration of the computer clusters used for this work. Technical assistances of Anita Rágyanszki, Klára Gerlei, Kyungseop Lee and Csaba Hatvani are greatly appreciated.

References

1. Csizmadia IG, Harrison MC, Sutcliffe BT (1963) Non-empirical LCAO-SCF calculation of formyl fluoride with gaussian atomic orbitals. *Q Prog Rep* 50:1
2. Csizmadia IG, Harrison MC, Moskowitz JW, Sutcliffe BT (1966) Non-empirical LCAO-MO-SCF-CI calculations on organic molecules with gaussian type functions. Part I *Theor Chim Acta* 6:191
3. Gerlei KZ, Jákli I, Szőri M et al (2013) Atropisomerism of the Asn α radicals revealed by Ramachandran surface topology. *J Phys Chem B* 117:12402–12409. doi:[10.1021/jp4070906](https://doi.org/10.1021/jp4070906)
4. Robb MA, Csizmadia IG (1968) Non-empirical LCAO-MO-SCF-CI calculations on organic molecules with gaussian type functions. *Theor Chim Acta* 10:269
5. Mucsi Z, Tsai A, Szori M et al (2007) A quantitative scale for the extent of conjugation of the amide bond. Amidity percentage as a chemical driving force. *J Phys Chem A* 111:13245–13254. doi:[10.1021/jp0759325](https://doi.org/10.1021/jp0759325)
6. Weintraub SJ, Manson SR (2004) Asparagine deamidation: a regulatory hourglass. *Mech Ageing Dev* 125:255–257. doi:[10.1016/j.mad.2004.03.002](https://doi.org/10.1016/j.mad.2004.03.002)
7. Robinson NE, Robinson AB (2001) Deamidation of human proteins. *Proc Natl Acad Sci USA* 98:12409–12413. doi:[10.1073/pnas.221463198](https://doi.org/10.1073/pnas.221463198)
8. Ritz-Timme S, Laumeier I, Collins M (2003) Age estimation based on aspartic acid racemization in elastin from the yellow ligaments. *Int J Legal Med* 117:96–101. doi:[10.1007/s00414-002-0355-2](https://doi.org/10.1007/s00414-002-0355-2)
9. Ohtani S, Yamamoto T (2010) Age estimation by amino acid racemization in human teeth. *J Forensic Sci* 55:1630–1633. doi:[10.1111/j.1556-4029.2010.01472.x](https://doi.org/10.1111/j.1556-4029.2010.01472.x)
10. Thorpe CT, Streeter I, Pinchbeck GL et al (2010) Aspartic acid racemization and collagen degradation markers reveal an accumulation of damage in tendon collagen that is enhanced with aging. *J Biol Chem* 285:15674–15681. doi:[10.1074/jbc.M109.077503](https://doi.org/10.1074/jbc.M109.077503)
11. Peters B, Trout BL (2006) Asparagine deamidation: pH-dependent mechanism from density functional theory. *Biochemistry* 45:5384–5392. doi:[10.1021/bi052438n](https://doi.org/10.1021/bi052438n)

12. Catak S, Monard G, Aviyente V, Ruiz-López MF (2008) Computational study on nonenzymatic peptide bond cleavage at asparagine and aspartic acid. *J Phys Chem A* 112:8752–8761. doi:[10.1021/jp8015497](https://doi.org/10.1021/jp8015497)
13. Fujii N, Tajima S, Tanaka N et al (2002) The presence of D-beta-aspartic acid-containing peptides in elastic fibers of sun-damaged skin: a potent marker for ultraviolet-induced skin aging. *Biochem Biophys Res Commun* 294:1047–1051. doi:[10.1016/S0006-291X\(02\)00597-1](https://doi.org/10.1016/S0006-291X(02)00597-1)
14. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95–99. doi:[10.1016/S0022-2836\(63\)80023-6](https://doi.org/10.1016/S0022-2836(63)80023-6)
15. Perczel A, Angyan JG, Kajtar M et al (1991) Peptide models. 1. Topology of selected peptide conformational potential energy surfaces (glycine and alanine derivatives). *J Am Chem Soc* 113:6256–6265. doi:[10.1021/ja00016a049](https://doi.org/10.1021/ja00016a049)
16. Csizmadia IG (1997) Basic principles for introductory organic chemistry. Quirk Press, Toronto, pp 135–141
17. Perczel A, Kajtar M, Marcoccia JF, Csizmadia IG (1991) The utility of the four-dimensional Ramachandran map for the description of peptide conformations. *J Mol Struct THEOCHEM* 232:291–319. doi:[http://dx.doi.org/10.1016/0166-1280\(91\)85261-5](http://dx.doi.org/10.1016/0166-1280(91)85261-5)
18. Berman HM, Westbrook J, Feng Z et al (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
19. Radom L, Hehre JW, Pople JA (1972) Molecular orbital theory of the electronic structure of organic compounds. XIII. Fourier component analysis of internal rotation potential functions in saturated molecules. *J Am Chem Soc* 94:2371–2381. doi:[10.1021/ja00762a030](https://doi.org/10.1021/ja00762a030)
