Modulating Protein Stability by Engineering the n→π* Interactions

Daniel Gray Literature Seminar 18 November 2020

Protein primary structure encodes all the necessary chemical information to specify a particular three-dimensional structure, which then stipulates a function (1). If this chemical information were completely decoded one could predict the folding of any protein with complete accuracy. Numerous questions in biology and protein design would be answered with such an algorithm, making the protein folding challenge a prevailing issue in science today (2). Enumerating the forces affecting protein folding begins with the so-called primary (or canonical) forces: the hydrophobic effect, conventional hydrogen bonding, van der Waals interactions, and Coulombic (electrostatic) forces (3). Secondary forces include a wide array of non-covalent interactions involving both the main chain and side chain residues. These secondary forces were at times ignored in the literature even of the recent past (4). Despite their relative energetic weakness compared to primary forces, secondary forces play a key role in protein stability. It is important to recognize that proteins typically have a free energy of folding of only about 10-15 kcal/mol, while their enthalpic and entropic differences are in the range of hundreds of kcal/mol (5). This means even small energetic perturbations can have large structural effects on a folded protein. Figure 1 illustrates the energetic contributions (per 100 residues) of the most well-studied secondary forces in protein folding, with n→π* interactions being the strongest. Although individual n→π* interactions have energies of ~0.3 kcal/mol the total contribution of the n→π*

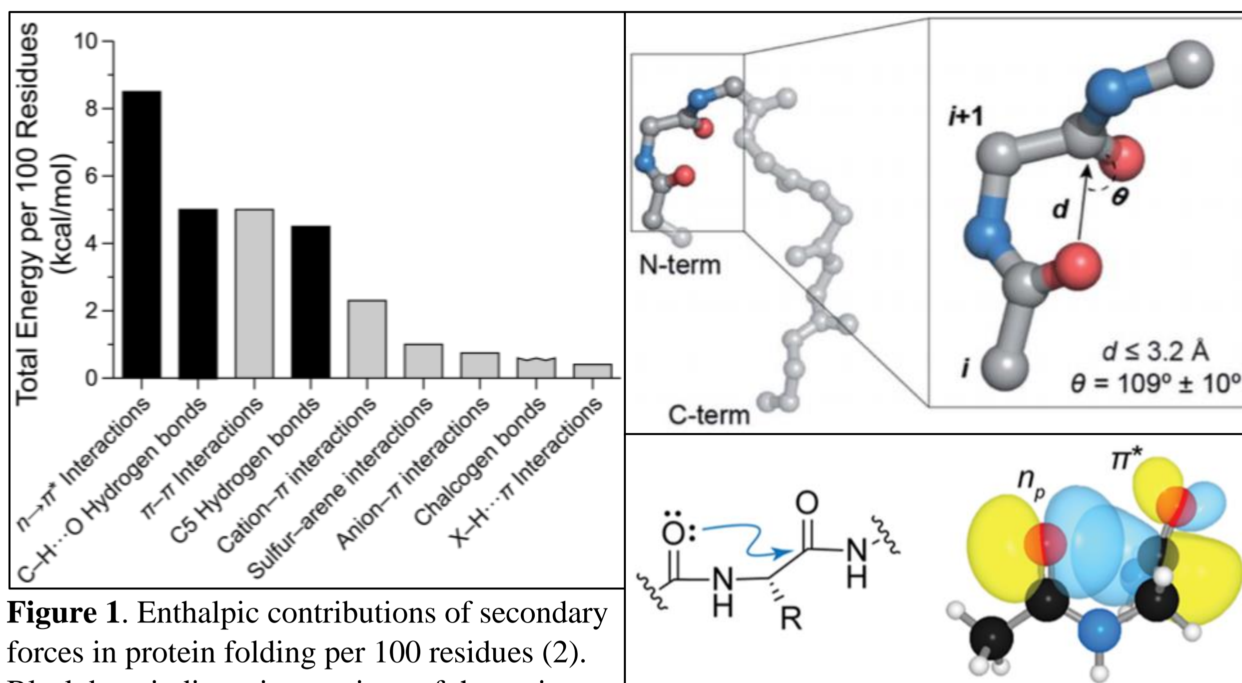

Figure 1. Enthalpic contributions of secondary forces in protein folding per 100 residues (2). Black bars indicate interactions of the main chain, while gray bars indicate side chain interactions. Note for a protein of several hundred residues, the net n→π* interactions will be greater than the free energy of folding.

