PCCP

Cite this: *Phys. Chem. Chem. Phys.*, 2012, **14**, 1507–1516

www.rsc.org/pccp

PAPER

The role of entropy in initializing the aggregation of peptides: a first principle study on oligopeptide oligomerization†

Gábor Pohl,[a] Imre Jákli,[b] Imre G. Csizmadia,[a] Dóra Papp,[a] Garibotto Francisco Matías[c] and András Perczel*[ab]

Received 5th September 2011, Accepted 22nd November 2011
DOI: 10.1039/c2cp22821a

The initiation and progression of Alzheimer's disease is coupled to the oligo- and polymerization of amyloid peptides in the brain. Amyloid like aggregates of protein domains were found practically independent of their primary sequences. Thus, the driving force of the transformation from the original to a disordered amyloid fold is expected to lie in the protein backbone common to all proteins. In order to investigate the thermodynamics of oligomerization, full geometry optimizations and frequency calculations were performed both on parallel and antiparallel β-pleated sheet model structures of $[HCO–(Ala)_{1–6}–NH_2]_2$ and $(For–Ala_{1–2}–NH_2)_{1–6}$ peptides, both at the B3LYP and M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) levels of theory, both in vacuum and in water. Our results show that relative entropy and enthalpy both show a hyperbolic decrease with increasing residue number and with increasing number of strands as well. Thus, di- and oligomerization are always thermodynamically favored. Antiparallel arrangements were found to have greater stability than parallel arrangements of the polypeptide backbones. During our study the relative changes in thermodynamic functions are found to be constant for long enough peptides, indicating that stability and entropy terms are predictable. All thermodynamic functions of antiparallel di- and oligomers show a staggered nature along the increasing residue number. By identifying and analyzing the 6 newly emerging dimer vibrational modes of the 10- and 14-membered building units, the staggered nature of the entropy function can be rationalized. Thus, the vanishing rotational and translational modes with respect to single strands are converted into entropy terms "holding tight" the dimers and oligomers formed, rationalizing the intrinsic adherence of natural polypeptide backbones to aggregate.

1. Introduction

Alzheimer's disease is categorized as a molecular conformational illness caused by protein misfolding. The initiation and progression of Alzheimer's disease is accounted for the formation and aggregation of the amyloid peptides. Furthermore, even at slightly abnormal cellular conditions, proteins and peptides can adopt a second, alternative fold, a multiple stranded β-pleated sheet like nanostructure, regardless of the length of the polypeptide chain and its amino acid composition.[1,2] These β-layers are also the common structural motif of the threadlike aggregates of the aforementioned amyloid peptide (Fig. 1).

During the aggregation process, more and more extended β-layers are formed resulting in macroscopic aggregates called plaques,[3–5] which serve then as seeds for further aggregation. There is a notion that through prohibiting the further growth of a plaque or breaking its structure the progress of Alzheimer's disease could be stopped or even reversed.[6] The inhibition or reversion of plaque formation has been attempted by

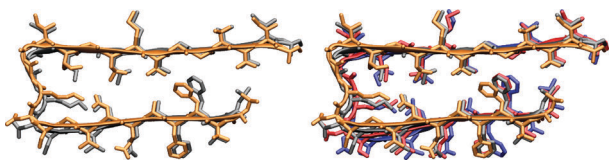

Fig. 1 Structure of a 2 stranded (left) and a 4 stranded (right) amyloid aggregate (PDB ID: 2BEG) (strands form parallel β-pleated sheet superstructures.).

[a] *Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd University, Pázmány Péter sétány 1/A, H-1117, Budapest, Hungary. E-mail: perczel@chem.elte.hu; Web: http://www.chem.elte.hu/departments/protnmr/index.html*
[b] *Protein Modeling Group of MTA-ELTE, Institute of Chemistry, Eötvös University, P.O. Box 32, H-1538 Budapest, Hungary*
[c] *Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis, Chacabuco 915, 5700 San Luis, Argentina*

† Electronic supplementary information (ESI) available: Fitting parameters, backbone torsional angles of the optimized dimer and oligomer structures discussed in this paper as well as their thermodynamic parameters as obtained at the B3LYP/6-31G(d), M05-2X/6-31G(d) and M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) levels of theory. See DOI: 10.1039/c2cp22821a

numerous research groups, however no lead molecules have yet been reported.

The present work aims to deepen our understanding of the theoretical background of oligomerization and plaque formation. By understanding the thermodynamics of β-pleated sheet formation from monomer strands, the driving forces of aggregation might be unveiled. Understanding the entropy changes upon dimerization would take us one step closer to predict the circumstances under which the aggregation process is reversed resulting in the dissociation of plaques. For the dimerization equilibrium of a polypeptide chain, the enthalpy change (ΔH) is composed of two parts (eqn (1)): ΔG measures the driving force of the process, T is the temperature and ΔS measures the extent of order ($\Delta S < 0$) or disorder ($\Delta S > 0$) introduced:

$$\Delta H = \Delta G + T\Delta S. \quad (1)$$

The total entropy of a thermodynamic system can be divided into two parts. The extrinsic entropy (or conformational entropy) can be estimated from the possible conformational states of a protein[7,8] and as such it can be useful in the characterization of the folding process itself.[9–11] However, an already folded protein exists in a well-defined structure. The entropy of that given conformation, intrinsic entropy, can also be computed and it provides valuable information about the stability of that given conformer.[12] Performing a full vibrational analysis, the intrinsic entropy characteristics of helix formation and the stability of other secondary structures were successfully investigated recently.[13] The intrinsic entropy of a conformer can be further divided into different components, translational (S_{Trans}), rotational (S_{Rot}) and vibrational (S_{Vib}):

$$S_{\mathrm{TOT}} = S_{\mathrm{Trans}} + S_{\mathrm{Rot}} + S_{\mathrm{Vib}} \quad (2)$$

$$S_{\mathrm{Trans}} + S_{\mathrm{Rot}} = S_{\mathrm{motion}} \quad (3)$$

Upon dimerization 3 translational and 3 rotational modes of motion are "lost" (corresponding to $\Delta S_{\mathrm{motion}}$), while 6 new vibrational modes appear, all corresponding to ΔS_{vib}. The loss of 3 translational and 3 rotational modes of motions results in an "entropy loss" and thus, an increase in order, while the 6 new vibrational modes contribute to an "entropy gain" or an increase of disorder. The identification and comprehensive analysis of these new vibrational terms are necessary to the proper characterization of the entropy change associated to polypeptide chain dimerization ($\Delta S_{\mathrm{dimerization}}$).

The present work aims to:

- study the variation of the $-T\Delta S$, ΔH and ΔG thermodynamic functions upon dimerization with respect to the number of amino acid residues for both the parallel and antiparallel β-pleated sheets,
- find a way to predict $-T\Delta S$, ΔH and ΔG functions for longer systems,
- explain the staircase-like increase of entropy observed for antiparallel sheets,
- identify the six new dimeric vibrational modes in 10-, 14- and 12-membered hydrogen bonded pseudo-ring systems,
- study the variation of the $-T\Delta S$, ΔH and ΔG thermodynamic functions upon polymerization with respect to the number of strands for parallel and antiparallel β-pleated sheets.

