Published in final edited form as:
J Phys Chem B. 2012 December 6; 116(48): 14017–14022. doi:10.1021/jp3094947.

The Folding of Acetyl(Ala)$_{28}$NH$_2$ and Acetyl(Ala)$_{40}$NH$_2$ Extended Strand Peptides into Antiparallel β-Sheets. A Density Functional Theory Study of β-Sheets with β-Turns

Jorge Ali-Torres and **J. J. Dannenberg**
Department of Chemistry, City University of New York - Hunter College and the Graduate School, 695 Park Avenue, New York NY 10065

J. J. Dannenberg: jdannenberg@gc.cuny.edu

Abstract

We report ONIOM calculations using B3LYP/D95** and AM1 on β-sheet formation from acetyl(Ala)$_N$NH$_2$ (N=28 or 40). The sheets contain from one to four β-turns for N=28 and up to six for N=40. We have obtained four types of geometrically optimized structures. All contain only β-turns. They differ from each other in the types of β-turns formed. The unsolvated sheets containing two turns are most stable. Aqueous solvation (using the SM5.2 and CPCM methods) reduces the stabilities of the folded structures compared to the extended strands.

Keywords

ONIOM; loop; solvation; CPCM; optimized structures; B3LYP; AM1

INTRODUCTION

Several reports from our group[1–3] and others[4–9] have dealt with molecular orbital studies of the energetics and structures of both parallel and antiparallel β-sheets. However, they have generally focused upon the aggregation of two or more strands into sheet-like structures, thereby neglecting the effects upon the stability attributable to the turns that are generally present in antiparallel β-sheets. In this paper, we report ONIOM DFT/AM1 calculations upon two sets of isomeric conformations of Acetyl(Ala)$_{28}$NH$_2$ and Acetyl(Ala)$_{40}$NH$_2$ (hereafter, referred to as A28 and A40, respectively) that contain different numbers of turns and strands. We used two different length peptides to obtain some information upon the effect of the peptide length upon the energetics and structures of the various isomeric structures. We chose the peptide lengths (24 and 40 residues) since these allow for folding into sheets with equal number of residues for most structures.

The β-sheet structures considered contain from one to four turns for A28 and one to six turns for A40. Each sheet has one more strand than the number of turns. We chose conformations that have close to the same number of residues per strand of the sheet as they provide the most H-bonds within the sheet. We found several different conformations of the sheets that differ in the nature of the turns, but focus on the lowest enthalpy structures.

Correspondence to: J. J. Dannenberg, jdannenberg@gc.cuny.edu.

Supporting Information Available: Complete data including structures '**b**', '**c**' and '**d**' and complete reference 12. This material is available free of charge via the Internet at http://pubs.acs.org.

CALCULATIONAL DETAILS

We used the ONIOM[10,11] method as programmed in the GAUSSIAN 09[12] suite of computer programs. ONIOM divides the system into up to three segments which can be treated at different levels of calculational complexity. Thus, one can treat the essential part of the system at the high level, while the less critical parts of the system might be calculated at the medium or low level. For this study we only used two levels (high and medium). We treated the backbones of each of the peptides at the high level, with only the side chains (methyls) at the medium level. The high level used hybrid DFT methods at the B3LYP/D95(d,p) level. This method combines Becke's 3-parameter functional,[13] with the non-local correlation provided by the correlation functional of Lee, Yang and Parr.[14] In the ONIOM method, there are unsatisfied valences in the high level at the interface between it and medium level. These valences were satisfied by using the default method of capping them with a hydrogen atom in the direction of the connecting atom in the medium level with a C-H distance of 0.723886 times the C-C distance. We used the AM1[15] semiempirical molecular orbital method for the ONIOM medium level.

With the ONIOM method, GAUSSIAN performs a high level (DFT) calculation only on those atoms designated with the valences terminated as described above.

All geometries were completely optimized in all (up to 1227) internal degrees of freedom. Vibrational calculations confirmed that the reported geometries are true minima on the PESs as there are no imaginary vibrational frequencies. We used these frequencies to calculate the enthalpies and free energies of the optimized species. We evaluated the counterpoise corrections (CP) for basis set superposition error (BSSE) only for the H-bonds between the strands using the single point a posteriori procedure. To do this, we cut the C-C bonds that connect strands to the turns and terminated them with H's. We then calculated the CP for each individual strand using the ghost orbitals of all the other strands and took the sum as the total CP following the procedure we have previously suggested. [16] We did not apply the CP-opt procedure[17] (which optimizes on a CP corrected surface) as this would not be possible after the turns are cut away.