20. Radom L, Lathan WA, Hehre WJ, Pople JA (1973) Molecular orbital theory of the electronic structure of organic compounds. XVII. Internal rotation in 1,2-disubstituted ethanes. *J Am Chem Soc* 95:693–698. doi:[10.1021/ja00784a008](https://doi.org/10.1021/ja00784a008)
21. Kehoe TAK, Peterson MR, Chass GA, et al (2003) The fitting and functional analysis of a double rotor potential energy surface for the R and S enantiomers of 1-chloro-3-fluoro-isobutane. *J Mol Struct THEOCHEM* 666–667:79–87. doi:<http://dx.doi.org/10.1016/j.theochem.2003.08.015>
22. Demaré GR, Peterson MR, Csizmadia IG, Strausz OP (1980) Conformational energy surfaces of triplet-state isomeric methyloxiranes. *J Comp Chem* 1:141–148. doi:[10.1002/jcc.540010206](https://doi.org/10.1002/jcc.540010206)
23. Peterson R, Mare RDE, Roosevelt AFD (1981) Conformations formaldehyde, acetone of triplet carbonyl compounds: and the reaction of O (3P) atoms with olefins proceeds via a C % Ci, triplet biradical [I], which can undergo intersystem crossing to the So state and be stabilized as an epoxide, 86:131–147
24. Rágyanszki A, Surányi A, Csizmadia IG, et al (2014) Fourier type potential energy function for conformational change of selected organic functional groups. *Chem Phys Lett* 599:169–174. doi:<http://dx.doi.org/10.1016/j.cplett.2014.03.029>
25. Pohl G, Perczel A, Vass E et al (2007) A matrix isolation study on Ac-Gly-NHMe and Ac-L-Ala-NHMe, the simplest chiral and achiral building blocks of peptides and proteins. *Phys Chem Chem Phys* 9:4698–4708
26. Pohl G, Perczel A, Vass E et al (2008) A matrix isolation study on Ac-L-Pro-NH₂: a frequent structural element of beta- and gamma-turns of peptides and proteins. *Tetrahedron* 64(9):2126–2133
27. Jákli I, Knak Jensen SJ, Csizmadia IG, Perczel A (2012) Variation of conformational properties at a glance. True graphical visualization of the Ramachandran surface topology as a periodic potential energy surface. *Chem Phys Lett* 547:82–88. doi:[10.1016/j.cplett.2012.08.002](https://doi.org/10.1016/j.cplett.2012.08.002)
28. Viviani W, Rivail J-L, Csizmadia I (1993) Peptide models II. Intramolecular interactions and stable conformations of glycine, alanine, and valine peptide analogues. *Theor Chim Acta* 85:189–197. doi:[10.1007/BF01374587](https://doi.org/10.1007/BF01374587)
29. McAllister MA, Perczel A, Császár P et al (1993) Peptide models 4. Topological features of molecular mechanics and ab initio 2D-Ramachandran maps. *J Mol Struct THEOCHEM* 288:161–179. doi:[10.1016/0166-1280\(93\)87048-I](https://doi.org/10.1016/0166-1280(93)87048-I)

30. Endrédi G, Perczel A, Farkas O, et al (1997) Peptide models XV. The effect of basis set size increase and electron correlation on selected minima of the ab initio 2D-Ramachandran map of For-Gly-NH₂ and For-L-Ala-NH₂. *J Mol Struct THEOCHEM* 391:15–26. doi:[http://dx.doi.org/10.1016/S0166-1280\(96\)04695-7](http://dx.doi.org/10.1016/S0166-1280(96)04695-7)
31. Viviani W, Rivail JL, Perczel A, Csizmadia IG (1993) Peptide models. 3. Conformational potential energy hypersurface of formyl-L-valinamide. *J Am Chem Soc* 115:8321–8329. doi:[10.1021/ja00071a046](http://dx.doi.org/10.1021/ja00071a046)
32. Perczel A, Farkas O, Csizmadia IG (1995) Peptide models. 17. The role of the water molecule in peptide folding. An ab initio study on the right-handed helical conformations of N-formylglycinamide and N-formyl-L-alaninamide monohydrates [H(CONH-CHR-CONH)H·H₂O; R=H or CH₃]. *J Am Chem Soc* 117:1653–1654. doi:[10.1021/ja00110a028](http://dx.doi.org/10.1021/ja00110a028)
33. Perczel A, Daudel R, Ángyán JG, Csizmadia IG (1990) A study on the backbone/side-chain interaction in N-formyl-(L)serineamide. *Can J Chem* 68:1882–1888. doi:[10.1139/v90-291](http://dx.doi.org/10.1139/v90-291)
34. Farkas Ö, Perczel A, Marcoccia J-F, et al (1995) Peptide models XIII. Side-chain conformational energy surface $E=E(\chi_1, \chi_2)$ of N-formyl-L-serinamide (For-L-Ser-NH₂) in its γ L or C7 eq backbone conformation. *J Mol Struct THEOCHEM* 331:27–36. doi:[http://dx.doi.org/10.1016/0166-1280\(94\)03929-F](http://dx.doi.org/10.1016/0166-1280(94)03929-F)