2. Methods

2.1. Molecular computation

All calculations were carried out in vacuum using the Gaussian 09 program package.[14] In order to investigate the entropy change upon dimerization, full geometry optimizations followed by frequency calculations were performed on HCO–$(Ala)_n$–NH_2 and $[HCO–(Ala)_n–NH_2]_2$ (n = 1–6) peptides at the M05-2X/6-311++G(d,p)//M05-2X/6-31G(d)[15] level of theory. The M05-2X functional developed by Truhlar *et al.*[15] was opted for instead of the commonly used B3LYP method, because it provides a more accurate description of dispersion forces, which are also important in the correct structural characterization of the aggregates.

For our calculations, the basis set superposition error (BSSE) was treated as follows:

- full geometry optimization and frequency calculations were performed at the M05-2X/6-31G(d) level of theory,
- followed by single point energy calculations on the optimized at the M05-2X/6-311G++(d,p) level of theory,
- then recalculating corrected enthalpy and Gibbs free energy values using the energy values obtained at the M05-2X/6-311G++(d,p) level of theory.

It has been shown that at this level of theory BSSE is not too significant (less than 0.5 kcal mol^{-1} per residue[16])

In order to obtain a better picture, the same calculations (full geometry optimization and frequency calculation) were carried out at the same level of theory for oligomers of mono- and dialanine, containing 1–6 strands.

2.2. Calculation of thermodynamic functions in vacuum

Thermodynamic functions were computed at standard temperature and pressure (*i.e.* at 298.15 K and 1 atm (101 325 Pa)). Results of the frequency calculations mentioned above provide the zero-point vibrational energy, the thermal correction to energy as well as the enthalpy (H), entropy (S) and Gibbs free energy (G) terms. Using these values, the enthalpy, entropy and Gibbs free energy changes upon dimerization can be computed as:

$$\Delta X = X_{\mathrm{dimer}} - 2X_{\mathrm{monomer}}, \text{ where } X = H, S \text{ or } G \quad (4)$$

The entropy contributions of the vibrations characteristic to the hydrogen bonds formed directly contribute to the ΔG value of dimerization (eqn (1)). As ΔG measures the driving force of a reaction the nature of these new dimerization interactions can make the difference between the aggregation being thermodynamically favored ($\Delta G < 0$) or disfavored ($\Delta G > 0$).

2.3. Calculation of thermodynamic functions in water

It has been shown that applying solvent models for optimizations has negligible effect on the backbone structures obtained, thus a solvent model was only applied for the single point energy calculations, while optimization was performed in vacuum.

For our calculations, the effect of solvation was treated as follows:

- full geometry optimization and frequency calculations were performed at the M05-2X/6-31G(d) level of theory in vacuum,

• followed by single point energy calculations on the optimized structure in water, using the IEF-PCM solvent model, at the M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) level of theory.

2.4. Vibrational analysis

In our first approach for the assignment of "dimeric vibrational modes", it was presumed that the atomic movements are dominated by synchronous translations within the strands, not affecting the overall β-strand geometry. By superimposing the atoms using the Quaternion fit[17,18] before and after applying the calculated translational vectors, the sum of the atomic distances measures the magnitude of the change in geometry in each of the strands. After sorting the vibrational modes by the quality of the fit, the first six items contain the sought dimeric vibrations (least difference in strand geometry after applying the translation vectors). Our results showed that a noticeable difference can be found between the best fitting normal modes and the remaining ones.

The other option (JustFit) of the dimer vibrational mode analysis is that they are considered as unique entities of the dimers, and thus they can't be assigned in any of the monomers. In order to make the vibrational modes of the monomer and the dimer comparable, each atom of the monomer has to be matched to each one of the dimeric strands. Subsequently, the translational vectors of the monomer were scaled to that of the dimer, to make the largest motional vectors equal. Finally, the difference of the scaled translational vectors, considering the normal and the opposite direction of the monomeric motional vector, was calculated. The magnitude of the difference was summarized for each monomeric vibrational mode and additionally for the "zero" vibrational mode, which contains all null translational vectors. The monomeric normal modes were sorted by the "sum" of the differences and the best fitting "valid" frequency was assigned to the dimeric normal mode. Frequency is considered valid, if it is not zero, or the difference between the normal mode frequency of the dimer and monomer is smaller by 20% of the dimer frequency or 200 cm^{-1}, if the preceding one is less than 200 cm^{-1}. If no valid fit is found, or if it is not the best one, it indicates the presence of a "dimeric vibrational mode".

In some cases the dimeric vibrational modes can't simply be described by means of a single monomeric vibration, rather by the combination of several ones (LinComb). In order to consider the latter option, after the atom assignment and normal mode translation vector transformation (as described above), the dimeric normal mode vibrations were approximated by the combination of the transformed normal modes. The number of the modes included should be reduced to a reasonable number, in order to perform the calculation within a reasonable time frame. The filter criterion is similar to the one used above: the dimeric and the monomeric normal mode frequency difference has to be less than 10%, or at least 200 cm^{-1}. Only 3 of the filtered monomeric vibrational modes (of positive and negative sign) were included in the linear combinations, achieved by the algorithm described previously in the CCA+ program.[19] The normal mode was assigned as a "dimer vibration" if the percentage of the "zero frequency" component is above 50%, or if the contribution of the largest component is smaller than 1.5 times the second largest component (indicating that no dominant monomer vibrational mode can be identified).

2.5. Structural considerations

Dimers of polypeptides can possess one type of parallel or two types of antiparallel arrangements.[2] Parallel and antiparallel β-pleated sheets have a different hydrogen bonding pattern: in parallel β-pleated sheets every hydrogen bonded "pseudo-ring" contains 12 atoms, while in antiparallel β-pleated sheets, the number of atoms forming the hydrogen bonded "pseudo-rings" is either 10 or 14 (Fig. 2). Being slightly different in nature, the entropy contributions of this three hydrogen bonded nanosystems are also expected to be different. After the thermodynamic properties of the dimerization of alanine and glycine diamides are cleared, oligoalanine strands and β-pleated sheets were investigated as they are more suitable models for the amyloid peptide. In order to obtain the thermodynamic functions of the formation of extended β-pleated sheet structures of oligoalanines, geometry optimizations and frequency calculations on monomeric $HCO–(Ala)_n–NH_2$, and both parallel (containing 12-membered H-bonded "pseudo-rings") and antiparallel (containing alternating 10- or 14-membered H-bonded "pseudo-rings") β-pleated sheets (Fig. 3) of $[HCO–(Ala)_n–NH_2]_2$, n = 1–6, were carried out at the M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) level of theory. The same calculations were carried out for mono- and dialanine oligomers, $[HCO–Ala–NH_2]_m$ and $[HCO–(Ala)_2–NH_2]_m$, m = 2–6 model systems. In the case of strands with odd number of residues and monoalanine oligomers with even number of

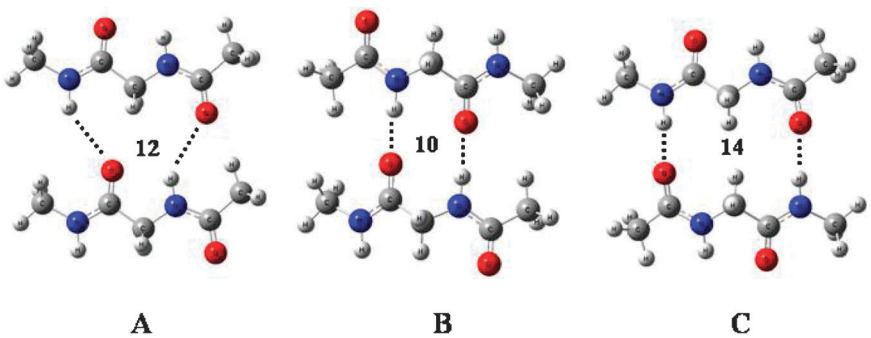

Fig. 2 Illustration of the 3 possible hydrogen bond systems formed during dimerization of Ac–Gly–NHMe: the 12-membered pseudo-ring (A) of a parallel, while the (B) 10- and (C) 14-membered pseudo-rings of the antiparallel arrangements.