In a previous study of five 17-amino acid peptides, [18] we found little difference in relative energies between this procedure and another where the side chains (in this case, the methyls) were subsequently optimized using DFT, with the (previously optimized) peptide chain held fixed. The current procedure also gave relative energies that agreed well with complete DFT optimizations for a series of five small 3_{10}-helical peptides. [19] We recently found this procedure to give reasonable results for peptides when compared to results from experimental databases, while several DFT methods that are specifically parametrized to include dispersion do not provide results in accord with experiments.[20] We believe this may be due to the interaction of induction and dispersion, as we have discussed elsewhere. [21] B3LYP/D95** also gives reasonable results for H-bonding interactions.[22,23]

We calculated free energies using both the SM5.2[24] and CPCM[25,26] models, and geometrically optimized using CPCM with the Pauling radii.[27] We used the AMPAC 8.16[28] program to calculate single point AM1 energies and solvation free energies of the peptides in their ONIOM optimized structures as we have done previously.[29–31] We used only the solvation free energies from these calculations. Solvated enthalpies combine the ONIOM ΔH's with the SM5.2 solvation ΔG's, while solvated free energies combine the ONIOM ΔG's with the SM5.2 solvation ΔG's. We followed the (optimized) CPCM, calculations with vibrational calculations to verify the minima and to obtain thermodynamic properties. We added the CPCM solvation free energies to the ΔH's and ΔG's obtained from the

vibrations at the optimized CPCM geometries to obtain the solvated ΔH's and ΔG's reported.

RESULTS AND DISCUSSION

We found four different kinds of local minima for the β-sheets containing turns which we have illustrated for A40 and labeled (in order of stability) A, B, C and D in figure 1. All four contain only β-turns using the definition that the distance between the α-carbons of residues i and i+3 be less than 7 angstroms and the structure not be helical,[32,33] which is still widely used.[34] However, they differ in the types of β-turns present as seen from the H-bonding topologies in the turns.

Venkatachalam[35] originally classified β-turns as types I, I′, II and II′ according to their φ and ψ dihedral angles at i+1 and i+2. This classification has been widely used since.[33] However, this classification requires both C_{14} and C_{10} H-bonds between the i^{th} and $i+3^{rd}$ residues in the turn in all these types and that the i+1st and i+2nd residues of type II or II′ have opposite chirality (or have one glycine which is achiral). Each turn contains up to three C=O's, one of which cannot form an H-bond due to its conformation. The other two have the potential to form H-bonds. We note here that we and others have previously found C_{10} H-bonds to be particularly weak.[3,6] We have recently proposed a new classification of β-turns and we shall apply that classification here.[36] This classification divides β-turns into two general classes: type A (which contain a C_{10} H-bond between the i^{th} and $i+3^{rd}$ residues) and type B (which lack this C_{10} H-bond, but may contain one or more C_7 H-bonds). The C_7 H-bonds that we observed may be more stable than the C_{10} (C_N H-bonds refer to cyclic H-bonding systems that contain N atoms).

The most stable, '**a**', structure contains only B_I and A_{II} turns, which are the most stable structures for isolated Type A and B turns, respectively.[36] While the A_{II} turns have φ and ψ dihedral angles close to those that are used to define type II′ turns in the Venkatachalam classification,[36] the type II′ turn requires that the $i+1^{st}$ residue to be D (or glycine).[33] Both the '**b**' and '**c**' structures contain only A_I and B_{II}, turns while '**d**' contains mostly A_I and A_{III} turns (there is one B_{II} turn in the A40 structure containing five turns). Thus, the C_{10} H-bonds appear in only some of the turns of the '**a**', '**b**' and '**c**' structures. We have calculated each kind of sheet that we report using each of these geometries for the turns (which are all completely optimized to local minima). In each case we have found the '**a**' type turn to give the lowest energy structures. For this reason, we consider only the '**a**' structures the following discussion unless specifically noted. The data for the '**b**', '**c**' and '**d**' structures can be found in the supplementary information.