35. Perczel A, Farkas Ö, Csizmadia IG (1996) Peptide models XVI. The identification of selected HCO—L—SER—NH₂ conformers via a systematic grid search using ab initio potential energy surfaces. *J Comput Chem* 17:821–834. doi:[10.1002/\(SICI\)1096-987X\(199605\)17:7<821::AID-JCC6>3.0.CO;2-U](http://dx.doi.org/10.1002/(SICI)1096-987X(199605)17:7<821::AID-JCC6>3.0.CO;2-U)
36. Perczel A, Farkas Ö, Csizmadia IG (1996) Peptide models. 18. hydroxymethyl side-chain induced backbone conformational shifts of L-serine amide. All ab Initio conformers of for-L-Ser-NH₂. *J Am Chem Soc* 118:7809–7817. doi:[10.1021/ja960464q](http://dx.doi.org/10.1021/ja960464q)
37. Perczel A, Farkas Ö, Jákli I, Csizmadia IG (1998) Peptide models XXI. Side-chain/backbone conformational interconversions in HCO-L-Ser-NH₂. Tracing relaxation paths by ab initio modeling. An exploratory study. *J Mol Struct THEOCHEM* 455:315–338
38. Imre Jákli AP (2000) Peptide models XXIII. Conformational model for polar side-chain containing amino acid residues: a comprehensive analysis of RHF, DFT, and MP2 properties of HCO-L-SER-NH₂. *J Comput Chem* 21:626–655. doi:[10.1002/\(SICI\)1096-987X\(200006\)21:83.0.CO;2-P](http://dx.doi.org/10.1002/(SICI)1096-987X(200006)21:83.0.CO;2-P)
39. Fang D-C, Fu X-Y, Tang T-H, Csizmadia IG (1998) Ab initio modelling of peptide biosynthesis. *J Mol Struct THEOCHEM* 427:243–252. doi:[http://dx.doi.org/10.1016/S0166-1280\(97\)00257-1](http://dx.doi.org/10.1016/S0166-1280(97)00257-1)
40. Galant NJ, Lee DR, Fiser B, et al (2012) Disulfidicity: a scale to characterize the disulfide bond strength via the hydrogenation thermodynamics. *Chem Phys Lett* 539–540:11–14. doi:<http://dx.doi.org/10.1016/j.cplett.2012.05.017>
41. Galant NJ, Song HC, Jákli I et al (2014) A theoretical study of the stability of disulfide bridges in various β -sheet structures of protein segment models. *Chem Phys Lett* 593:48–54
42. Vank JC, Sosa CP, Perczel A, Csizmadia IG (2000) Peptide models XXVII. An exploratory ab initio study on the 21st amino acid side-chain conformations of N-formyl-L-selenocysteinamide (For-L-Sec-NH₂) and N-acetyl-L-selenocysteine-N-methylamide (Ac-L-Sec-NHMe) in their γ L backbone conformation. *Can J Chem* 78:395–408. doi:[10.1139/v00-029](http://dx.doi.org/10.1139/v00-029)
43. Fiser B, Mucsi Z, Viskolcz B et al (2013) Controlled antioxidative steps of the cell. The concept of chalcogenicity. *Chem Phys Lett* 590:83–86. doi:[10.1016/j.cplett.2013.10.033](http://dx.doi.org/10.1016/j.cplett.2013.10.033)
44. Láng A, György K, Csizmadia IG, Perczel A (2003) A conformational comparison of N- and C-protected methionine and N- and C-protected homocysteine. *J Mol Struct THEOCHEM* 666–667:219–241. doi:<http://dx.doi.org/10.1016/j.theochem.2003.08.029>
45. Láng A, Csizmadia IG, Perczel A (2005) Peptide models XLV: conformational properties of N-formyl-L-methioninamide and its relevance to methionine in proteins. *Proteins* 58:571–588. doi:[10.1002/prot.20307](http://dx.doi.org/10.1002/prot.20307)
46. Sheraly AR, Chass GA, Csizmadia IG (2003) The multidimensional conformational analysis for the backbone across the disrotatory axis at selected side-chain conformers of

- N-Ac-homocysteine-NHMe—an ab initio exploratory study. *J Mol Struct THEOCHEM* 666–667:243–249. doi:<http://dx.doi.org/10.1016/j.theochem.2003.08.030>
47. Sahai MA, Motiwala SS, Chass GA et al (2003) An ab initio exploratory study of the full conformational space of MeCO-L-threonine-NH-Me. *J Mol Struct THEOCHEM* 666–667:251–267. doi:[10.1016/j.theochem.2003.08.031](http://dx.doi.org/10.1016/j.theochem.2003.08.031)
48. Sahai MA, Fejer SN, Viskolcz B et al (2006) First-principle computational study on the full conformational space of L-threonine diamide, the energetic stability of cis and trans isomers. *J Phys Chem A* 110:11527–11536. doi:[10.1021/jp0680488](http://dx.doi.org/10.1021/jp0680488)
49. Lam JSW, Koo JCP, Hudaky I, et al. Predicting the conformational preferences of N-acetyl-4-hydroxy-L-proline-N'-methylamide from the proline residue. *J Mol Struct Theochem* 666–667:285–289
50. Rassolian M, Chass GA, Setiadi DH, Csizmadia IG (2003) Asparagine—ab initio structural analyses. *J Mol Struct THEOCHEM* 666–667:273–278. doi:<http://dx.doi.org/10.1016/j.theochem.2003.08.032>
