

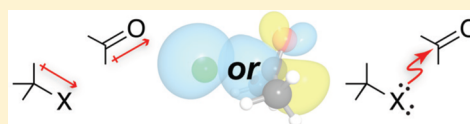
# Intimate Interactions with Carbonyl Groups: Dipole–Dipole or $n \rightarrow \pi^*$ ?

Kimberli J. Kamer,<sup>†</sup> Amit Choudhary,<sup>§,‡</sup> and Ronald T. Raines<sup>\*,†,||</sup>

<sup>†</sup>Department of Biochemistry, <sup>§</sup>Graduate Program in Biophysics, and <sup>||</sup>Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States

## Supporting Information

**ABSTRACT:** Amide carbonyl groups in proteins can engage in  $C=O \cdots C=O$  and  $C-X \cdots C=O$  interactions, where X is a halogen. The putative involvement of four poles suggests that these interactions are primarily dipolar. Our survey of crystal structures with a  $C-X \cdots C=O$  contact that is short (i.e., within the sum of the X and C van der Waals radii) revealed no preferred  $C-X \cdots C=O$  dihedral angle. Moreover, we found that structures with a short  $X^- \cdots C=O$  contact display the signatures of an  $n \rightarrow \pi^*$  interaction. We conclude that intimate interactions with carbonyl groups do not require a dipole.



## INTRODUCTION

The amide carbonyl groups in the main chain of proteins are the foci for many noncovalent interactions. For example, this main-chain carbonyl group engages in a  $C=O \cdots C=O$  interaction with another main-chain carbonyl group in common secondary structures.<sup>1</sup> Many protein–ligand complexes are stabilized by analogous  $C-X \cdots C=O$  interactions, where X is a halogen located on the ligand.<sup>2</sup> Two well-known examples include a serine protease–inhibitor complex, which is stabilized by a  $C-F \cdots C=O$  interaction,<sup>3</sup> and the histamine *N*-methyltransferase–quinacrine complex, which is stabilized by a  $C-Cl \cdots C=O$  interaction.<sup>4</sup> The possible involvement of at least four poles in these interacting pairs has led to the proposal that such interactions are primarily dipolar.<sup>2</sup> According to this argument, the functional group interacting favorably with the carbonyl group must have a dipole.

We were skeptical of this proposal for several reasons. In many interacting pairs, the van der Waals surface of the negative pole and that of the carbonyl carbon interpenetrate, forming a short contact. The ensuing orbital overlap places such interactions in the realm of quantum mechanics rather than classical electrostatics. In addition, a short contact is observed even when the interacting dipoles are not oriented favorably. For example, the two adjacent carbonyl dipoles are in a repulsive orientation in an  $\alpha$ -helix.<sup>5</sup> Yet, a short contact is observed between these carbonyl groups.<sup>1</sup>

An intimate  $C=O \cdots C=O$  interaction can be modeled as an  $n \rightarrow \pi^*$  interaction. This  $n \rightarrow \pi^*$  interaction involves delocalization of a lone pair ( $n$ ) of the donor carbonyl group into the antibonding orbital ( $\pi^*$ ) of the acceptor carbonyl group. Evidence for  $n \rightarrow \pi^*$  interactions has been detected in small molecules,<sup>6</sup> peptides,<sup>7</sup> peptoids,<sup>8</sup> proteins,<sup>1,9</sup> and nucleic acids,<sup>10</sup> and they have been postulated to stabilize transition states.<sup>11</sup> These interactions have three signatures. First, a short contact exists between the donor atom and the acceptor carbonyl carbon, allowing for orbital overlap. Second, the donor atom approaches the carbonyl group along the Bürgi–Dunitz

trajectory,<sup>12</sup> maximizing that overlap. Finally, this interaction pyramidalizes the acceptor carbonyl group.