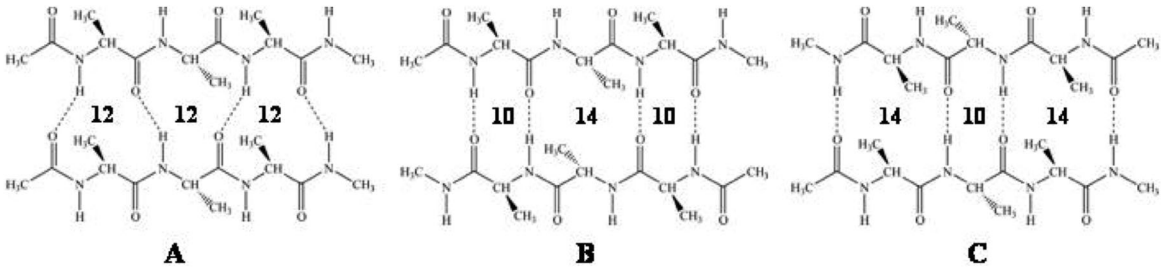

Fig. 3 Schematic representation of parallel (A) and the two types of antiparallel (B and C) arrangements of Ac–(Ala)$_3$–NHMe. Hydrogen bonds are denoted by dashed lines, numbers indicate the total number of atoms involved in the H-bonded, ''pseudo-ring'' formation. The total number of inter-chain H-bonds is $(m + 1)$, where m stands for the residue number within a polypeptide chain.

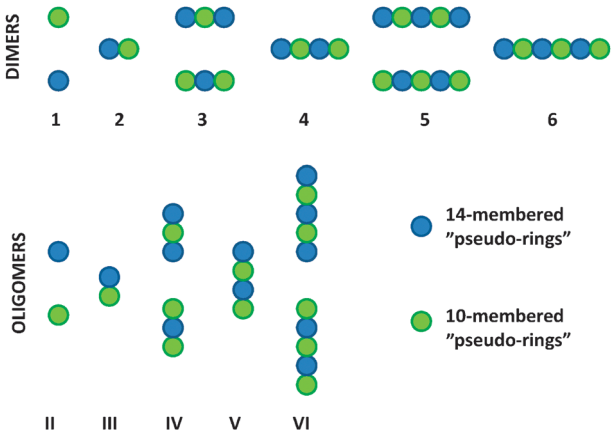

Fig. 4 Schematic representation of the structure of antiparallel oligoalanine dimers and monoalanine oligomers studied. The arabic numbers sign the number of amino acid residues in dimers and the roman numbers sign the number of monoalanine strands.

strands antiparallel β-pleated sheets will contain uneven number of 10- and 14-membered ''rings'' thus, both possible arrangements (more 10-membered or more 14-membered "rings") should be investigated (Fig. 3). Dialanine oligomers in antiparallel arrangements possess an even number of 10- and 14-membered hydrogen bonded pseudo-rings regardless the number of strands.

All this implies that there is only one sequence of parallel β-pleated sheets, but the situation with antiparallel β-pleated sheets is more complicated. In the case of antiparallel H-bonded β-pleated sheets, there is a single structure when the number of residues in the monomers is even and the number of strands in the oligomers is odd. However, the odd number containing oligoalanine dimers and even number containing monoalanine oligomers have two distinct arrangements as illustrated in Fig. 4.

3. Results and discussion

3.1. Change of thermodynamic functions upon the formation of parallel and antiparallel β-pleated sheets

Variations of the $\Delta X = -T\Delta S$, ΔH and ΔG thermodynamic functions upon dimerization with respect to the number of amino acid residues for both the parallel and antiparallel β-pleated sheets are shown in Fig. 5, and Tables S5 and S10 (ESI†) (fitting parameters in Table S17, ESI†). The variation of dimerization entropy (ΔS) as a function of the number of

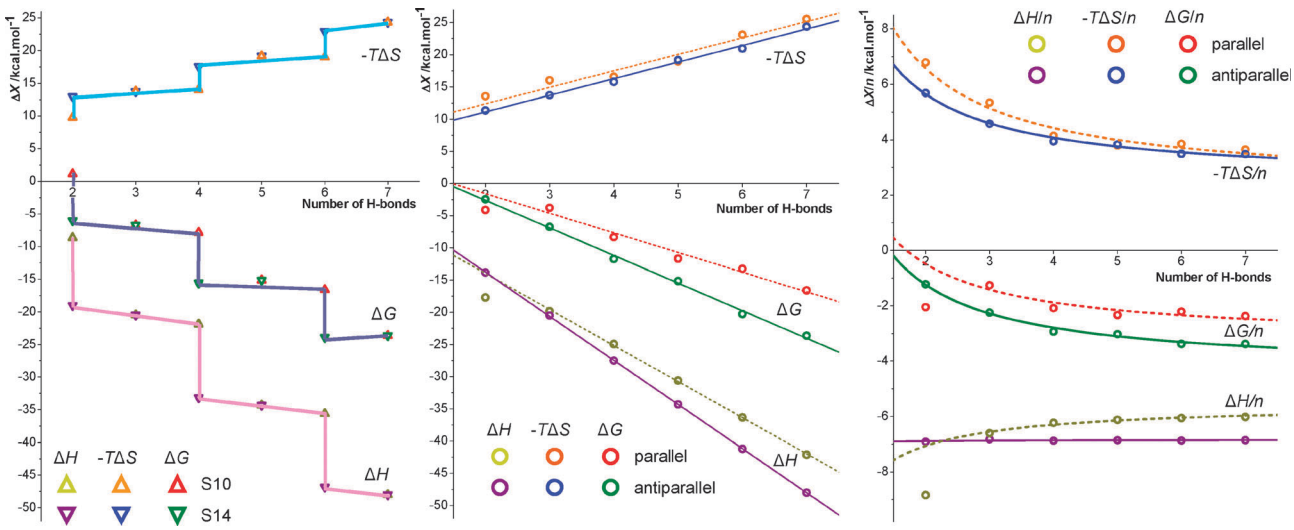

Fig. 5 Left: $-T\Delta S$, ΔH and ΔG *vs.* the number of inter-chain H-bonds for antiparallel β-pleated sheets. Middle: $-T\Delta S$, ΔH and ΔG *vs.* the number of inter-chain H-bonds for parallel dimers (dashed line, 12-membered rings) and for the average of the two antiparallel dimers (solid line, 10- and 14-membered pseudo-rings). Right: $-T\Delta S/n$, $\Delta H/n$ and $\Delta G/n$ *vs.* the number of inter-chain H-bonds in parallel β-pleated sheets (dashed line, 12-membered pseudo-rings) and for the average of the two antiparallel dimers (solid line, 10- and 14-membered pseudo-rings).