51. Berg M, Salpietro SJ, Perczel A et al (2000) Peptide models XXVI Side chain conformational analysis of N-formyl-L-asparagin amide and N-acetyl-L-asparagin N-methylamide in their gL backbone conformation. *J Mol Struct* 504:127–140
52. Masman MF, Zamora MA, Rodriguez AM et al (2002) Exploration of the full conformational space of N-acetyl-L-glutamate-N-methylamide. *Eur Phys J D-At Mol Opt Plasma Phys* 20:531–542. doi:[10.1140/epjd/e2002-00150-y](http://dx.doi.org/10.1140/epjd/e2002-00150-y)
53. Salpietro SJ, Perczel A, Farkas Ö et al (2000) Peptide models XXV. Side-chain conformational potential energy surface, of N-formyl-L-aspartic acidamide and its conjugate base N-formyl-L-aspartatamide in their γ L backbone conformations. *J Mol Struct THEOCHEM* 497:39–63. doi:[10.1016/S0166-1280\(99\)00196-7](http://dx.doi.org/10.1016/S0166-1280(99)00196-7)
54. Koo JCP, Chass GA, Perczel A et al (2002) Exploration of the four-dimensional-conformational potential energy hypersurface of N-Acetyl-L-aspartic Acid N'-Methylamide with its internally hydrogen bonded side-chain orientation. *J Phys Chem A* 106:6999–7009. doi:[10.1021/jp014514b](http://dx.doi.org/10.1021/jp014514b)
55. Koo JCP, Chass GA, Perczel A et al (2002) N-acetyl-L-aspartic acid-N'-methylamide with side-chain orientation capable of external hydrogen bonding. *Eur Phys J D-At Mol Opt Plasma Phys* 20:499–511. doi:[10.1140/epjd/e2002-00148-5](http://dx.doi.org/10.1140/epjd/e2002-00148-5)
56. Koo JCP, Lam JSW, Chass GA, et al. (2003) Ramachandran backbone potential energy surfaces of aspartic acid and aspartate residues: implications on allosteric sites in receptor–ligand complexations. *J Mol Struct THEOCHEM* 666–667:279–284. doi:<http://dx.doi.org/10.1016/j.theochem.2003.08.055>
57. Farkas Ö, McAllister MA, Ma JH et al (1996) Peptide models XIX: side-chain conformational energy surface and amide I vibrational frequencies of N-formyl-L-phenylalaninamide (For-Phe-NH₂) in its γ L or γ inv or C7 eq backbone conformation. *J Mol Struct THEOCHEM* 369:105–114. doi:[10.1016/S0166-1280\(96\)04548-4](http://dx.doi.org/10.1016/S0166-1280(96)04548-4)
58. Perczel A, Farkas Ö, Csizmadia IG, Császár AG (1997) Peptide models XX. Aromatic side-chain–backbone interaction in phenylalanine-containing diamide model system. A systematic search for the identification of all the ab initio conformers of N-formyl-L-phenylalanine-amide. *Can J Chem* 75:1120–1130. doi:[10.1139/v97-134](http://dx.doi.org/10.1139/v97-134)
59. Jákli I, Perczel A, Farkas Ö et al (1998) Peptide models XXII. A conformational model for aromatic amino acid residues in proteins. A comprehensive analysis of all the RHF/6–31 + G* conformers of for-L-Phe–NH₂. *J Mol Struct THEOCHEM* 455:303–314
60. Chass GA, Mirasol RS, Setiadi DH et al (2005) Characterization of the conformational probability of N-acetyl-phenylalanyl-NH₂ by RHF, DFT, and MP2 computation and AIM analyses, confirmed by jet-cooled infrared data. *J Phys Chem A* 109:5289–5302. doi:[10.1021/jp040720i](http://dx.doi.org/10.1021/jp040720i)
61. Chass GA, Lovas S, Murphy RF, Csizmadia IG (2002) The role of enhanced aromatic-electron donating aptitude of the tyrosyl sidechain with respect to that of phenylalanyl in intramolecular interactions. *Eur Phys J D* 20:481–497. doi:[10.1140/epjd/e2002-00155-6](http://dx.doi.org/10.1140/epjd/e2002-00155-6)

62. Sahai MA, Kehoe TAK, Koo JCP et al (2005) First principle computational study on the full conformational space of l-proline diamides. *J Phys Chem A* 109:2660–2679. doi:[10.1021/jp040594i](https://doi.org/10.1021/jp040594i)
63. Füüzéry AK, Csizmadia IG (2000) An exploratory density functional study on N- and C-protected trans- α , β -didehydroalanine. *J Mol Struct THEOCHEM* 501–502:539–547. doi:[10.1016/S0166-1280\(99\)00469-8](https://doi.org/10.1016/S0166-1280(99)00469-8)
64. Baldoni HA, Zamarride GN, Enriz RD, et al (2000) Peptide models XXIX. cis–trans Isomerism of peptide bonds: ab initio study on small peptide model compound; the 3D-Ramachandran map of formylglycinamide. *J Mol Struct THEOCHEM* 500:97–111. doi:[http://dx.doi.org/10.1016/S0166-1280\(00\)00372-9](http://dx.doi.org/10.1016/S0166-1280(00)00372-9)