We sought a means to determine whether a dipole was truly necessary for an intimate interaction with a carbonyl group. We reasoned that focusing on the interaction of a *monopole* with a carbonyl group could provide insight. Accordingly, we analyzed  $X^- \cdots C=O$  short contacts in detail (where  $X^-$  is a halide ion), as well as  $C-X \cdots C=O$  short contacts. The latter serve as surrogates for intimate  $C=O \cdots C=O$  interactions, as both contain two interacting dipoles and have the potential for an  $n \rightarrow \pi^*$  interaction. Herein, we report on the results of these analyses.

## RESULTS AND DISCUSSION

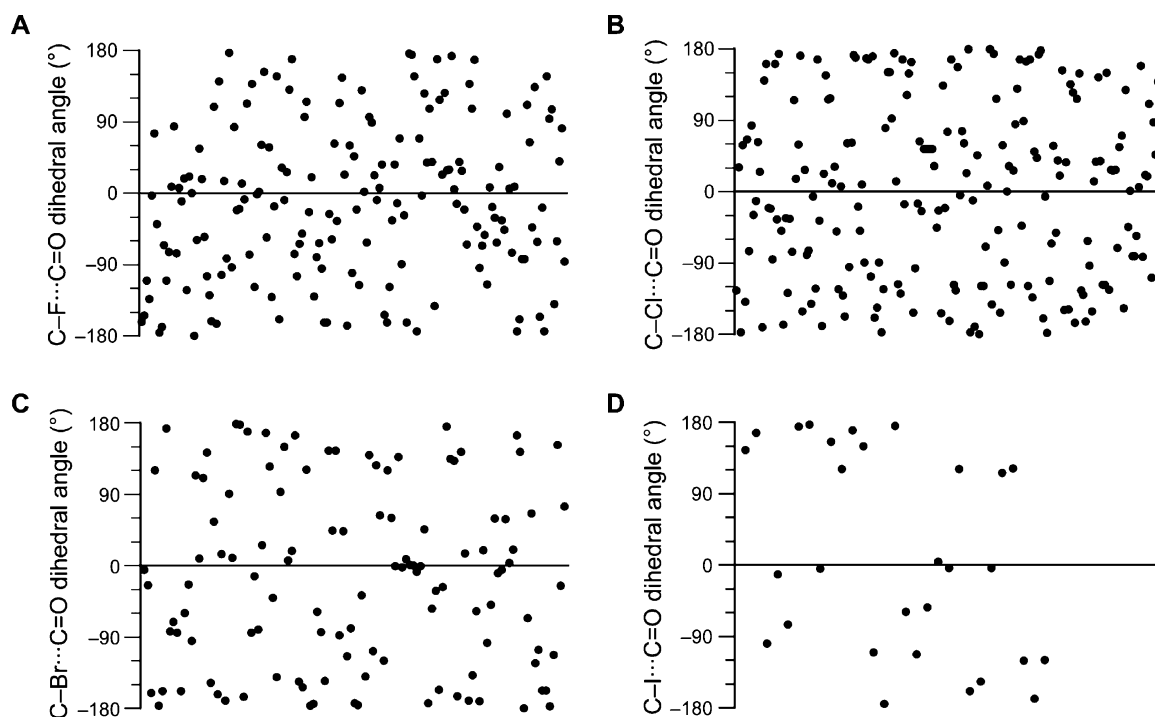
We began our analysis with a survey of  $C-X \cdots C=O$  short contacts in crystal structures of the Cambridge Structural Database (CSD). The  $n \rightarrow \pi^*$  interaction, unlike the dipolar interaction, restricts the orientation of the negative but not the positive pole of the dipole that interacts with the carbonyl group. Typically, the negative pole of the dipole interacting with the carbonyl group is located along the Bürgi–Dunitz trajectory.<sup>12,13</sup> Our analysis did not, however, reveal a preferred  $C-X \cdots C=O$  dihedral angle (Figure 1). In accord with this finding, we have observed an orientational preference in the  $O \cdots C=O$  angle, but not the  $C=O \cdots C=O$  dihedral angle, in protein secondary structures with  $C=O \cdots C=O$  interactions.<sup>71</sup> Because there is no orientational restriction on the positive pole of the dipole that interacts with the carbonyl group, this result is contrary to the expectations from a meaningful dipole–dipole interaction.