residues is staggered for antiparallel β-pleated sheets. When the $-T\Delta S$ of odd residue number containing antiparallel β-pleated sheet dimers are averaged, a straight line is obtained (R^2 = 0.993 for antiparallel). The Pearson correlation coefficient presented above and that of the parallel β-pleated sheets (R^2 = 0.946) show a rigorous correlation. Thus, the dimerization associated entropy term ($-T\Delta S$) correlates linearly with the number of inter-chain H-bonds, with high statistical significance. When the $-T\Delta S$ term of dimerization is normalized *per* hydrogen bonds (n) (Fig. 5), the curves obtained converge to a value of 2.50 for antiparallel, and 2.28 kcal mol^{-1} for parallel β-pleated sheets (Table 1), respectively.

The variation of dimerization enthalpy (ΔH) as a function of the number of residues is also staggered for antiparallel β-pleated sheets.[2] The average, however, is again a straight line with R^2 = 0.999 for antiparallel, and R^2 = 0.999 for parallel β-pleated sheets. When ΔH is normalized for H-bonds (Fig. 5) the function converges for $\Delta H_\infty/n$ to −6.83 for antiparallel and to −5.54 kcal mol^{-1} for parallel β-pleated sheets (Table 1), respectively. It can also be seen that in the case of antiparallel β-pleated sheets the dimerization enthalpy gain *per* H-bond formed shows a very small dependence on the number of inter-chain H-bonds.

This staggered function can be seen also for the antiparallel dimerization ΔG values (Fig. 5). The average is again a straight line of R^2 = 0.997 for antiparallel and R^2 = 0.975 for parallel β-pleated sheets, respectively. The normalized $\Delta G/n$ (Fig. 5) converges to a value $\Delta G_\infty/n$ = −4.34 kcal mol^{-1} for antiparallel and $\Delta G_\infty/n$ = −3.26 kcal mol^{-1} for parallel β-pleated sheets (Table 1), respectively.

Furthermore, the dependence of entropy on the number of hydrogen bonds is greater than that of the enthalpy term. Thus, for lower residue numbers, the formation of β-pleated sheets is less favored entropically, but this effect rapidly decreases with increasing residue number. While both entropy and enthalpy functions are converging to a constant number, indicating that for hexapeptides and above, the entropy loss and enthalpy gain are independent of the residue number. The enthalpy gain (−6.83 kcal mol^{-1} for antiparallel, −5.54 kcal mol^{-1} for parallel β-pleated sheets) will always exceed the entropy loss (2.50 kcal mol^{-1} for antiparallel and 2.28 kcal mol^{-1} for parallel β-pleated sheets) and thus, making β-pleated sheet formation thermodynamically favored.

In the case of parallel β-pleated sheets the last amino acid residue of one strand in the parallel β-pleated sheet structures computed in vacuum always adopts a γ-turn local conformation (Table S7, ESI†). This phenomenon results in a significant "artificial" lowering of the overall energy of the one residue long dimer. (This data point is omitted from the fitting, resulting in a slightly lower correlation coefficient.)

Table 1 Limits of the normalized thermodynamic (Y/n) functions, *vs.* the number of inter-chain H-bonds (n) as obtained at the M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) level of theory in vacuum and water[a]

	Parallel		Antiparallel	
(Y/n)	Vacuum	Water	Vacuum	Water
$-T\Delta S/n$	2.28	2.28	2.50	2.50
$\Delta H/n$	−5.54	−3.84	−6.83	−4.77
$\Delta G/n$	−3.26	−1.80	−4.34	−2.56

[a] See Methods.

Table 2 Comparison between the thermodynamic parameters of the generic building units of parallel and antiparallel β-pleated sheets as obtained at the M05-2X/6-311++G(d,p)//M05-2X/6-31G(d,p) level of theory

	Parallel TD function	Antiparallel TD function	Difference
ΔS/cal K^{-1} mol^{-1}	−7.65	−8.38	−0.73
$-T\Delta S$/kcal mol^{-1}	+2.28	+2.50	+0.22
ΔH/kcal mol^{-1}	−5.54	−6.83	−1.29
ΔG/kcal mol^{-1}	−3.26	−4.34	−1.08

It appears that there is not a too large difference in thermodynamic functions associated with the formation of parallel and antiparallel β-pleated sheets, leading to the oligo- and polymerization and subsequent plaque formation (Table 2). For longer polypeptides (larger plaques) the antiparallel arrangement is the more favored one out of the two systems; ΔG = −4.34 kcal mol^{-1} for antiparallel and −3.26 kcal mol^{-1} for parallel. However, experimental studies on amyloid plaques (*e.g.* Aβ1-42)[20] show that those aggregates are built up of parallel β-pleated sheets. Our results showed that parallel β-pleated sheets have a lower stability, which implies that only a relatively small perturbation is needed to dissociate plaques.

One reason for the discrepancy between theoretical and experimental results might be the effect of side chains. Experimental results were obtained on systems with considerable side chain stacking effects, which is out of the scope of the present study. The small difference in Gibbs free energy values between parallel and antiparallel sheets might be compensated by side chain interactions. This notion is further supplemented by the results of Hwang *et al.*[21] They have found on the oligomerization of five different β-sheet forming peptides (MD study) that the formation of aggregates initiated by side-chain/side-chain interactions is closely followed by the arrangement of the H-bond system. Nevertheless, antiparallel β-sheets were still found to be more stable.

Another reason might be a possible cooperativity in sheet formation: for 2–3 stranded arrangements of mono- or dialanine systems, parallel arrangement is more stable than antiparallel one. The initial parallel arrangements might subsequently drive the further aggregation of parallel sheets.

The M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) method is suitable enough to describe energy differences of the magnitude reported in the previous paragraph. The difference between the enthalpy values of antiparallel and parallel sheets (1.3 kcal mol^{-1}) is more than twice of the estimated BSSE at this level ~0.5 kcal mol^{-1}. Even though the difference in ΔG values is smaller the conclusions drawn from it are sound as the effect of BSSE error on frequencies (and thus entropy values) is negligible.

3.2. Change of thermodynamic functions and normalized thermodynamic terms for the generic building unit upon the formation of double stranded parallel and antiparallel β-pleated sheets in water

In order to further confirm our results obtained in vacuum, single point energy calculations were completed at the

M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) level of theory, in water by using the IEF-PCM solvent model (Table 1). Conclusions drawn from our calculations in vacuum hold for the solvent perturbed cases (Fig. 6): both for the averaged and normalized thermodynamic functions. The presence of a solvent results in the slight destabilization of sheets: $\Delta H = -4.78$ kcal mol^{-1}, $\Delta G = -2.20$ kcal mol^{-1} for the antiparallel and $\Delta H = -3.91$ kcal mol^{-1}, $\Delta G = -1.55$ kcal mol^{-1} for the parallel arrangements. Nevertheless, antiparallel arrangements are still found to be more favorable than parallel structures are.