65. Sahai MA, Szöri M, Viskolcz B et al (2007) Transition state infrared spectra for the trans \rightarrow Cis isomerization of a simple peptide model. *J Phys Chem A* 111:8384–8389. doi:[10.1021/jp074991f](https://doi.org/10.1021/jp074991f)
66. Van Alsenoy C, Cao M, Newton SQ et al (1993) Conformational analysis and structural study by ab initio gradient geometry optimizations of the model tripeptide N-formyl L-alanyl L-alanine amide. *J Mol Struct THEOCHEM* 286:149–163. doi:[10.1016/0166-1280\(93\)87160-F](https://doi.org/10.1016/0166-1280(93)87160-F)
67. McAllister MA, Perczel A, Császár P, Csizmadia IG (1993) Peptide models 5. Topological features of molecular mechanics and ab initio 4D-Ramachandran maps. *J Mol Struct THEOCHEM* 288:181–198. doi:[10.1016/0166-1280\(93\)87049-J](https://doi.org/10.1016/0166-1280(93)87049-J)
68. Perczel A, McAllister MA, Császár P, Csizmadia IG (1994) Peptide models. IX. A complete conformational set of For-Ala-Ala-NH₂ from ab initio computations. *Can J Chem* 72:2050–2070. doi:[10.1139/v94-262](https://doi.org/10.1139/v94-262)
69. Perczel A, McAllister MA, Csaszar P, Csizmadia IG (1993) Peptide models 6. New.beta.-turn conformations from ab initio calculations confirmed by x-ray data of proteins. *J Am Chem Soc* 115:4849–4858. doi:[10.1021/ja00064a053](https://doi.org/10.1021/ja00064a053)
70. Perczel A, Jáklí I, McAllister MA, Csizmadia IG (2003) Relative stability of major types of beta-turns as a function of amino acid composition: a study based on Ab initio energetic and natural abundance data. *Chemistry* 9:2551–2566. doi:[10.1002/chem.200204393](https://doi.org/10.1002/chem.200204393)
71. Sahai MA, Setiadi DH, Chass GA et al (2003) A model study of the IgA hinge region: an exploratory study of selected backbone conformations of MeCO-l-Pro-l-Thr-NH-Me. *J Mol Struct THEOCHEM* 666–667:311–319. doi:[10.1016/j.theochem.2003.08.036](https://doi.org/10.1016/j.theochem.2003.08.036)
72. Sahai MA, Viskolcz B, Pai EF, Csizmadia IG (2007) Quantifying the intrinsic effects of two point mutation models of pro-pro-pro triamino acid diamide. A first-principle computational study. *J Phys Chem B* 111:13135–13142. doi:[10.1021/jp074046r](https://doi.org/10.1021/jp074046r)
73. Perczel A, Endrédi G, McAllister MA et al (1995) Peptide models VII the ending of the right-handed helices in oligopeptides [For-(Ala)_n-NH₂ for 2 \leq n \leq 4] and in proteins. *J Mol Struct THEOCHEM* 331:5–10. doi:[10.1016/0166-1280\(94\)03972-N](https://doi.org/10.1016/0166-1280(94)03972-N)
74. Endrédi G, McAllister MA, Farkas Ö et al (1995) Peptide models XII Topological features of molecular mechanics and ab-initio 8D-Ramachandran maps. Conformational data for Ac-(l-Ala)4-NHMe and For-(l-Ala)4-NH₂. *J Mol Struct THEOCHEM* 331:11–26. doi:[10.1016/0166-1280\(94\)03915-8](https://doi.org/10.1016/0166-1280(94)03915-8)
75. Sahai MA, Sahai MR, Chass GA et al (2003) An ab initio exploratory study on selected conformational features of MeCO-l-Ala-l-Ala-l-Ala-NH-Me as a XxxYyyZzz tripeptide motif within a protein structure. *J Mol Struct THEOCHEM* 666–667:327–336. doi:[10.1016/j.theochem.2003.08.041](https://doi.org/10.1016/j.theochem.2003.08.041)
76. Masman MF, Rodríguez AM, Svetaz L et al (2006) Synthesis and conformational analysis of His-Phe-Arg-Trp-NH₂ and analogues with antifungal properties. *Bioorg Med Chem* 14:7604–7614. doi:[10.1016/j.bmc.2006.07.007](https://doi.org/10.1016/j.bmc.2006.07.007)
77. Perczel A, Gáspári Z, Csizmadia IG (2005) Structure and stability of beta-pleated sheets. *J Comput Chem* 26:1155–1168. doi:[10.1002/jcc.20255](https://doi.org/10.1002/jcc.20255)
78. Perczel A, Farkas Ö, Marcoccia JF, Csizmadia IG (1997) Peptide models. XIV. Ab initio study on the role of side-chain backbone interaction stabilizing the building unit of right- and left-handed helices in peptides and proteins. *Int J Quantum Chem* 61:797–814. doi:[10.1002/\(SICI\)1097-461X\(1997\)61:5<797::AID-QUA6>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-461X(1997)61:5<797::AID-QUA6>3.0.CO;2-R)

79. Topol IA, Burt SK, Deretey E et al (2001) α - and 3(10)-helix interconversion: a quantum-chemical study on polyalanine systems in the gas phase and in aqueous solvent. *J Am Chem Soc* 123:6054–6060