Next, we examined  $X^- \cdots C=O$  short contacts. Our premise was that the absence or presence of intimate interactions between halide ions (which are monopoles) and carbonyl

**Special Issue:** Howard Zimmerman Memorial Issue

**Received:** October 22, 2012

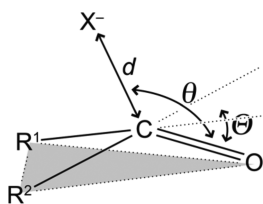
**Published:** November 19, 2012



**Figure 1.** Values of C–X...C=O dihedral angles. The abscissa is used to distribute the values randomly. (A) X = F. (B) X = Cl. (C) X = Br. (D) X = I.

groups would reveal whether a dipole is required for a favorable interaction with a carbonyl group. In addition, we reasoned that an examination of orientational restriction on the halide ions with respect to the carbonyl group in these interacting pairs would illuminate the nature of an  $X^- \cdots C=O$  interaction. Toward this end, we sought small-molecule crystal structures containing short halide–carbonyl group contacts.

For fluoride, we found four structures that met these search criteria, containing five fluoride–carbonyl short contacts. Likewise, there were 110 structures for chloride with 130 short contacts, 22 structures for bromide with 27 short contacts, and 6 structures for iodide with 7 short contacts. For each of these short contacts, we determined the distance ( $d$ ) and the angle ( $\theta$ ) of approach of the halide to the carbonyl (Figure 2).