The lowest energy structures found for A28 conformations containing from one to four turns and for A40 containing one to six turns appear in figures 2 and 3. We tried to keep the numbers of residues in each strand as close to equal as possible to make the numbers of interstrand H-bonds follow the number of turns. We confirmed all unsolvated structures and all but three structures optimized with CPCM as true minima by calculating the vibrational frequencies. Two of the remaining three have three low frequency imaginary frequencies and the other seven. Repeated attempts at optimization of these three solvated structures failed to find true minima. Since the imaginary frequencies are all quite small ($<15\ cm^{-1}$), we do not expect the energies to be substantially affected. However, these imaginary frequencies do affect the enthalpies and free energies calculated from the vibrations. Since the minimum contribution to the molar enthalpy for a low frequency vibration will be RT, we added this for each imaginary frequency to the calculated enthalpies. Table 1 displays the energetic folding data for both polyalanines, Figures 4 and 5 present the differences (per residue) in the ΔH's and ΔG's between the folded and extended structures (with and

without solvation) for the A28 and A40. As seen from figures 4 and 5, the most stable sheets for both A28 and A40 contain two turns (three strands). The small enthalpic preference for two versus three turns for unsolvated A28 increases for A40 which suggests that this property will remain for larger systems

For the unsolvated structures, the ΔH's per residue become less negative as the number of turns increases beyond two, as seen from figure 4. Figure 5 displays a similar trend for ΔG. The foregoing results from an increase in the number of H-bonds with the number of turns while each H-bond becomes less stabilizing due to the strain induced by the turns. The increased number of H-bonds dominates for small numbers of turns, but the decrease in stability per H-bond dominates for larger numbers of turns. We have included detailed data on the H-bond lengths and average H-bond stabilities for each sheet in the supplementary information.

Aqueous solvation tends to decrease the stability of the sheets versus the extended strands. One might expect this behavior as the extended strands have more solvent-exposed H-bond acceptors and donors. The number of exposed acceptors and donors decreases with increasing numbers of turns. As seen from figures 4 and 5, the energetic behavior of the optimized CPCM solvated structures and those calculated using (single point) SM5.2 differ somewhat. While both methods show the sheets to be less solvated than the extended strands, we see from the ΔHs of figure 4 that, compared to CPCM, SM5.2 better solvates sheets with fewer turns and more poorly solvates those with more turns. In fact, the plots for the solvated enthalpies vs. number of turns cross for both A28 and A40 in figure 4.

Comparison of the enthalpic and free energy data (table 1 and figures 4–5) suggest that the increasing rigidity of the sheets over the strands decreases the relative entropy of the former. The unsolvated A28 with one turn has a slightly lower ΔG/residue than the analogous structure for A40. We note that calculating the entropy (thus, the free energy) of these structures from the harmonic vibrations can introduce errors. These errors become particularly problematic for low barrier torsional modes, such as those involved in methyl group rotations. As we only use relative free energies, we expect the methyl rotational errors to cancel as the number of such modes does not change upon folding.

The solvation free energies cannot readily be calculated from first principles. When quantum mechanical methods are used, solvation generally is calculated using a continuum model of some kind which relies on empirical parameters. The other popular alternative, molecular dynamics, generally depends upon parametrized force fields and does not usually rely on quantum mechanical first principles. Such methods can be useful, but treat the problem from an entirely different approach.

The solvation ΔG relative to the extended strand becomes more positive as the number of turns increases, presumably due to the decrease in H-bond donors (N-H's) and acceptors (C=O's) that are exposed to the solvent as more H-bonds within the sheets form with increasing numbers of turns. Since the A40 has more exposed H-bond donors and acceptors than A28 for the same number of turns, the aqueous solvation stabilization for the former is always greater. Large β-sheets without polar side chains tend to become insoluble as they become larger, which is exemplified by silk and amyloids. The observed tend in ΔG_{solv} accords with this observation.