3.3. Vibrational analysis of antiparallel β-pleated sheets

The origin of the staggered nature of ΔH and $T\Delta S$ as a function of the length of the polypeptide chain in the case of antiparallel sheets can be attributed to several reasons. Perczel *et al.*[2] have hypothesized that the shorter distance between the two oxygen atoms is the cause of the smaller enthalpy gain upon the formation of a 10-membered hydrogen bonded dimer. Dannenberg *et al.*[22] were explaining the energy differences between 10- and 14-membered hydrogen bonded dimers of polyglycin with the additional weak external C-5 hydrogen bonds formed in the latter structure.

The analysis of the newly emerging vibrational modes, found in the dimers but not in the monomers, enabled us to decipher the origin of differences in entropy of the 10- and 14-membered hydrogen bonded pseudo-rings. As shown in Table 3, all the 6 new dimeric vibrational modes (Fig. 7) could be identified by using any of the methods for vibrational analysis described above.

As the newly emerging dimer vibrations are vibrational modes along weak and noncovalent bonds, they are expected to have low frequencies. Thus, it seems sufficient to discuss the 10 lowest vibrational modes ($\nu < 200$ cm^{-1}) of the dimer. Both the frequency and the entropy contributions of these dimer vibrational modes are different for the 10- and the 14-membered nanosystems. In both structures the 5th, 9th and 10th vibrational modes are hydrogen bond "distortions". The H-bond distance in the 10-membered dimer is larger, resulting in weaker vibrations, and thus greater entropy contribution. 1st, 2nd and 3rd vibrational modes are associated with the Cα atoms in strands. Similar to the H-bonds, the Cα–Cα distance in a 10-membered system is also larger (Fig. 2B and C) than it is in a 14-membered one. Thus the corresponding vibrational modes are weaker and have therefore a greater entropy contribution ($T\Delta S_{\mathrm{VIB}}$ is 5.52 kcal mol^{-1} for 14-, while 8.56 kcal mol^{-1} for 10-membered rings).

As expected, calculations show that TS_{motion} (Table 4) decreases considerably and thus a great loss of complexity is noticed upon dimerization. However, this decrease from 40.89 kcal mol^{-1} is almost the same for both 10- and 14-membered H-bonded dimers (22.41 and 22.53 kcal mol^{-1}, respectively). It is expected, as the TS_{motion} entropy contribution depends on the molecular weight, which is the same for both the antiparallel and parallel β-pleated sheets composed of the very same number of atoms.

The formation of dimers not only results in the formation of 6 new dimeric vibrational modes, but also the splitting of the monomeric modes into symmetric and antisymmetric ones. This of course affects the entropy contribution of the latter vibrations. Interestingly enough, the splitting is not symmetric as the vibrations are more blue- than red-shifted and thus, the sum of the differences adds up to a negative value: -1.95 kcal mol^{-1} and -1.73 kcal mol^{-1}, respectively (Table 4). (Blue shifted frequencies result in lower entropy contribution.) This indicates that during dimerization not only the TS_{motion}, but also the contribution of the monomeric vibrational modes, $T\Delta S_{\mathrm{monomer}}$, has decreased. Thus, only the $T\Delta S_{\mathrm{VIB}}$ vibrational entropy contribution of the six new dimeric vibrational modes compensates the total entropy loss upon dimerization ($T\Delta S_{\mathrm{dimer}} = -12.96$ kcal mol^{-1} for 14 and

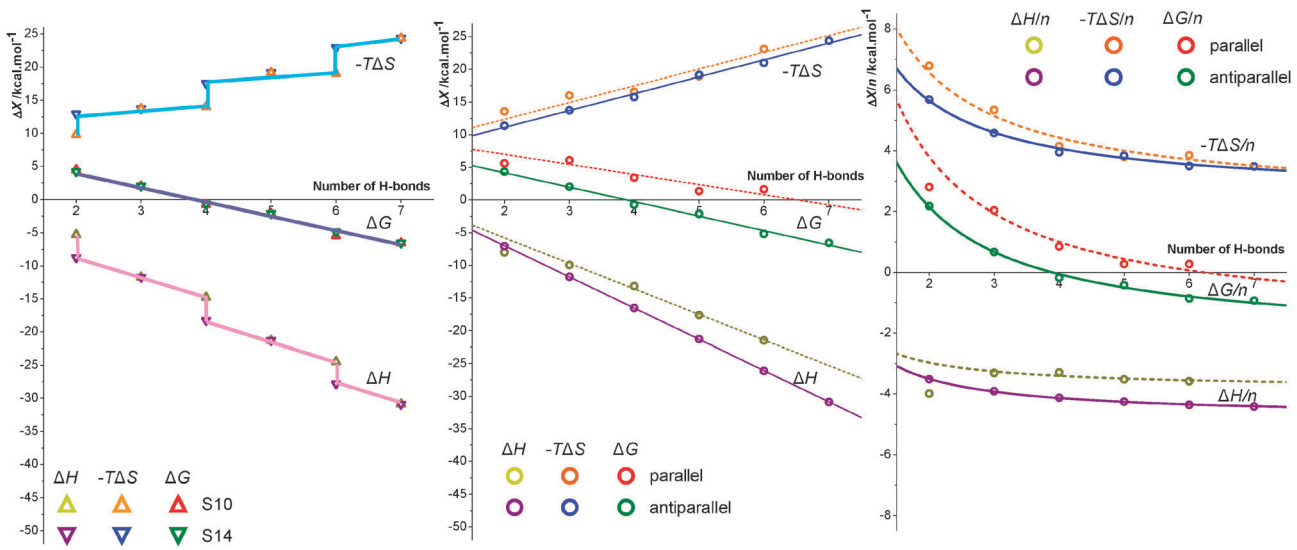

Fig. 6 Left: $-T\Delta S$, ΔH and ΔG *vs.* the number of inter-chain H-bonds for antiparallel β-pleated sheets, in water. Middle: $-T\Delta S$, ΔH and ΔG *vs.* the number of inter-chain H-bonds for parallel dimers (dashed line, 12-membered rings) and for the average of the two antiparallel dimers (solid line, 10- and 14-membered pseudo-rings), in water. Right: $-T\Delta S/n$, $\Delta H/n$ and $\Delta G/n$ *vs.* the number of inter-chain H-bonds in parallel β-pleated sheets (dashed line, 12-membered pseudo-rings) and for the average of the two antiparallel dimers (solid line, 10- and 14-membered pseudo-rings), in water.