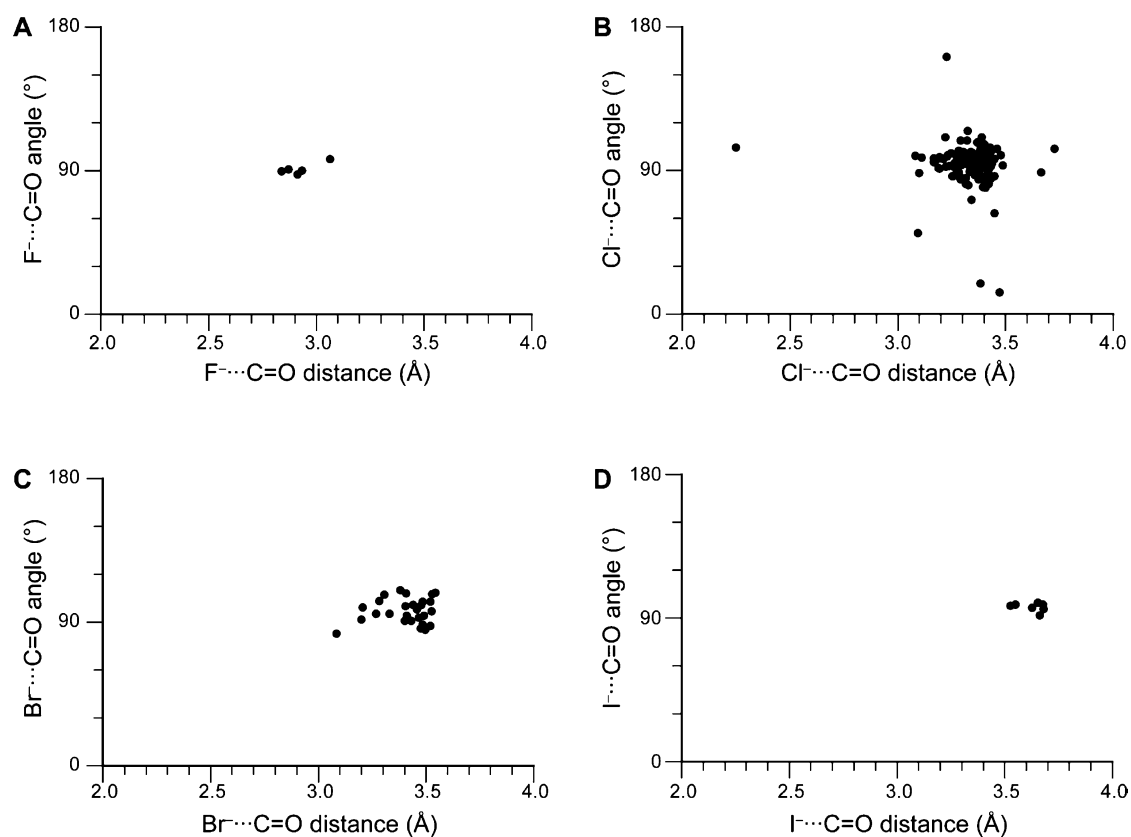


**Figure 2.** Definition of distance  $d$  and angles  $\theta$  and  $\Theta$ .

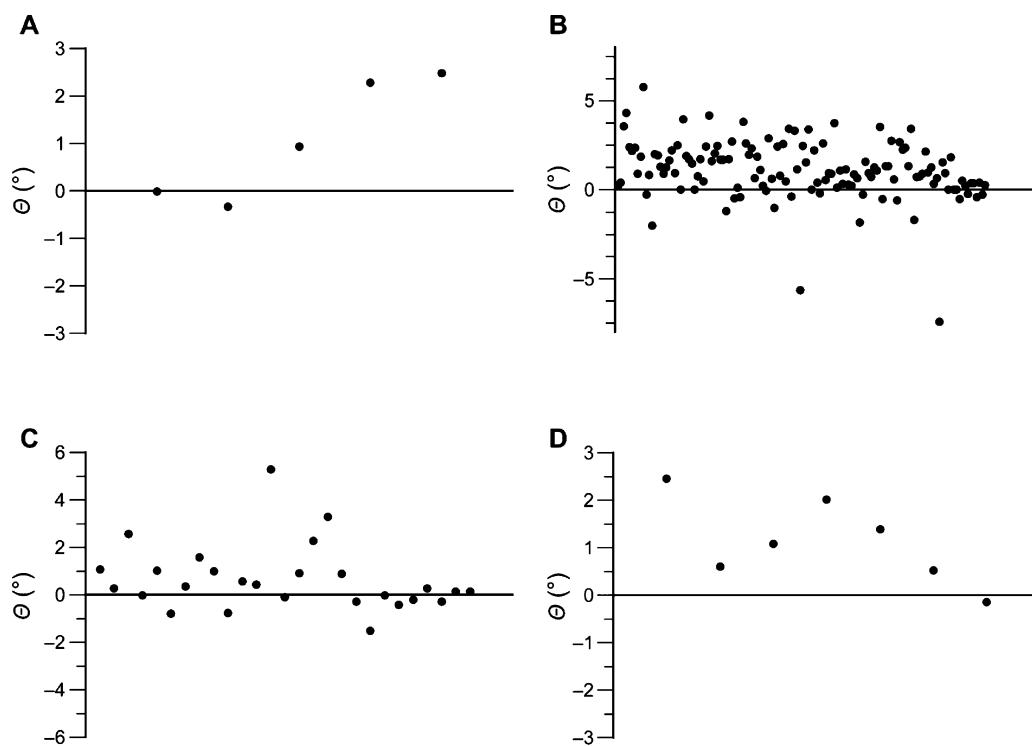
The  $X^- \cdots C=O$  angles between the halide donors and carbonyl acceptors fall mainly within the Bürgi–Dunitz trajectory ( $\theta \approx 107^\circ$ ; Figure 3). Steric effects could contribute to an orientational preference. Accordingly, we re-examined the crystal structures for carbonyl pyramidalization, which can be measured by the parameter  $\Theta$  (Figure 2). A positive  $\Theta$  value indicates that the acceptor carbonyl oxygen is displaced away from the donor halide, while a negative value indicates that the displacement is toward the halide donor. In general, we found  $\Theta$  to be  $>0$  (Figure 4), consistent with the halide donating