We note that the results presented here apply to polyalanines and that sheets containing other residues may be stabilized by specific interactions between residues, i.e., glutamines.[37–40]

CONCLUSIONS

Unsolvated polyalanines containing 28 and 40 residues prefer β-sheets containing two turns (three strands) enthalpically. Considering the free energy reduces the stabilities of the sheets with increasing numbers of turns. Not surprisingly, the results suggest that longer peptides can sustain β-sheets with more turns. Aqueous solvation reduces the stabilities of the folded sheets in comparison with the extended strand.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The work described was supported by Award Number SC1AG034197 from the National Institute on Aging.

References

1. Plumley JA, Dannenberg JJ. J Phys Chem B. 2011; 115:10560. [PubMed: 21797271]
2. Plumley JA, Tsai MI-H, Dannenberg JJ. J Phys Chem B. 2011; 115:1562. [PubMed: 21261311]
3. Viswanathan R, Asensio A, Dannenberg JJ. J Phys Chem A. 2004; 108:9205.
4. Zhao Y-L, Wu Y-D. J Am Chem Soc. 2002; 124:1570. [PubMed: 11853419]
5. Suhai S. Int J Quantum Chem. 1991; 40:559.
6. Perczel A, Gaspari Z, Csizmadia IG. J Comput Chem. 2005; 26:1155. [PubMed: 15952205]
7. Scheiner S. J Phys Chem B. 2006; 110:18670–18679. [PubMed: 16970498]
8. Wang C-S, Sun C-L. J Comput Chem. 2010; 31:1036. [PubMed: 19821516]
9. Ireta J. J Chem Theory and Comp. 2011; 7:2630.
10. Morokuma K. Bull Korean Chem Soc. 2003; 24:797.
11. Vreven T, Morokuma K. J Chem Phys. 2000; 113:2969.
12. Frisch, MJ.; Trucks, GW.; Schlegel, HB.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, GA., et al. Gaussian 09, Revision A2. Gaussian, Inc; Wallingford CT: 2009.
13. Becke AD. J Chem Phys. 1993; 98:5648.
14. Lee C, Yang W, Parr RG. Phys Rev B. 1988; 37:785.
15. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP. J Am Chem Soc. 1985; 107:3902.
16. Turi L, Dannenberg JJ. J Phys Chem. 1993; 97:2488.
17. Simon S, Duran M, Dannenberg JJ. J Chem Phys. 1996; 105:11024.
18. Wieczorek R, Dannenberg JJ. J Am Chem Soc. 2003; 125:8124. [PubMed: 12837081]
19. Wieczorek R, Dannenberg JJ. J Am Chem Soc. 2004; 126:14198. [PubMed: 15506786]
20. Marianski M, Asensio A, Dannenberg JJ. J Chem Phys. 2012; 137:044109. [PubMed: 22852599]
21. Roy D, Marianski M, Maitra N, Dannenberg JJ. J Chem Phys. 2012; 137:134109. [PubMed: 23039587]
22. Plumley JA, Dannenberg JJ. J Comput Chem. 2011; 32:1519. [PubMed: 21328398]
23. Simon S, Duran M, Dannenberg JJ. J Phys Chem A. 1999; 103:1640.
24. Hawkins GD, Cramer CJ, Truhlar DG. J Phys Chem B. 1998; 102:3257.
25. Barone V, Cossi M. J Phys Chem A. 1998; 102:1995.
26. Cossi M, Rega N, Scalmani G, Barone V. J Comput Chem. 2003; 24:669. [PubMed: 12666158]
27. Besler BH, Merz KM, Kollman PA. J Comput Chem. 1990; 11:431.
28. AMPAC 8.16. Semichem, Inc; Shawnee KS:
29. Marianski M, Dannenberg JJ. J Phys Chem B. 2012; 116:1437. [PubMed: 22201227]
30. Tsai MI-H, Xu Y, Dannenberg JJ. J Phys Chem B. 2009; 113:309. [PubMed: 19072621]

