Proceedings of the 24th American Peptide Symposium
Ved Srivastava, Andrei Yudin, and Michal Lebl (Editors)
American Peptide Society, 2015 *http://dx.doi.org/10.17952/24APS.2015.020*

The Main-Chain Oxygen: Unappreciated Effects on Peptide and Protein Structure

Robert W. Newberry[1] and Ronald T. Raines[1,2]
[1]Dept. of Chemistry, University of Wisconsin-Madison, Madison, WI, 53706; [2]Dept. of Biochemistry, University of Wisconsin-Madison, Madison, WI, 53706, USA

Introduction

Current limitations in protein structure prediction and design suggest an incomplete understanding of the forces governing protein folding. As such, noncovalent interactions in proteins, particularly hydrogen bonds, have received great attention [1,2]. In common secondary structure patterns like the α-helix and β-sheet, main-chain N–H hydrogen bond donors approach their carbonyl acceptors approximately along the carbonyl bond axis [3], despite conventional wisdom that hydrogen bond energies are maximized when donors approach at 120° to the carbonyl bond axis [4]. This observation can be rationalized using a modern, quantum-mechanically based model of the carbonyl lone pairs that indicates that the two orbitals differ from the sp^2-hybridized VSEPR "rabbit ears" assumed commonly. Specifically, one lone pair, approximately *sp*-hybridized, is oriented along the carbonyl bond axis, while the second, purely *p*-orbital orients orthogonally (Figure 1).

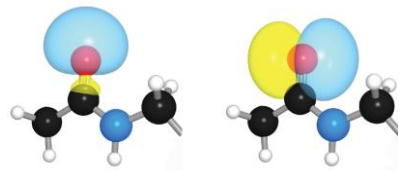

Fig. 1. s-Type (left) and p-type (right) carbonyl lone pairs.

Canonical hydrogen bonds in protein secondary structure therefore often employ the *s*-rich lone pair; however, the role of the *p*-type lone pair is less clear. We have previously noted that backbone $n\rightarrow\pi^*$ interactions are well poised to exploit this *p*-type lone pair [5]. In an $n\rightarrow\pi^*$ interaction, the filled *p*-type lone pair of a carbonyl oxygen interacts with the empty π^* orbital of an adjacent carbonyl group, and the mixing of these orbitals releases energy. These interactions have energies generally greater than 0.27 kcal/mol each [6], and are ubiquitous in folded proteins [7,8], particularly in the α-helix [9]. Yet, no analogous role for the *p*-type carbonyl lone pair has been identified in β-sheets. We now posit that a previously unappreciated hydrogen bond occurs within the backbone of individual residues in β-sheets.

Results and Discussion

Upon inspection of an idealized β-sheet, we noted close proximity of the *p*-type carbonyl lone pair with the amide N–H group of the same residue (Figure 2) and hypothesized an attraction between them that could be analogous to canonical hydrogen bonds. Compared to traditional hydrogen bonds, these putative interactions are highly distorted, so we first set out to determine if these interactions have the properties typical of other hydrogen bonds.

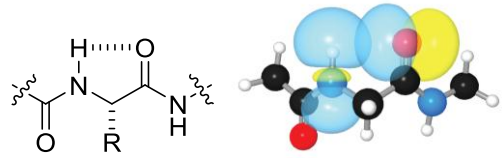

Fig. 2. Putative C5 hydrogen bond in the peptide backbone.

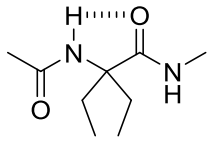

Fig. 3. Diethylglycine scaffold for studying C5 hydrogen bonds.

To probe a single interaction, we preorganized the putative donor and acceptor using a diethylglycine model system (Figure 3); diethylglycines have been shown by computation [10,11], as well as NMR [12-14] and vibrational spectroscopies [12,15], to adopt the "C5" geometry [16], which is an extended conformation that places the carbonyl oxygen in close proximity to the amide proton. Having realized the necessary geometry, we probed the putative interaction by replacing an amide hydrogen bond acceptor with an ester, which is known to attenuate *bona fide* hydrogen bonds. We found that attenuating the putative C5 hydrogen bond caused an increase in the stretching frequency of the donor in the infrared spectrum. In addition, we found that replacement of the amide acceptor with an ester

caused an upfield chemical shift of the donor proton, despite the greater electron-withdrawing character of the ester. Finally, we found that an amide acceptor was much more effective at protecting the donor proton from H/D exchange than was the ester. Together, these data show that these interactions do constitute hydrogen bonding.