Table 3 Analysis of the lowest 10 vibrational modes of (For–Ala–NH)$_2$ within a 10- and 14-membered H-bonded pseudo-rings: S10 and S14, respectively[a]

	S14				S10			
			Method applied[b]				Method applied[b]	
Vibration	ν/cm^{-1}	S/cal mol^{-1} K^{-1}	JustFit	LinComb	ν'/cm^{-1}	S/cal mol^{-1} K^{-1}	JustFit	LinComb
1	37.4	5.392	G	G	13.2	7.467	A	A
2	38.9	5.313	H	H	24.0	6.27	B	B
3	50.9	4.782	I	I	38.1	5.353	C	C
4	79.6	3.901	M1	M1	50.5	4.796	M1	M1
5	86.7	3.734	J	J	59.6	4.469	D	D
6	90.8	3.642	M2	M2	67.6	4.223	E	E
7	95.4	3.547	M1	M1	74.2	4.039	M2	M2
8	102.1	3.414	M2	M2	81.9	3.844	M1	M1
9	129.3	2.956	K	K	84.9	3.774	M2	M2
10	135.4	2.868	L	L	114.52	3.191	F	F

[a] ν indicates wavenumber, S signals the entropy contribution of the selected vibrational mode. [b] A, B, C, D, E and F stand for dimer vibrational modes in 10-membered systems (Fig. 8); G, H, I, J, K and L stand for dimer vibrations in 14-membered systems (Fig. 8); M1 and M2 stand for the symmetric and antisymmetric coupling of the lowest monomeric and 2nd lowest monomeric vibrational modes, respectively.

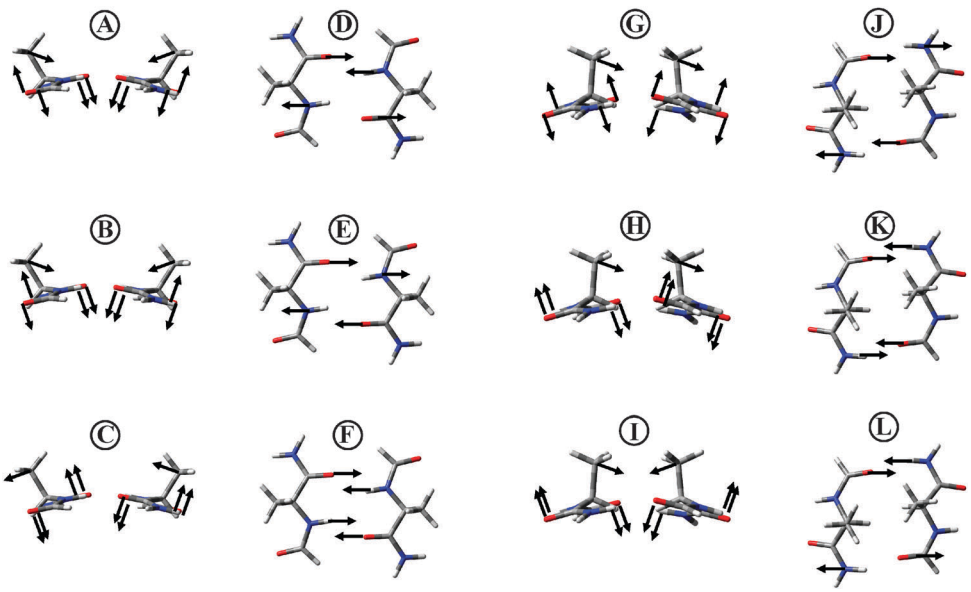

Fig. 7 (A–F) Schematic representation of the 6 "new" vibrational modes in a 10-membered hydrogen bonded pseudo-ring system appearing upon dimerization. (G–L) Schematic representation of the 6 "new" vibrational modes in the 14 membered hydrogen bonded systems appearing upon dimerization (see Table 2).

Table 4 The total-, motional- and vibrational-entropy terms (TS_{TOT}, TS_{motion}, TS_{VIB}) and their changes ($T\Delta S_{TOT}$, $T\Delta S_{motion}$, $T\Delta S_{VIB}$) upon the formation of the two possible antiparallel β-pleated sheet arrangements of (For–Ala–NH)$_2$ as obtained at the M05-2X/6-31G(d) level of theory[a]

	TS_{TOT}/ kcal mol^{-1}	$T\Delta S_{TOT}$/ kcal mol^{-1}	TS_{motion}/ kcal mol^{-1}	$T\Delta S_{motion}$/ kcal mol^{-1}	TS_{VIB}/ kcal mol^{-1}	$T\Delta S_{VIB}$/ kcal mol^{-1}	$T\Delta S_{dimer}$/ kcal mol^{-1}	$T\Delta S_{monomer}$/ kcal mol^{-1}
Mono (2×)	54.87	0	40.89	0	13.98	0	0	0
S14	41.91	−12.96	22.41	−18.48	19.50	5.52	7.47	−1.95
S10	45.08	−9.79	22.53	−18.36	22.54	8.56	10.29	−1.72

[a] At standard temperature and pressure (T = 298.15 K, p = 1 atm).

−9.79 kcal mol^{-1} for 10 membered rings). The net difference in vibrational entropies of 10- and 14-membered systems (as outlined before, $T\Delta\Delta S$ = −3.04) equals the "step size" of the entropy function (Fig. 5), which indicates that this difference is the major cause of entropy differences between the two types of antiparallel arrangements in β-pleated sheets.

In conclusion, as the entropy loss from losing 3 translational and 3 rotational degrees of freedom is the same for both systems the greater entropy gain from dimer vibrations lessens the total entropy loss of the formation of a 10-membered system making it entropically favored over the 14-membered one. However, compared to the 10-membered system, the

formation of a 14-membered system is still more favored thermodynamically because of the significantly greater enthalpy gain.

The two algorithms worked out for the analysis of dimeric vibrational modes were successfully applied here for dimers composed of single amino acid pairs only. However, due to the complex nature of these vibrational modes for longer dimers, a similar analysis was impossible. That being said, the origin of the staggered nature of the entropy function can be explained solely based on the results obtained for the above-described dimers of "monopeptides". In addition for the scope of the present paper the investigation of longer nanosystems is not needed.

3.4. Change of thermodynamic functions and normalized thermodynamic terms for the generic building unit upon the formation of multiple stranded parallel and antiparallel β-pleated sheets

As we have shown before, the formation of β-sheets is thermodynamically favored during dimerization of oligopeptides and is quite accurately predictable for longer nanosystems.

Upon studying the multiple stranded arrangements, in the case of monoalanine, the variation of thermodynamics as a function of the number of strands shows the same staggered motif for antiparallel β-pleated sheets (Fig. 8A) as it can be observed for the dimerization. Comparing the average values of these functions for antiparallel sheets with the values for parallel ones in Fig. 8B, above three strands the antiparallel sheets are found to be more stable (the fitted linear equation has a larger m value). However at an "early stage" of aggregation (fewer strands) the formation of the parallel β-sheets is favored. To see the relative gain of the aforementioned thermodynamic functions, their normalized values have been plotted (Fig. 8C), and as in the double stranded case, these hyperbolic curves obtained also converge to constant values (Table 5). Thus, predictions also could be made for larger systems for multiple stranded cases. Once again, the antiparallel β-pleated sheets are more favored to be formed.

Up to this point, the differences between the antiparallel and the parallel multiple stranded monoalanine systems have been studied. In order to get a better picture of the size effect the antiparallel sheets formed from mono- and dialanine are compared (Fig. 9A). We have also depicted the average values of the variation of the thermodynamic functions for monoalanine strands. For dialanines the same is not required, as rising the size of the aggregates by one strand, a 10- and a 14-membered ring will be formed and hence there is only one type of antiparallel dialanine type to be considered. As expected, the formation of the sheet is also favorable for dialanines because of the formation of an extra H-bond with respect to monoalanine oligomers, and thus further stability is gained. Here the values also have been normalized by the number of strands, and the hyperbolic curves obtained also converge to constant values (Fig. 9B), which reports larger stability for the β-sheets formed of dialanine strands (Table 5).