electron density into the carbonyl  $\pi^*$  orbital in an  $n \rightarrow \pi^*$  interaction and thus pulling the carbonyl carbon out of its canonical planar geometry and toward the halide donor. To confirm that these trends in the  $\theta$  and  $\Theta$  values are indeed meaningful, we examined “control” structures with chloride ions located at 0.90–1.00 Å beyond the sum of the van der Waals radii of carbon and chlorine. We measured the distance,  $d$ , and the angles  $\theta$  and  $\Theta$  for these long  $Cl^- \cdots C=O$  interactions. The values of  $d$  and  $\theta$  indicate that the chloride ions are not necessarily located along the Bürgi–Dunitz trajectory (Figure 5A); moreover, the values of  $\Theta$  are not indicative of carbonyl group pyramidalization toward the chloride ions (Figure 5B). We note that the location and effect of the halide could be altered in some structures by the accompanying cation (which maintains electrical neutrality). Moreover, crystal packing forces could have some influence on our observed geometrical preferences.<sup>14</sup> We did not observe a strong correlation between  $d$  and  $\Theta$ . The degree of pyramidalization,  $\Theta$ , is a function of many variables besides  $X^- \cdots C=O$  distance. For example, the degree of pyramidalization depends on the Bürgi–Dunitz angle ( $\theta$ ) and the elasticity of the acceptor carbonyl group. For the same  $X^- \cdots C=O$  distance, the degree of pyramidalization for an amide carbonyl will be different from that of a ketone. The structures examined in our search vary not only in the nature of the carbonyl group but also in the  $\theta$  angle. Hence, we should not expect a strong correlation between the degree of pyramidalization and the  $X^- \cdots C=O$  distance.

Finally, we note that an  $n \rightarrow \pi^*$  interaction should not be confused<sup>15</sup> with an  $n \rightarrow \pi^*$  electronic transition, which is distinct electronically and energetically (Figure 6). An  $n \rightarrow \pi^*$  electronic transition within a carbonyl group refers to the excitation of its nonbonded electron ( $n$ ) to its  $\pi^*$  orbital. An  $n \rightarrow \pi^*$  interaction, on the other hand, refers to electron delocalization between a donor atom and an acceptor carbonyl group. Moreover, an

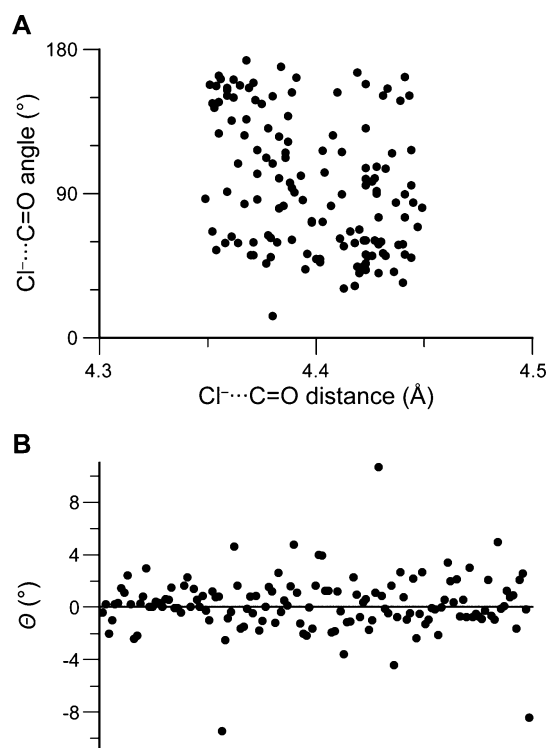


**Figure 3.** Plots of  $X \cdots C=O$  distance ( $d$ ) and angle ( $\theta$ ). (A)  $X = F$ . (B)  $X = Cl$ . (C)  $X = Br$ . (D)  $X = I$ .