31. Salvador P, Asensio A, Dannenberg JJ. J Phys Chem B. 2007; 111:7462. [PubMed: 17552560]
32. Lewis PN, Momany FA, Scheraga HA. Proc Nat Acad Sci U S A. 1971; 68:2293.
33. Rose GD, Glerasch LM, Smith JA. Adv Protein Chem. 1985; 37:1. [PubMed: 2865874]
34. Kuo-Chen C. Anal Biohem. 2000; 286:1.
35. Venkatachalam CM. Biopolymers. 1968; 6:1425. [PubMed: 5685102]
36. Roy D, Pohl G, Ali-Torres J, Marianski M, Dannenberg JJ. Biochemistry. 2012; 51:5387. [PubMed: 22731966]
37. Plumley JA, Dannenberg JJ. J Am Chem Soc. 2010; 132:1758–1759. [PubMed: 20088582]
38. Roy D, Dannenberg JJ. Chem Phys Lett. 2011; 512:255. [PubMed: 21927063]
39. Sikorski P, Atkins E. Biomacromolecules. 2005; 6:425. [PubMed: 15638548]
40. Perutz MF, Johnson T, Suzuki M, Finch JT. Proc Nat Acad Sci U S A. 1994; 91:5355.

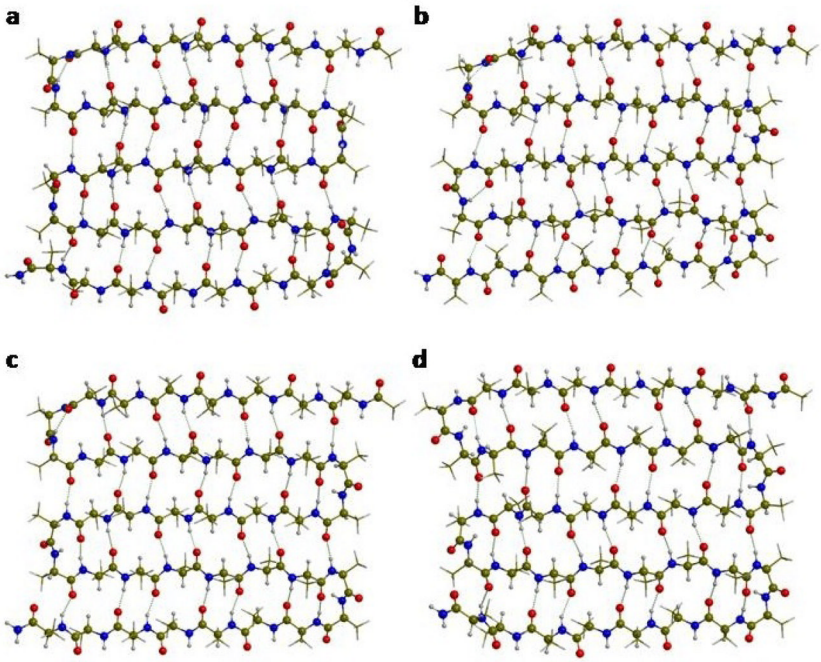

Figure 1.

Examples of the four kinds of structures found using A40 (presented in increasing energetic order): the lowest energy structure, **a**, contains C_7 H-bonds in the 1^{st} and th turns and C_{10} H-bonds in the 2^{nd} and 3^{rd} ; structure **b** contains C_7 H-bonds in the 1^{st} and 3^{rd} turns and C_{10} H-bonds in the 2^{nd} and 3^{rd} ; structure **c** has only one C_7 H-bonding turn, the 1^{st}, and C_{10} H-bonds in the others, while the highest energy structure, **d,** contains C_{10} H-bonds in all turns.