80. Viskolcz B, Fejer SN, Knak Jensen SJ et al (2007) Information accumulation in helical oligopeptide structures. *Chem Phys Lett* 450:123–126. doi:[10.1016/j.cplett.2007.11.001](https://doi.org/10.1016/j.cplett.2007.11.001)
81. Jákli I, Csizmadia IG, Fejer SN et al (2013) Helix compactness and stability: Electron structure calculations of conformer dependent thermodynamic functions. *Chem Phys Lett* 563:80–87. doi:[10.1016/j.cplett.2013.01.060](https://doi.org/10.1016/j.cplett.2013.01.060)
82. Galant NJ, Wang H, Lee DR et al (2009) Thermodynamic role of glutathione oxidation by peroxide and peroxybcarbonate in the prevention of Alzheimer's disease and cancer. *J Phys Chem A* 113:9138–9149. doi:[10.1021/jp809116n](https://doi.org/10.1021/jp809116n)
83. Ding VZY, Dawson SSH, Lau LWY et al (2011) A computational study of glutathione and its fragments: N-acetylcisteinylglycine and γ -glutamylmethanamide. *Chem Phys Lett* 507:168–173. doi:[10.1016/j.cplett.2011.03.067](https://doi.org/10.1016/j.cplett.2011.03.067)
84. Fiser B, Szori M, Jójárt B et al (2011) Antioxidant potential of glutathione: a theoretical study. *J Phys Chem B* 115:11269–11277. doi:[10.1021/jp2049525](https://doi.org/10.1021/jp2049525)
85. Fiser B, Jójárt B, Csizmadia IG, Viskolcz B (2013) Glutathione–hydroxyl radical interaction: a theoretical study on radical recognition process. *PLoS ONE* 8:e73652. doi:[10.1371/journal.pone.0073652](https://doi.org/10.1371/journal.pone.0073652)
86. Veeravalli K, Boyd D, Iverson BL et al (2011) Laboratory evolution of glutathione biosynthesis reveals natural compensatory pathways. *Nat Chem Biol* 7:101–105. doi:[10.1038/nchembio.499](https://doi.org/10.1038/nchembio.499)
87. Ganguly D, Srikanth CV, Kumar C et al (2003) Why is glutathione (a tripeptide) synthesized by specific enzymes while TSH releasing hormone (TRH or thyroliberin), also a tripeptide, is produced as part of a prohormone protein? *IUBMB Life* 55:553–554. doi:[10.1080/15216540310001623064](https://doi.org/10.1080/15216540310001623064)
88. Fahey RC, Sundquist AR (1991) Evolution of glutathione metabolism. *Adv Enzymol Relat Areas Mol Biol* 64:1–53
89. Yi H, Ravilious GE, Galant A et al (2010) From sulfur to homogluthathione: thiol metabolism in soybean. *Amino Acids* 39:963–978. doi:[10.1007/s00726-010-0572-9](https://doi.org/10.1007/s00726-010-0572-9)
90. Hermes-Lima M (2004) Oxygen in biology and biochemistry: role of free radicals. *Funct Metab* 319–368
91. Meuwly P, Thibault P, Rauser WE (1993) Gamma-glutamylcysteinylglutamic acid—a new homologue of glutathione in maize seedlings exposed to cadmium. *FEBS Lett* 336:472–476
92. Pálfi V, Perczel A (2008) How stable is a collagen triple helix? An ab initio study on various collagen and beta-sheet forming sequences. *J Comput Chem* 29(9):1374–1386
93. Bella J, Berman HM (1996) Crystallographic evidence for $\text{C}\alpha\text{--H}\cdots\text{O}=\text{C}$ hydrogen bonds in a collagen triple helix. *J Mol Biol* 264:734–742. doi:[10.1006/jmbi.1996.0673](https://doi.org/10.1006/jmbi.1996.0673)
94. Henkelman RM, Stanisz GJ, Kim JK, Bronskill MJ (1994) Anisotropy of NMR properties of tissues. *Magn Reson Med* 32:592–601
95. Pálfi V, Perczel A (2010) Stability of the hydration layer of tropocollagen: a QM study. *J Comput Chem* 31(4):764–777
96. Masman MF, Eisel ULM, Csizmadia IG et al (2009) Model peptides for binding studies on peptide–peptide interactions: the case of amyloid β (1–42) aggregates. *J Phys Chem A* 113:11710–11719
97. Eakin CM, Berman AJ, Miranker AD (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nat Struct Mol Biol* 13:202–208. doi:[10.1038/nsmb1068](https://doi.org/10.1038/nsmb1068)
98. Nelson R, Sawaya MR, Balbirnie M et al (2005) Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435:773–778. doi:[10.1038/nature03680](https://doi.org/10.1038/nature03680)
99. Wright CF, Teichmann SA, Clarke J, Dobson CM (2005) The importance of sequence diversity in the aggregation and evolution of proteins. *Nature* 438:878–881

100. Beke T, Csizmadia IG, Perczel A (2006) Theoretical study on tertiary structural elements of beta-peptides: nanotubes formed from parallel-sheet-derived assemblies of beta-peptides. *J Am Chem Soc* 128:5158–5167. doi:[10.1021/ja0585127](https://doi.org/10.1021/ja0585127)