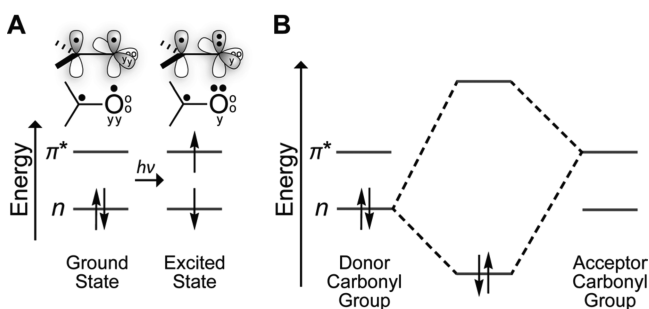


**Figure 4.** Values of carbonyl group pyramidalization induced by a sub-van der Waals  $X \cdots C=O$  interaction. The abscissa is used to distribute the values randomly. (A)  $X = F$ . (B)  $X = Cl$ . (C)  $X = Br$ . (D)  $X = I$ .

$n \rightarrow \pi^*$  electronic transition requires the addition of energy to the system, whereas an  $n \rightarrow \pi^*$  interaction releases energy from the system.



**Figure 5.** Parameters for structures with a long  $\text{Cl}\cdots\text{C}=\text{O}$  interaction. (A) Plot of  $\text{Cl}\cdots\text{C}=\text{O}$  distance ( $d$ ) and angle ( $\theta$ ). (B) Values of carbonyl group pyramidalization. The abscissa is used to distribute the values randomly.



**Figure 6.** Orbital energy diagrams for two distinct electronic phenomena of carbonyl groups. (A)  $n \rightarrow \pi^*$  electronic transition of a single carbonyl group. The inset depicts the Zimmerman "circle-dot-y" notation for the ground state and excited state in which "o" represents an s-rich electron, "•" represents a  $\pi$  electron, and "y" represents a p-rich electron.<sup>16</sup> (B)  $n \rightarrow \pi^*$  interaction between two carbonyl groups.

## CONCLUSION

We have argued previously that intimate interactions between carbonyl groups with interpenetrating van der Waals surfaces involve significant delocalization of an electron pair of the oxygen of the donor carbonyl group into the antibonding orbital ( $\pi^*$ ) of the acceptor carbonyl group. Our current findings support earlier ones<sup>1,6–11</sup> and suggest that these intimate interactions involving carbonyl groups do not require a dipole.

## EXPERIMENTAL SECTION

ConQuest and Mercury software were used to search CSD version 5.31 (updated November 2009) for structures with a sub-van der Waals contact between a halide ion and a carbonyl group.<sup>13c</sup> Specifically, we searched for structures where the distance ( $d$  in

Figure 2) between the halide and the carbonyl carbon was  $\leq 3.17$  Å,  $\leq 3.45$  Å,  $\leq 3.55$  Å, and  $\leq 3.68$  Å for fluoride, chloride, bromide, and iodide, respectively. All of these structures have  $R \leq 5\%$ , and none have any errors or disorder. In addition, none were polymeric, and none were organometallic or powder structures.

## ASSOCIATED CONTENT

### Supporting Information

Data used to construct Figures 1 and 3–5. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [rtraines@wisc.edu](mailto:rtraines@wisc.edu).

### Present Address

‡Society of Fellows, Harvard University, 78 Mount Auburn Street, Cambridge, Massachusetts 02138, United States.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank G. J. Bartlett, C. N. Bradford, B. R. Caes, G. R. Krow, M. D. Shoulders, B. VanVeller, and D. N. Woolfson for assistance and contributive discussions. This work was supported by Grants R01 AR044276 (NIH) and CHE-1124944 (NSF). K.J.K. was supported by a Mary Shine Peterson Award from the Department of Biochemistry.

## DEDICATION

This paper is dedicated to the memory of our colleague, Howard E. Zimmerman (1926–2012).