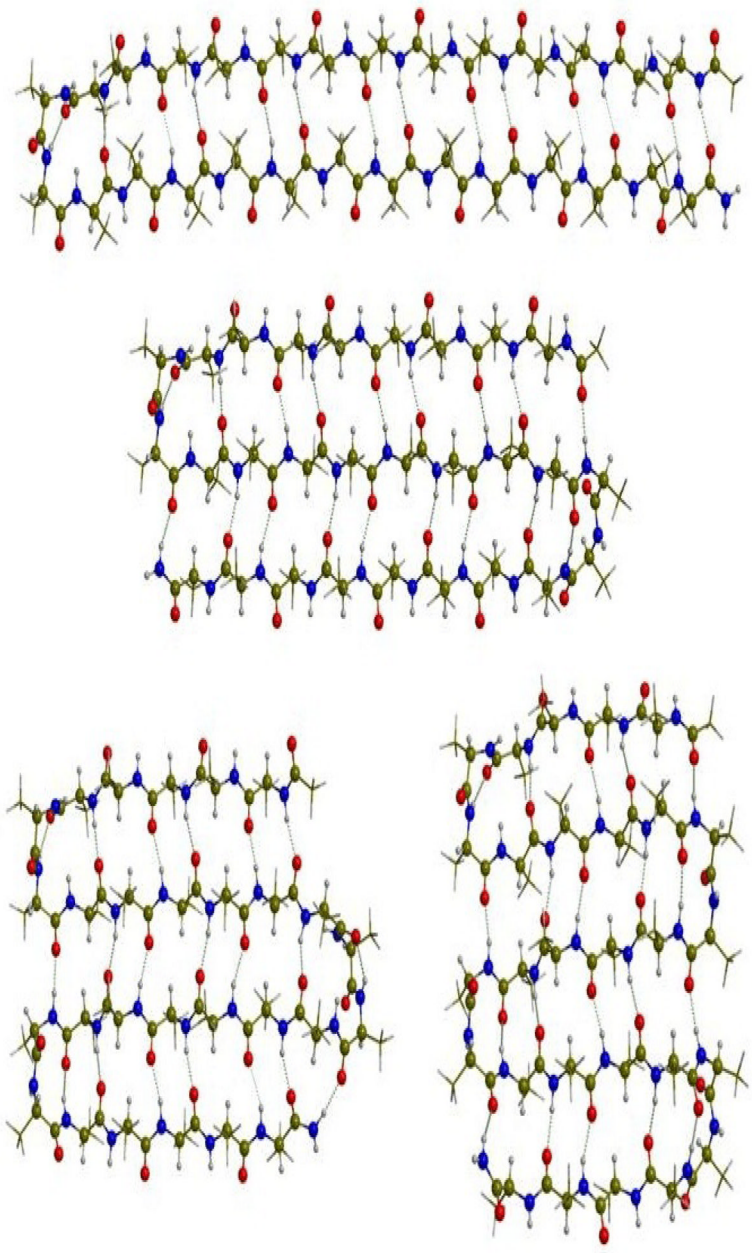

Figure 2.
Structures of A28 containing one to four β-turns.

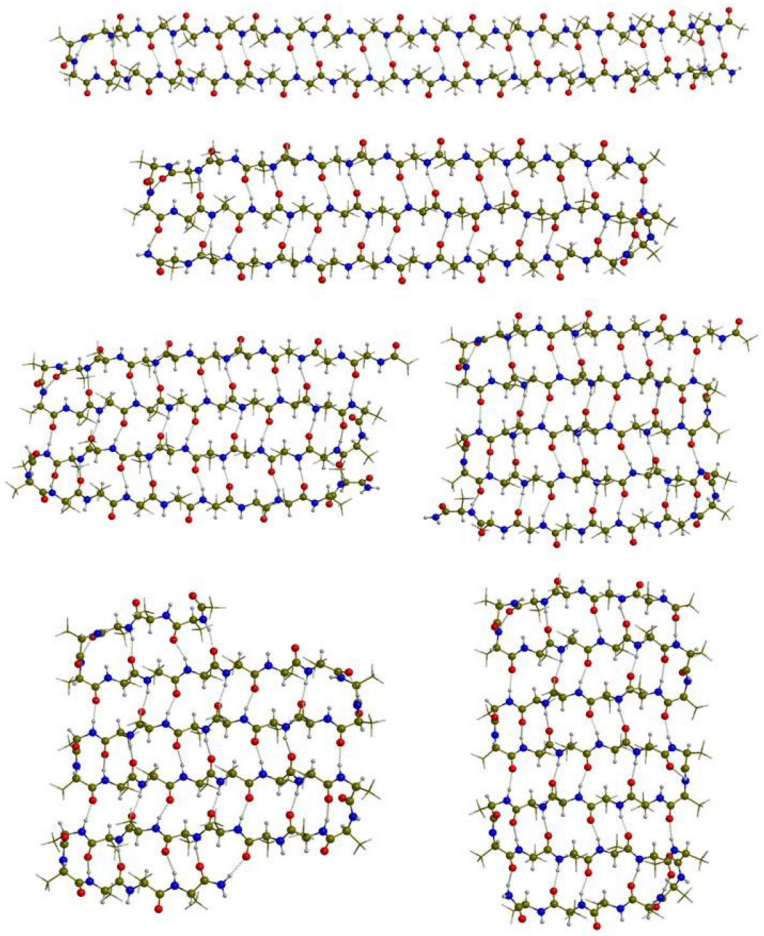

Figure 3.
Structures of A40 containing from one to six turns.