101. Czajlik A, Perczel A (2004) Peptide models XXXII. Computed chemical shift analysis of penetratin fragments. *J Mol Struct THEOCHEM* 675:129–139. doi:[10.1016/j.theochem.2003.12.036](https://doi.org/10.1016/j.theochem.2003.12.036)
102. Salpietro SJ, Viskolcz B, Csizmadia IG (2003) An exploratory ab initio study on the entropy of various backbone conformers for the HCO-Gly-Gly-Gly-NH₂ tripeptide motif. *J Mol Struct THEOCHEM* 666–667:89–94. doi:[10.1016/j.theochem.2003.08.016](https://doi.org/10.1016/j.theochem.2003.08.016)
103. Viskolcz B, Fejer SN, Csizmadia IG (2006) *J Phys Chem A* 110:3808
104. Fejer SN, Csizmadia IG, Viskolcz B (2006) Thermodynamic functions of conformational changes: conformational network of glycine diamide folding, entropy lowering, and informational accumulation. *J Phys Chem A* 110:13325–13331. doi:[10.1021/jp065595k](https://doi.org/10.1021/jp065595k)
105. Pohl G, Jákl I, Csizmadia IG et al (2012) The role of entropy in initializing the aggregation of peptides: a first principle study on oligopeptide oligomerization. *Phys Chem Chem Phys* 14:1507
106. Hudáky I, Hudáky P, Perczel A (2004) Solvation model induced structural changes in peptides. A quantum chemical study on Ramachandran surfaces and conformers of alanine diamide using the polarizable continuum model. *J Comput Chem* 25:1522–1531. doi:[10.1002/jcc.20073](https://doi.org/10.1002/jcc.20073)
107. Hudáky P, Perczel A (2004) Conformation dependence of pKa: Ab initio and DFT investigation of histidine. *J Phys Chem A* 108:6195–6205. doi:[10.1021/jp048964q](https://doi.org/10.1021/jp048964q)
108. Barnham KJ, Masters CL, Bush AI (2004) Neurodegenerative diseases and oxidative stress. *Nat Rev Drug Discov* 3:205–214. doi:[10.1038/nrd1330](https://doi.org/10.1038/nrd1330)
109. Owen MC, Szori M, Csizmadia IG, Viskolcz B (2012) Conformation-dependent ¹OH/H₂O₂ hydrogen abstraction reaction cycles of Gly and Ala residues: a comparative theoretical study. *J Phys Chem B* 116:1143–1154. doi:[10.1021/jp2089559](https://doi.org/10.1021/jp2089559)
110. Büeler H (2009) Impaired mitochondrial dynamics and function in the pathogenesis of Parkinson's disease. *Exp Neurol* 218:235–246. doi:[10.1016/j.expneurol.2009.03.006](https://doi.org/10.1016/j.expneurol.2009.03.006)
111. Easton CJ (1997) Free-radical reactions in the synthesis of alpha-amino acids and derivatives. *Chem Rev* 97:53–82
112. Stadtman ER (2006) Protein oxidation and aging. *Free Radic Res* 40:1250–1258. doi:[10.1080/10715760600918142](https://doi.org/10.1080/10715760600918142)
113. Berlett BS, Stadtman ER (1997) Protein oxidation in aging, disease, and oxidative stress. *J Biol Chem* 272:20313–20316
114. Gregersen N, Bolund L, Bross P (2005) Protein misfolding, aggregation, and degradation in disease. *Mol Biotechnol* 31:141–150. doi:[10.1385/MB:31:2:141](https://doi.org/10.1385/MB:31:2:141)
115. Harris E (1992) Regulation of antioxidant enzymes. *FASEB J* 6:2675–2683
116. Shapira R, Chou CH (1987) Differential racemization of aspartate and serine in human myelin basic protein. *Biochem Biophys Res Commun* 146:1342–1349
117. Masters PM (1983) Stereochemically altered noncollagenous protein from human dentin. *Calcif Tissue Int* 35:43–47
118. Fujii N (2005) D-amino acid in elderly tissues. *Biol Pharm Bull* 28:1585–1589
119. Shapiro SD, Endicott SK, Province MA et al (1991) Marked longevity of human lung parenchymal elastic fibers deduced from prevalence of D-aspartate and nuclear weapons-related radiocarbon. *J Clin Invest* 87:1828–1834. doi:[10.1172/JCI115204](https://doi.org/10.1172/JCI115204)
120. Gerlei KZ, Élo L, Fiser B et al (2014) Impairment of a model peptide by oxidative stress: thermodynamic stabilities of asparagine diamide C α -radical foldamers. *Chem Phys Lett* 593:104–108
121. Owen MC, Viskolcz B, Csizmadia IG (2011) Quantum chemical analysis of the unfolding of a penta-glycyl 3(10)-helix initiated by HO(•), HO₂(•), and O₂(-•). *J Chem Phys* 135:035101. doi:[10.1063/1.3608168](https://doi.org/10.1063/1.3608168)

122. Czajlik A, Meskó E, Penke B, Perczel A (2002) Investigation of penetratin peptides. Part 1. The environment dependent conformational properties of penetratin and two of its derivatives. *J Pept Sci* 8:151–171. doi:[10.1002/psc.380](https://doi.org/10.1002/psc.380)