Table 5 Limits of the normalized thermodynamic (Y/n) functions, *vs.* the number of peptide strands (n) of monoalanine and dialanine as obtained at the M05-2X/6-311++G(d,p)//M05-2X/6-31G(d) level of theory[a]

(Y/n)	Monoalanine (parallel)	Monoalanine (antiparallel)	Dialanine (antiparallel)
$-T\Delta S/n$	11.60	11.55	13.61
$\Delta H/n$	−13.90	−15.38	−21.95
$\Delta G/n$	−2.30	−3.83	−8.34

[a] See computational details.

The above results indicate that the size extension of β-pleated sheets in both dimensions is favorable for aggregation and the normalized values are predictable. Up till now the intrinsic entropy of the nanosystems was only studied. However, by considering the conformational entropy of the same nanosystem, one can give a prediction for the entire entropy change. The values of enthalpy, Gibbs free energy and intrinsic entropy can be obtained from quantum chemical calculations, but the conformational entropy could only be estimated by the probability definition of entropy from the Boltzmann description of the system. The entire Gibbs free energy of a thermodynamic system can be expressed in the following way:

$$\Delta\Gamma = \Delta H - T\Delta(S + \Sigma) \quad (5)$$

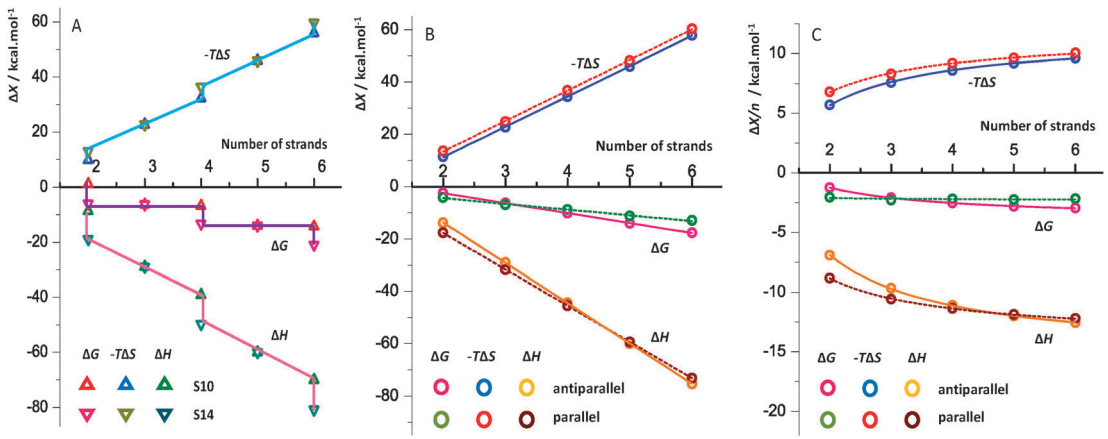

Fig. 8 Left: $-T\Delta S$, ΔH and ΔG *vs.* the number of strands for antiparallel β-pleated sheets of monoalanine. Middle: $-T\Delta S$, ΔH and ΔG *vs.* the number of strands for parallel oligomers (dashed line, 12-membered rings) and for the average of the two antiparallel oligomers for monoalanine (solid line, 10- and 14-membered pseudo-rings). Right: $T\Delta S/n$, $\Delta H/n$ and $\Delta G/n$ *vs.* the number of strands in parallel β-pleated sheets (dashed line, 12-membered pseudo-rings) and for the average of the two antiparallel oligomers for monoalanine (solid line, 10- and 14-membered pseudo-rings).

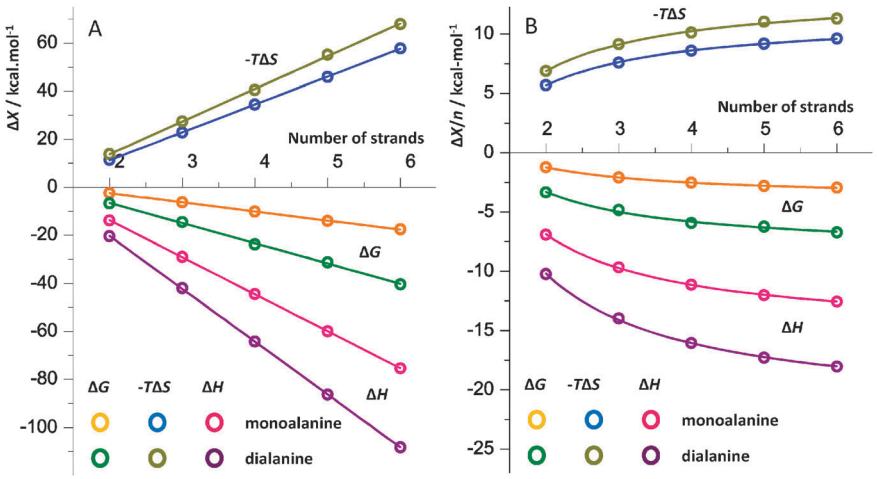

Fig. 9 Left: $-T\Delta S$, ΔH and ΔG *vs.* the number of strands for antiparallel oligomers of dialanine and for the average of the two antiparallel oligomers for monoalanine. Right: $T\Delta S/n$, $\Delta H/n$ and $\Delta G/n$ *vs.* the number of strands in antiparallel β-pleated sheets of dialanine and for the average of the two antiparallel oligomers for monoalanine.

where $\Delta\Sigma$ is the conformational entropy change derived from the folding of the protein and ΔS is the change in intrinsic entropy. Considering eqn (1) and (5), the entire Gibbs free energy of the system is:

$$\Delta\Gamma = \Delta G - T\Delta\Sigma \quad (6)$$

As an example, for the antiparallel octaalanine-octamer Gibbs free energy gained by the formation of a peptide strand containing 8 amino acids is $9 \times 3.08 + 6 \times 7 \times 3.83 = 188.58$ kcal mol^{-1} (Table 1 and 5).

The conformational entropy can be estimated in the following way: let us suppose that in an unfolded protein each amino acid residue has three possible conformers appearing with the same probability, thus $\Sigma = R\ln(3^{64})$, where $R = 8.314$ J mol^{-1} K^{-1}. If a strand of eight amino acids fold, the conformational entropy will decrease by $\Delta\Sigma = 8 \times R\ln(3^8)$, the value of $T\Delta\Sigma$ term on 298.16 K will be -41.70 kcal mol^{-1}. Thus, the corrected value of the Gibbs free energy is 188.58 kcal mol^{-1} $-$ 41.70 kcal mol^{-1} $=$ 146.88 kcal mol^{-1}, a relatively high value, which provides a significant stability for the β-pleated sheet containing 64 amino acid residues or a total of 63 H-bonds.

With respect to their unfolded state, the stability of folded proteins is relatively small, about 20 kcal mol^{-1}, they can easily be denatured by heat. However, in the case of larger β-pleated sheets this value is greater with an order of magnitude, and thus unexpected to easily unfold. This clearly explains why protein aggregates resist dissociating, regardless of their amino acid composition.