Figure 2. Top: an n→π* interaction in a protein backbone with relevant distance and angular criteria (9). Bottom: orbital overlap that underlies the n→π* interaction (2).

interactions are typically exceeding 10-15 kcal/mol (2). As such, it should be clear they have a significant effect on overall protein stability.

The n→π* interaction was first invoked by Raines, *et. al.* to help justify the conformational stability of collagen (6) and has since been recognized in many other systems (7). Fundamentally, it is a nucleophilic attack on a carbonyl, so it must agree with the classic Bürgi-Dunitz trajectory for nucleophilic attack (8, figure 2). This restricts the possible ϕ and ψ backbone angles that engage in n→π* interactions which happen to align better with α-helices than β-sheets (9). In fact, it is estimated that over 70% of the residues in α-helices engage in n→π* interactions opposed to <5% of residues in β-sheets (10). The n→π* interactions can further be seen in crystal structures by the pyramidalization of the carbonyl carbon (11). For example, the crystal structure of a *trans*-polyproline II helix shows n→π* interactions play a role in helix stability which can be quantified by calculating the pyramidalization of the carbonyl carbon (figure 3, 12).

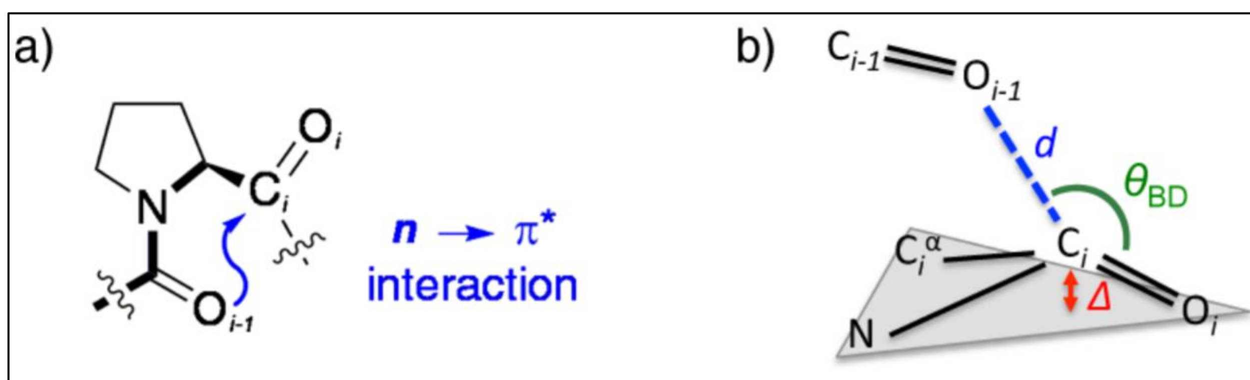

Figure 3. Diagram showing pyramidalization of carbonyl carbon and relevant parameters calculated from crystal structures. A value of Δ as high as 0.04 Å was obtained from crystal structures of *trans*-polyproline helices (12).

Experimentally testing predictions regarding n→π* interactions in native proteins remains difficult due to the weak nature of the interaction especially compared to the plethora of other, stronger interactions in the protein. However, there is disagreement over the extent of the importance of n→π* interactions. Since they are so local (between adjacent residues), they have been postulated to guide the earliest events in protein folding (2). One possible effect they have is to nucleate α-helices . These specific hypotheses have yet to be tested conclusively. However, progress has been made in understanding n→π* interactions more fully. For the first time, Khatri, *et. al.* engineered an individual n→π* interaction. They designed the solvent-exposed β-turn of GB1 via amino acid substitution, showing a relationship between the strength of the n→π* interaction as calculated by DFT from parameters obtained via the crystal structure and the stability of the overall structure as measured by thermal denaturation. Then, they showed via the same substitutions and analysis of 500 β-turn crystal structures that stronger n→π* interactions decrease the rotameric freedom of the turn (10). This is important because it shows the amino acid identity affects the ϕ and ψ angles of the backbone, which affect the strength of the n→π* interaction, which then gives a measurable effect on protein stability. Also, since the stronger n→π* interaction occurred in the α-helix region of backbone angles, they stated it agrees with the α-helix nucleation hypothesis. They furthered their study by incorporating a thioamide at the β-turn of the Pin 1 WW domain. This is to remove any doubt that the n→π* interaction is what strengthened the protein, as compared to other interactions via the side chains. The sulfur is a stronger electron-donor than