## REFERENCES

- (1) Bartlett, G. A.; Choudhary, A.; Raines, R. T.; Woolfson, D. N. *Nat. Chem. Biol.* **2010**, *6*, 615–620.
- (2) (a) Olsen, J. A.; Banner, D. W.; Seiler, P.; Obst-Sander, U.; D'Arcy, A.; Stihle, M.; Müller, K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 2507–2511. (b) Olsen, J. A.; Banner, D. W.; Seiler, P.; Wagner, B.; Tschopp, T.; Obst-Sander, U.; Kansy, M.; Müller, K.; Diederich, F. *ChemBioChem* **2004**, *5*, 666–675. (c) Hof, F.; Scofield, D. M.; Schweizer, W. B.; Diederich, F. *Angew. Chem., Int. Ed.* **2004**, *43*, 5056–5059. (d) Paulini, R.; Müller, K.; Diederich, F. *Angew. Chem., Int. Ed.* **2005**, *44*, 1788–1805. (e) Fischer, F. R.; Schweizer, W. B.; Diederich, F. *Angew. Chem., Int. Ed.* **2007**, *46*, 8270–8273. (f) Bissantz, C.; Kuhn, B.; Stahl, M. *J. Med. Chem.* **2010**, *53*, 5061–5084.
- (3) Adler, M.; Davey, D. D.; Phillips, G. B.; Kim, S. H.; Jancarik, J.; Rumennik, G.; Light, D. R.; Whitlow, M. *Biochemistry* **2000**, *39*, 12534–12542.
- (4) Horton, J. R.; Sawada, K.; Nishibori, M.; Zhang, X.; Cheng, X. *Structure* **2001**, *9*, 837–849.
- (5) (a) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1991**, *113*, 6256–6265. (b) Yang, A.-S.; Honig, B. *J. Mol. Biol.* **1995**, *252*, 351–365. (c) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry, Part III: The Behavior of Biological Macromolecules*; Freeman and Company: New York, NY, 2004.
- (6) (a) Lesarri, A.; Cocinero, E. J.; López, J. C.; Alonso, J. L. *J. Am. Chem. Soc.* **2005**, *127*, 2572–2579. (b) Blanco, S.; López, J. C.; Mata, S.; Alonso, J. L. *Angew. Chem., Int. Ed.* **2010**, *49*, 9187–9192. (c) Choudhary, A.; Kamer, K. J.; Raines, R. T. *J. Org. Chem.* **2011**, *76*, 7933–7937. (d) Cabezas, C.; Alonso, J. L.; López, J. C.; Mata, S. *Angew. Chem., Int. Ed.* **2012**, *51*, 1375–1378.
- (7) (a) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777–778. (b) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.;