To evaluate the relevance of these interactions for proteins, we probed their contributions to the conformational stability of a "tryptophan zipper" model β-hairpin peptide [17]. Upon selective attenuation of the C5 hydrogen bond using an amide-to-ester substitution [18,19], we observed a decrease in global thermostability by CD spectroscopy. Conversely, selective enhancement of a C5 hydrogen bond imparted additional thermostability to this peptide. Together, these results demonstrate that C5 hydrogen bonds are operative in β-sheet structures. Finally, to evaluate their cumulative contributions to conformational stability, we calculated the total energy of C5 hydrogen bonds in the structures of folded proteins and found that they contribute an average of 5 kcal/mol of stabilizing energy to a 100-residue protein.

Our results highlight the importance of a previously unappreciated force in protein folding: the C5 hydrogen bond. The discovery of this interaction explains how proteins make effective use of *both* lone pairs of the main-chain oxygen to stabilize secondary structure. We believe that the integration of these interactions into experimental and computational approaches would advance the understanding of the folding and stability of peptides and proteins.

Acknowledgments

We are grateful to G.J. Bartlett, D.N. Woolfson, and L.L. Kiessling for contributive discussions. We thank W.M. Westler for assistance with NMR spectroscopy, and T. Zhang and M.T. Zanni for assistance with IR spectroscopy. This work was funded by grants R01 AR044276 (NIH) and CHE-1124944 (NSF). This work made use of the National Magnetic Resonance Facility at Madison, which is supported by grant P41 GM103399 (NIH), and the Biophysics Instrumentation Facility at the University of Wisconsin-Madison, which was established by grants BIR-9512577 (NSF) and S10 RR013790 (NIH). High-performance computing was supported by grant CHE-0840494 (NSF). R.W.N. was supported by Biotechnology Training Grant T32 GM008349 (NIH) and by an ACS Division of Organic Chemistry Graduate Fellowship.

References

1. Dill, K.A. *Biochemistry* **29**, 7133-7155 (1990), http://dx.doi.org/10.1021/bi00483a001
2. Stickle, D.F., Presta, L.G., Dill, K.A., Rose, G.D. *J. Mol. Biol.* **226**, 1143-1159 (1992), http://dx.doi.org/10.1016/0022-2836(92)91058-W
3. Kabsch, W., Sander, C. *Biopolymers* **22**, 2577-2637 (1983), http://dx.doi.org/10.1002/bip.360221211
4. Steiner, T. *Angew. Chem. Int. Ed.* **41**, 48-76 (2002), http://dx.doi.org/10.1002/1521-3773(20020104)41:1<48::AID-ANIE48>3.0.CO;2-U
5. Bretscher, L.E., et al. *J. Am. Chem. Soc.* **123**, 777-778 (2001), http://dx.doi.org/10.1021/ja005542v
6. Newberry, R.W., VanVeller, B., Guzei, I.A., Raines, R.T. *J. Am. Chem. Soc.* **135**, 7843-7846 (2013), http://dx.doi.org/10.1021/ja4033583
7. Bartlett, G.J., Choudhary, A., Raines, R.T., Woolfson, D.N. *Nat. Chem. Biol.* **6**, 615-620 (2010), http://dx.doi.org/10.1038/nchembio.406
8. Newberry, R.W., et al. *Protein Sci.* **23**, 284-288 (2014), http://dx.doi.org/10.1002/pro.2413
9. Choudhary, A., Raines, R.T. *Protein Sci.* **20**, 1077-1081 (2011), http://dx.doi.org/10.1002/pro.627
10. Benedetti, E., et al. *Biopolymers* **27**, 357-371 (1988), http://dx.doi.org/10.1002/bip.360270302
11. Casanovas, J., et al. *J. Phys. Chem. B* **116**, 13297-13307 (2012), http://dx.doi.org/10.1021/jp3045115
12. Toniolo, C., et al. *Biopolymers* 27, 373-379 (1988), http://dx.doi.org/10.1002/bip.360270303
13. Bonora, G.M., et al. *J. Am. Chem. Soc.* **106**, 8152-8156 (1984), http://dx.doi.org/10.1021/ja00338a025
14. Formaggio, F., et al. *Org. Biomol. Chem.* **10**, 2413-2421 (2012), http://dx.doi.org/10.1039/C2OB06863J
15. Maekawa, H., Ballano, G., Toniolo, C., Ge, N.H. *J. Phys. Chem. B* **115**, 5168-5182 (2011), http://dx.doi.org/10.1021/jp105527n
16. Toniolo, C., Benedetti, E. *Crit. Rev. Biochem.* **9**, 1-44 (1980), http://dx.doi.org/10.3109/10409238009105471
17. Cochran, A.G., Skelton, N.J., Starovasnik, M.A. *Proc. Natl. Acad. Sci. USA* **98**, 5578-5583 (2001), http://dx.doi.org/10.1073/pnas.091100898
18. Deechongkit, S., Dawson, P.E., Kelly, J.W. *J. Am. Chem. Soc.* **126**, 16762-16771 (2004), http://dx.doi.org/10.1021/ja045934s
19. Deechongkit, S., et. al. *Nature* **430**, 101-105 (2004), http://dx.doi.org/10.1038/nature02611