the oxygen so would strengthen the n→π* interaction (13). As such, thioamide incorporation into the Pin 1 WW domain resulted in free energy decrease using both DFT and thermal denaturation (10).

While significant progress has been made in recent years, there is still much to learn regarding n→π* interactions. One research direction includes examining this interaction in sidechains, which could be relevant in some enzyme active sites. Further increasing the capacity to engineer n→π* interactions and potentially amplify them with thioamide incorporation could allow for tuning of protein stability. This engineering can also be used to restrict the rotameric freedom of the backbone in loops. However, the end goal of n→π* interaction research is to understand its role in protein folding. Since n→π* interactions are both local and ubiquitous, they likely play a role in the early stages of protein folding. However, disentangling its role from the myriad of other relevant forces is difficult and will require some scientific ingenuity for future researchers. Lastly, while n→π* interactions are arguably the most important secondary force in protein folding, furthering our knowledge of other secondary forces relevant to protein folding will be required before a structure can be definitively predicted from a sequence. Lastly, there may be yet-undiscovered forces affecting protein stability and/or folding that will prove important to a holistic understanding of structural biology.

References:

1. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **1973** *181*, 223−230.
2. Newberry, R. W.; Raines, R. T. Secondary Forces in Protein Folding. *ACS Chem. Biol.* **2019**, *14*, 1677−1686.
3. Dill, K. A. Dominant forces in protein folding. *Biochemistry* **1990**, *29*, 7133−7155.
4. Pace, C. N., Scholtz, J. M., and Grimsley, G. R. Forces stabilizing proteins. *FEBS Lett.* **2014**, *588*, 2177−2184.
5. Richards, F. M. Protein stability: Still an unsolved problem. *Cell. Mol. Life Sci.* **1997**, *53*, 790−802.
6. Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. Conformational Stability of Collagen Relies on a Stereoelectronic Effect *J. Am. Chem. Soc.* **2001**, *123*, 777–778.
7. Singh, S. K.; Das, A. The n→π* interaction: A rapidly emerging non-covalent interaction. *Phys. Chem. Chem. Phys.* **2015**, *17*, 9596−9612.
8. Burgi, H.B.; Dunitz, J. D.; Lehn, J. M.; Wipff G. Stereochemistry of reaction paths at carbonyl centres. *Tetrahedron*, **1974**, *30*, 1563−1572.
9. Bartlett, G. J.; Choudhary, A.; Raines, R. T.; & Woolfson, D. N. n-->pi* interactions in proteins. *Nature chemical biology*, **2010**, *6*, 615–620.
10. Khatri, B.; Majumder, P.; Nagesh, J.; Penmatsa, A.; Chatterjee, J. Increasing Protein Stability by Engineering the n→π* Interaction at the β-Turn. *Chem. Sci.* **2020**, *11*, 9480–9487.
11. Choudhary, A.; and Raines, R. T. Signature of n→π* interactions in α-helices. *Protein Sci.* **2011**, *20*, 1077−1081.
12. Wilhelm, P.; Lewandowski, B.; Trapp, N.; Wennemers, H. A Crystal Structure of an Oligoproline PPII-Helix, at Last. *J. Am. Chem. Soc.* **2014**, *136*, 15829–15832.
13. Newberry, R. W.; VanVeller, B.; Guzei, I. A.; Raines, R. T. n→π* Interactions of amides and thioamides: Implications for protein stability. *J. Am. Chem. Soc.* **2013**, *135*, 7843−7846.