- Raines, R. T.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497–2505.
- (c) Hinderaker, M. P.; Raines, R. T. *Protein Sci.* **2003**, *12*, 1188–1194.
- (d) Jenkins, C. L.; Lin, G.; Duo, J.; Rapolu, D.; Guzei, I. A.; Raines, R. T.; Krow, G. R. *J. Org. Chem.* **2004**, *69*, 8565–8573. (e) Horng, J.-C.; Raines, R. T. *Protein Sci.* **2006**, *15*, 74–83. (f) Hodges, J. A.; Raines, R. T. *Org. Lett.* **2006**, *8*, 4695–4697. (g) Sonntag, L.-S.; Schweizer, S.; Ochsenfeld, C.; Wennemers, H. *J. Am. Chem. Soc.* **2006**, *128*, 14697–14703. (h) Kümin, M.; Sonntag, L.-S.; Wennemers, H. *J. Am. Chem. Soc.* **2007**, *129*, 466–467. (i) Shoulders, M. D.; Guzei, I. A.; Raines, R. T. *Biopolymers* **2008**, *89*, 443–454. (j) Dai, N.; Wang, X. J.; Etzkorn, F. A. *J. Am. Chem. Soc.* **2008**, *130*, 5396–5397. (k) Chiang, Y.-C.; Lin, Y.-J.; Horng, J.-C. *Protein Sci.* **2009**, *18*, 1967–1977. (l) Choudhary, A.; Gandla, D.; Krow, G. R.; Raines, R. T. *J. Am. Chem. Soc.* **2009**, *131*, 7244–7246. (m) Dai, N.; Etzkorn, F. A. *J. Am. Chem. Soc.* **2009**, *131*, 13728–13732. (n) Kuemin, M.; Schweizer, S.; Ochsenfeld, C.; Wennemers, H. *J. Am. Chem. Soc.* **2009**, *131*, 15474–15482. (o) Choudhary, A.; Fry, C. G.; Raines, R. T. *ARKIVOC* **2010**, *8*, 251–262. (p) Kuemin, M.; Nagel, Y. A.; Schweizer, S.; Monnard, F. W.; Ochsenfeld, C.; Wennemers, H. *Angew. Chem., Int. Ed.* **2010**, *49*, 6324–6327. (q) Jakobsche, C. E.; Choudhary, A.; Miller, S. J.; Raines, R. T. *J. Am. Chem. Soc.* **2010**, *132*, 6651–6653. (r) Choudhary, A.; Pua, K. H.; Raines, R. T. *Amino Acids* **2011**, *41*, 181–186. (s) Choudhary, A.; Raines, R. T. *Protein Sci.* **2011**, *20*, 1077–1081. (t) Pollock, S. B.; Kent, S. B. H. *Chem. Commun.* **2011**, *47*, 2342–2344. (u) Erdmann, R. S.; Wennemers, H. *J. Am. Chem. Soc.* **2012**, *134*, 17117–17124.
- (8) (a) Gorske, B. C.; Bastian, B. L.; Geske, G. D.; Blackwell, H. E. *J. Am. Chem. Soc.* **2007**, *129*, 8928–8929. (b) Gorske, B. C.; Stringer, J. R.; Bastian, B. L.; Fowler, S. A.; Blackwell, H. E. *J. Am. Chem. Soc.* **2009**, *131*, 16555–16567. (c) Caumes, C.; Roy, O.; Faure, S.; Taillefumier, C. *J. Am. Chem. Soc.* **2012**, *134*, 9553–9556.
- (9) (a) Gao, J.; Kelly, J. W. *Protein Sci.* **2008**, *17*, 1096–1101. (b) Fufezan, C. *Proteins* **2010**, *78*, 2831–2838. (c) Choudhary, A.; Kamer, K. J.; Raines, R. T. *Protein Sci.* **2012**, *21*, 171–177.
- (10) Choudhary, A.; Kamer, K. J.; Powner, M. W.; Sutherland, J. D.; Raines, R. T. *ACS Chem. Biol.* **2010**, *5*, 655–657.
- (11) (a) Liu, P.; Yang, X.; Birman, V. B.; Houk, K. N. *Org. Lett.* **2012**, *14*, 3288–3291. (b) Wang, H.; Kohler, P.; Overman, L. E.; Houk, K. N. *J. Am. Chem. Soc.* **2012**, *134*, 16054–16058.
- (12) Bürgi, H. B.; Dunitz, J. D.; Shefter, E. *Acta Crystallogr.* **1974**, *B30*, 1517–1527.
- (13) (a) Allen, F. H. *Acta Crystallogr.* **2002**, *B58*, 380–388. (b) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. *Acta Crystallogr.* **2002**, *B58*, 389–397. (c) Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor, R.; Towler, M.; van de Streek, J. *J. Appl. Crystallogr.* **2006**, *39*, 453–457.
- (14) (a) Dauber, P.; Hagler, A. T. *Acc. Chem. Res.* **1980**, *13*, 105–112. (b) Dickerson, R. E.; Goodsell, D. S.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3579–3583. (c) Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *48*–76.
- (15) Worley, B.; Richard, G.; Harbison, G. S.; Powers, R. *PLoS ONE* **2012**, *7*, e42075.
- (16) (a) Zimmerman, H. E. *Abstracts of the 17th National Organic Symposium of the American Chemical Society, Bloomington, IN; American Chemical Society: Washington, DC, 1961; pp 31–41.* (b) Zimmerman, H. E. In *Advances in Photochemistry*; Noyes, W. A., Hammond, G. S., Pitts, J. N., Eds.; John Wiley: New York, NY, 1963; Vol. 1, pp 183–208. (c) Zimmerman, H. E. *Pure Appl. Chem.* **2006**, *78*, 2193–2203.