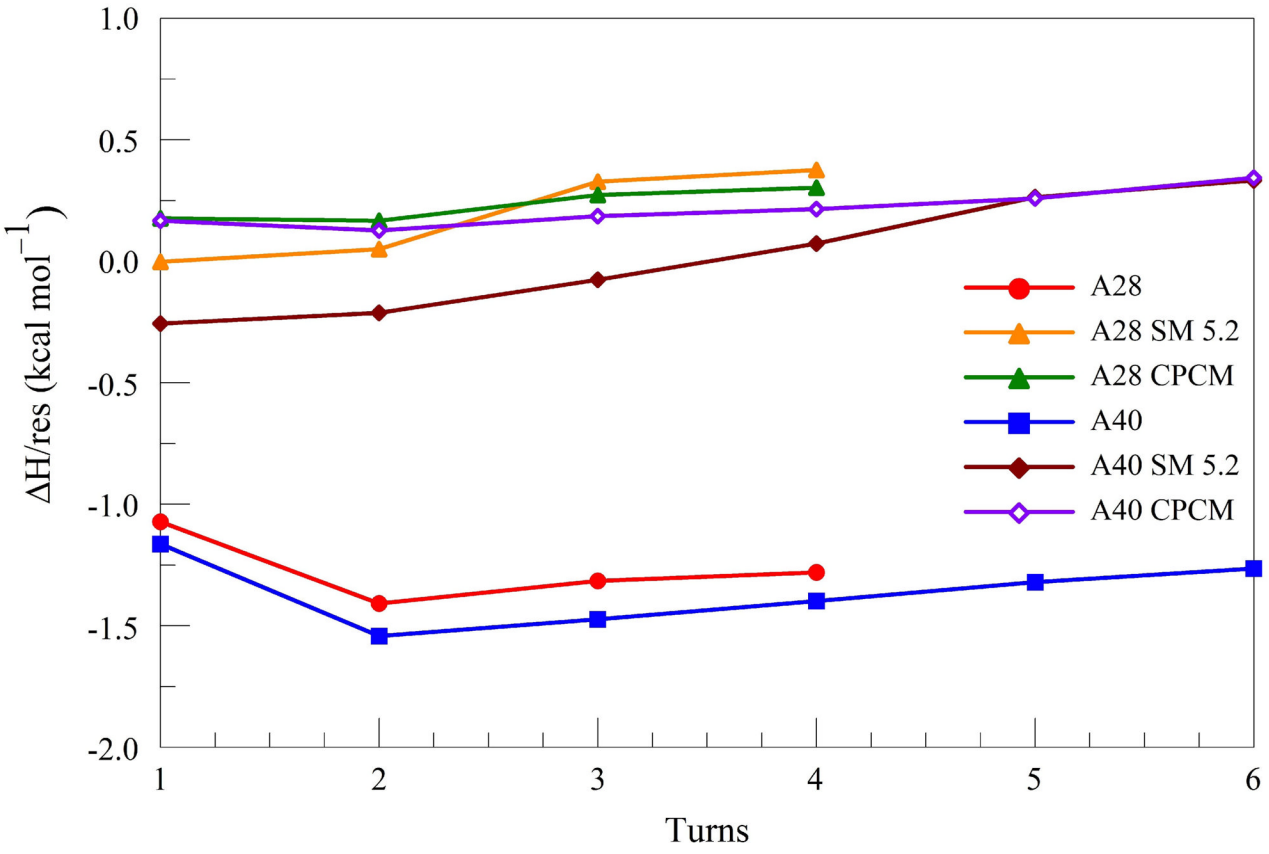

Figure 4.
ΔHs (relative to extended strands) versus number of turns with and without solvation (indicated as SM5.2 or CPCM).