123. Smirnov AY, Savel'ev S, Mourokh LG, Nori F (2007) Modelling chemical reactions using semiconductor quantum dots. *Europhys Lett* 80:67008. doi:[10.1209/0295-5075/80/67008](https://doi.org/10.1209/0295-5075/80/67008)
124. Aspuru-Guzik A, Dutoi AD, Love PJ, Head-Gordon M (2005) Simulated quantum computation of molecular energies. *Science* 309:1704–1707. doi:[10.1126/science.1113479](https://doi.org/10.1126/science.1113479)
125. Kassal I, Jordan SP, Love PJ et al (2008) Polynomial-time quantum algorithm for the simulation of chemical dynamics. *Proc Natl Acad Sci USA* 105:18681–18686. doi:[10.1073/pnas.0808245105](https://doi.org/10.1073/pnas.0808245105)
126. Kassal I, Whitfield JD, Perdomo-Ortiz A et al (2011) Simulating chemistry using quantum computers. *Annu Rev Phys Chem* 62:185–207. doi:[10.1146/annurev-physchem-032210-103512](https://doi.org/10.1146/annurev-physchem-032210-103512)
127. Wang H, Ashhab S, Nori F (2009) Efficient quantum algorithm for preparing molecular-system-like states on a quantum computer. *Phys Rev A* 79:042335. doi:[10.1103/PhysRevA.79.042335](https://doi.org/10.1103/PhysRevA.79.042335)
128. Lu D, Xu N, Xu R et al (2011) Simulation of chemical isomerization reaction dynamics on a NMR quantum simulator. *Phys Rev Lett* 107:20501
129. Lidar DA, Wang H (1999) Calculating the thermal rate constant with exponential speedup on a quantum computer. *Phys Rev E* 59:2429–2438
130. Perdomo A, Truncik C, Tubert-Brohman I et al (2008) Construction of model Hamiltonians for adiabatic quantum computation and its application to finding low-energy conformations of lattice protein models. *Phys Rev A* 78:12320
131. Floudas CA, Pardalos PM (2000) Optimization in computational chemistry and molecular biology: local and global approaches. Kluwer Academic Publication, Berlin
132. Alexander K, Hartmann HR (2004) New optimization algorithms in physics
133. Kolinski A, Skolnick J (1996) Lattice models of protein folding. Dynamics and thermodynamics. Chapman & Hall, Austin
134. Mirny L, Shakhnovich E (2001) Protein folding theory: from lattice to all-atom models. *Annu Rev Biophys Biomol Struct* 30:361–396. doi:[10.1146/annurev.biophys.30.1.361](https://doi.org/10.1146/annurev.biophys.30.1.361)
135. Pande VS, Grosberg AY, Tanaka T (2000) Heteropolymer freezing and design: towards physical models of protein folding. *Rev Mod Phys* 72:259–314
136. Hart WE, Istrail S (1997) Robust proofs of NP-hardness for protein folding: general lattices and energy potentials. *J Comput Biol* 4:1–22
137. Dill KA, Ozkan SB, Shell MS, Weikl TR (2008) The protein folding problem. *Annu Rev Biophys* 37:289–316. doi:[10.1146/annurev.biophys.37.092707.153558](https://doi.org/10.1146/annurev.biophys.37.092707.153558)
138. Lau KF, Dill KA (1989) A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules* 22:3986–3997. doi:[10.1021/ma00200a030](https://doi.org/10.1021/ma00200a030)
139. Epstein CJ, Goldberger RF, Anfinsen CB (1963) The genetic control of tertiary protein structure: studies with model systems. *Cold Spring Harb Symp Quant Biol* 28:439–449. doi:[10.1101/SQB.1963.028.01.060](https://doi.org/10.1101/SQB.1963.028.01.060)
140. Crescenzi P, Goldman D, Papadimitriou C et al (1998) On the complexity of protein folding. *J Comput Biol* 5:423–465
141. Amin MHS, Choi V (2009) First-order quantum phase transition in adiabatic quantum computation. *Phys Rev A* 80:062326. doi:[10.1103/PhysRevA.80.062326](https://doi.org/10.1103/PhysRevA.80.062326)
142. Farhi E, Goldstone J (2009) Quantum adiabatic algorithms, small gaps, and different paths. *Comput Res Repos—CORR*
143. Perdomo-Ortiz A, Venegas-Andraca SE, Aspuru-Guzik A (2010) A study of heuristic guesses for adiabatic quantum computation. *Quantum Inf Process* 10:33–52. doi:[10.1007/s11128-010-0168-z](https://doi.org/10.1007/s11128-010-0168-z)
144. Allinger NL (1977) Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *J Am Chem Soc* 99:8127–8134. doi:[10.1021/ja00467a001](https://doi.org/10.1021/ja00467a001)

Protein Modelling

Gamble, A.

2014, VIII, 329 p. 117 illus., 78 illus. in color., Hardcover

ISBN: 978-3-319-09975-0