4. Conclusion

Our results show that during oligomerization of polypeptides the formation of 10-, 12- or 14-membered H-bonded pseudo-rings is accompanied by specific changes of their thermodynamic functions. However, in the case of parallel β-pleated sheets, the addition of each new 12-membered ring changes the thermodynamic functions by a constant value, and thus a linear correlation with the number of hydrogen bonds is found. In the case of antiparallel β-pleated sheets, where 10- and 14-membered rings alternate, the addition of a new 10-membered ring has little enthalpy gain and entropy loss, whereas the addition of a 14-membered ring has a considerable enthalpy gain and entropy loss.

The reason of enthalpy differences has already been discussed.[2,22] Based on our results the staggered nature of the entropy function for an antiparallel β-pleated sheet can be explained by the different entropy contributions of the dimer vibrations in 10- and 14-membered systems. In the case of dimers of single residue pairs, the dimer vibrational modes were successfully separated and assigned by using our recently developed method. These results suggest that 3 dimer vibrational modes are associated with the newly formed H-bonds, while the 3 additional ones are associated with the motion of the Cα atoms. As the distance between the 2 strands in the 10-membered system is larger, the corresponding dimer vibrational modes are weaker, resulting in an increased entropy contribution. However this is not the case for longer polypeptides, due to the more complex "origin" and nature of vibrational modes.

Monomeric vibrational modes are similar for both systems, however as the distance between the two strands is greater for the 10-membered pseudo ring, the aforementioned dimer vibrations are weaker and thus, have a greater entropy contribution. This difference is responsible for the staggered nature of the entropy function for antiparallel β-pleated sheets.

Because of the alternative arrangements of 10- and 14-membered systems, two types of antiparallel β-pleated sheets of different stability can be made from polypeptides with odd number of residues. By using the average of thermodynamic values of the two types of antiparallel β-pleated sheets with odd number of residues, like for parallel β-pleated sheets a linear correlation of thermodynamic function with residue number is observed. Normalized entropy and enthalpy functions give a hyperbolic equation converging to $-T\Delta S$ of 2.50 kcal mol^{-1} and ΔH of -6.83 kcal mol^{-1} for antiparallel and $-T\Delta S$ of 2.28 kcal mol^{-1} and ΔH of -5.54 kcal mol^{-1} for parallel β-pleated sheets. In the case of antiparallel β-pleated

sheets this equation is close to linearity, signaling a minimal dependence of relative enthalpy gain per H-bond formed on the total number of H-bonds.

The entropy change upon dimerization shows a much greater decrease with increasing residue number compared to the enthalpy change, resulting in greater stability for dimers made of longer systems. Thus, for hexapeptide systems and above, β-sheet formation is always thermodynamically favored as $\Delta G < 0$ for both parallel and antiparallel β-pleated sheets. However the formation of antiparallel β-pleated sheets is slightly more favored: the average ΔG per hydrogen-bond formed is -4.34 kcal mol^{-1} for antiparallel and -3.26 kcal mol^{-1} for parallel β-pleated sheets. Although the longer the system the greater is the stabilization upon dimer formation, the size of possible dimers is limited by the size of the misfolding proteins.

The results above were also confirmed by performing calculations at the same level of theory in water by using the IEF-PCM solvent model. The trends and conclusions observed in vacuum remains the same. Even though calculations in solvent predict sheets of somewhat lower stability (ΔG per hydrogen-bond formed is -2.20 kcal mol^{-1} for antiparallel and -1.55 kcal mol^{-1} for parallel dimers), antiparallel sheets are still found to be more favored than parallel sheets.

Our observations on oligoalanine dimers also hold for multiple stranded β-pleated sheet nano-arrangements. The variation of thermodynamic functions with the increasing strand number was found to be staggered for antiparallel β-pleated sheets, and the stability of these larger systems could also be predicted based on normalized thermodynamic functions. As in the case of oligoalanine dimers, antiparallel arrangements were found to be more favored than parallel ones. Moreover, as expected, oligomers containing dialanine strands are found to possess greater stability than the ones formed of monoalanine strands.

The present theoretical study gives a quantum-level explanation on why peptides and proteins self-transform into an amyloid like state. Because of the small ΔG value for the formation of parallel sheets, only a minor change in entropy would be needed to compensate the enthalpy gain of dimerization, which indicates that the search for molecules that hinder or reverse plaque formation is a reasonable and feasible effort.

Acknowledgements

This work was supported by grants from ICGEB (CRP/HUN08-03) and the Hungarian Scientific Research Fund (OTKA K72973, NK67800 and TáMOP-4.2.1.B-09/1/KMR).

References

1 C. M. Dobson, *Nature*, 2005, **435**, 747.
2 A. Perczel, P. Hudáky and V. K. Pálfy, *J. Am. Chem. Soc.*, 2007, **129**, 14959.
3 R. Nelson, M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riekel, R. Grothe and D. Eisenberg, *Nature*, 2005, **435**, 773.
4 M. R. Sawaya, S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. O. Madsen, C. Riekel and D. Eisenberg, *Nature*, 2007, **447**, 453.
5 F. Chiti and C. M. Dobson, *Nat. Chem. Biol.*, 2009, **5**, 15.
6 D. M. Holtzman, *Nature*, 2008, **454**, 418.
7 A. G. Street and S. L. Mayo, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 9074.
8 A. M. Berezhkovskii, F. Tofoleanu and N.-V. Buchete, *J. Chem. Theory Comput.*, 2011, **7**, 2370.
9 J. B. Thompson, H. G. Hansma, P. K. Hansma and K. W. Plaxco, *J. Mol. Biol.*, 2002, **322**, 645.
10 A. Baumketner and J.-E. Shea, *Biophys. J.*, 2005, **89**, 1493.
11 W. Xu, J. Ping, W. Li and J. Mu, *J. Chem. Phys.*, 2009, **130**, 164709.
12 B. Viskolcz, S. N. Fejer, S. N. K. Jensen, A. Perczel and I. G. Csizmadia, *Chem. Phys. Lett.*, 2007, **450**, 123.
13 B. Viskolcz, I. G. Csizmadia, S. K. Jensen and A. Perczel, *Chem. Phys. Lett.*, 2010, **501**, 30–32.
14 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian 03, Revision C.02*, Gaussian, Inc., Wallingford, CT, 2004.
15 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Theory Comput.*, 2006, **2**, 364.
16 T. Beke, I. G. Csizmadia and A. Perczel, *J. Am. Chem. Soc.*, 2006, **128**, 5158.
17 S. K. Kearley, *J. Comput. Chem.*, 1990, **11**, 1187.
18 D. J. Heisterberg, unpublished results, 1990, http://www.ccl.net/cca/software/SOURCES/C/quaternion-mol-fit/index.shtml.
19 I. Jákli and A. Perczel, *J. Pept. Sci.*, 2009, **15**, 738.
20 T. Luhrs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Doeli, D. Schubert and R. Riek, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 17342.
21 W. Hwang, S. Zhang, R. D. Kamm and M. Karplus, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12916.
22 R. Viswanathan, A. Asensio and J. J. Dannenberg, *J. Phys. Chem. A.*, 2004, **108**, 9205.