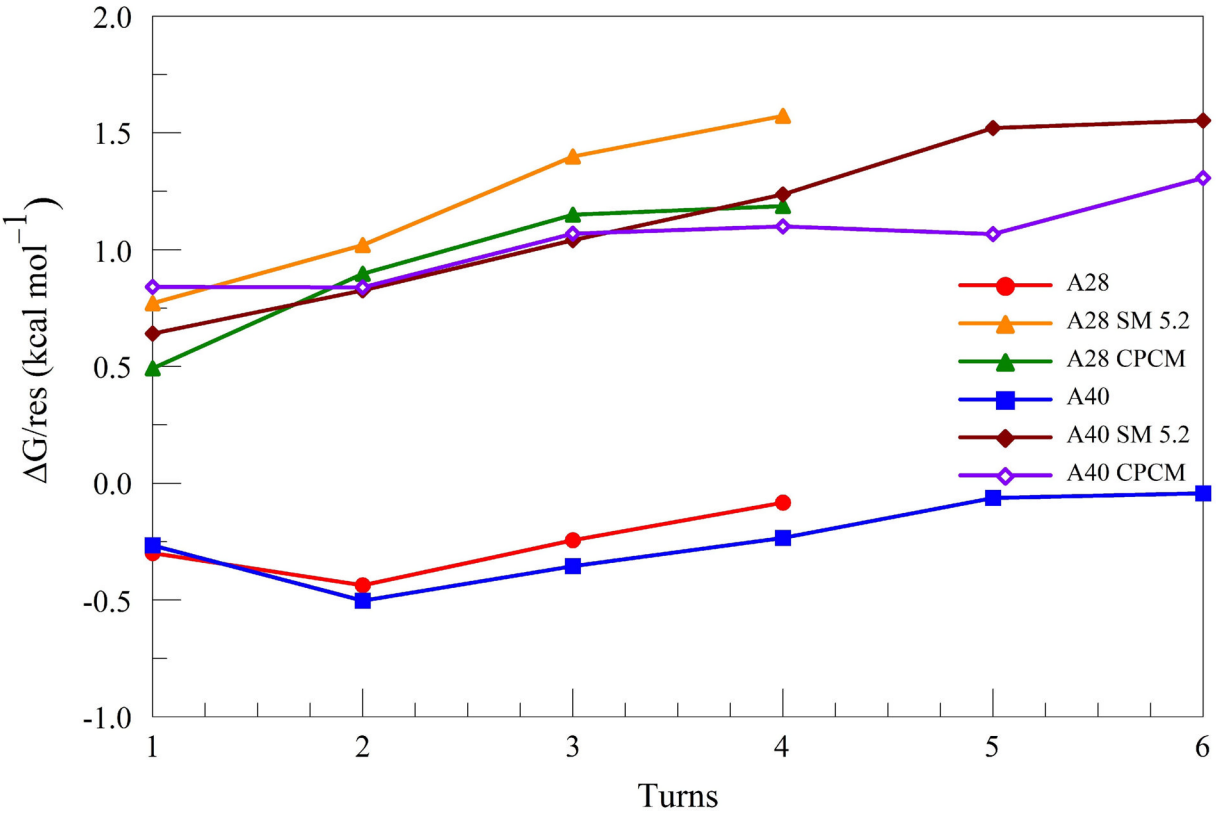

Figure 5.
ΔGs versus extended strands with and without solvation free energies (marked SM5.2 or CPCM).

Table 1

Energetics of folding and solvation (versus extended strand) using (single point) SM5.2 and (optimized) CCPM.

					ΔG_{solv}	
Turns	**H-bonds**	**ΔE**	**ΔH**	**ΔG**	**SM5.2**	**CPCM**
	A28					
1	14	−33.5	−30.0	−8.4	30.0	22.9
2	18	−45.8	−39.4	−12.2	40.8	39.5
3	20	−43.3	−36.8	−6.8	46.0	41.4
4	20	−42.2	−35.9	−2.3	46.4	38.2
	A40					
1	20	−51.9	−46.5	−10.6	36.3	45.2
2	26	−70.7	−61.7	−20.1	53.2	56.2
3	27	−68.2	−58.6	−14.2	55.8	59.3
4	28	−65.3	−55.9	−9.3	58.8	56.2
5	30	−62.9	−52.8	−2.5	63.4	48.6
6	30	−60.6	−50.6	−1.7	63.9	58.9