

## Evidence of an Unusual N–H···N Hydrogen Bond in Proteins

Ramkrishna Adhikary,<sup>†</sup> Jörg Zimmermann,<sup>†</sup> Jian Liu,<sup>†</sup> Ryan P. Forrest,<sup>‡</sup> Tesia D. Janicki,<sup>‡</sup> Philip E. Dawson,<sup>†</sup> Steven A. Corcelli,<sup>‡</sup> and Floyd E. Romesberg<sup>\*,†</sup>

<sup>†</sup>Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

<sup>‡</sup>Department of Chemistry & Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States

### S Supporting Information

**ABSTRACT:** Many residues within proteins adopt conformations that appear to be stabilized by interactions between an amide N–H and the amide N of the previous residue. To explore whether these interactions constitute hydrogen bonds, we characterized the IR stretching frequencies of deuterated variants of proline and the corresponding carbamate, as well as the four proline residues of an Src homology 3 domain protein. The C<sub>δ</sub>D<sub>2</sub> stretching frequencies are shifted to lower energies due to hyperconjugation with N<sub>i</sub> electron density, and engaging this density via protonation or the formation of the N<sub>i+1</sub>–H···N<sub>i</sub> interaction ablates this hyperconjugation and thus induces an otherwise difficult to explain blue shift in the C–D absorptions. Along with density functional theory calculations, the data reveal that the N<sub>i+1</sub>–H···N<sub>i</sub> interactions constitute H-bonds and suggest that they may play an important and previously underappreciated role in protein folding, structure, and function.

Protein structure underlies function, and the overall structure of a protein is determined by the path traced out by the backbone. The contribution of individual residues is conveniently described by the torsion angles around the N–C<sub>α</sub> bond ( $\phi$ ) and the C<sub>α</sub>–C bond ( $\psi$ ), which collectively define the allowed and disallowed regions of the Ramachandran plot.<sup>1</sup> While simple steric arguments adequately describe most of the conformations adopted, a complete understanding remains elusive. For example, conformations that fall within the so-called “bridge region,” where  $\phi$  and  $\psi$  are centered around  $-90^\circ$  and  $0^\circ$ , respectively, are generally considered unfavorable due to a steric repulsion between the amide nitrogen (N<sub>i</sub>) and the proton of the next amide (N<sub>i+1</sub>–H).<sup>1,2</sup> However, the distance between the nitrogen and the proton is close to the normal contact distance,<sup>1</sup> and the proton appears positioned to interact maximally with the N<sub>i</sub> electron density. Moreover, the specific  $\phi$ ,  $\psi$ , and N–C<sub>α</sub>–C angles observed appear to be correlated in a manner that preserves this interaction.<sup>3,4</sup> Interestingly, the N<sub>i+1</sub>–H···N<sub>i</sub> interaction typically appears to be part of a bifurcated (three-centered) hydrogen bond (H-bond)<sup>5</sup> involving the amide oxygen of another residue or water molecule, in addition to the N<sub>i</sub> atom, as H-bond acceptors. Nonetheless, the amide nitrogen is a poor H-bond acceptor, and the potential contribution of these interactions to protein structure and stability has largely been ignored.<sup>6</sup>

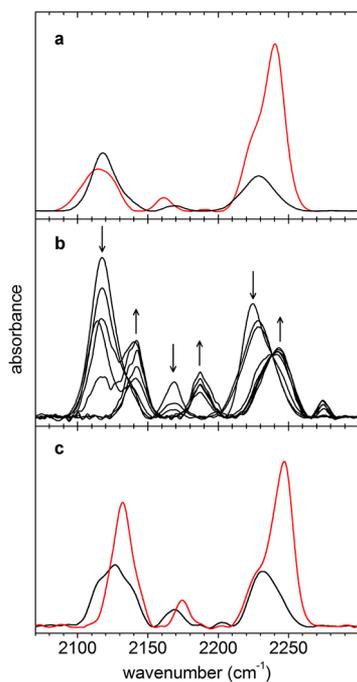
In addition to structural data, evidence for H-bond formation has typically come from the observation of a red shift of the donor heavy atom-hydrogen stretching vibration.<sup>5</sup> Unfortunately, the spectroscopic characterization of an individual N–H bond in a protein is challenging due to spectral congestion. However, individual carbon–deuterium (C–D) bonds are observable within a protein,<sup>7–13</sup> and in some cases they may serve as probes of H-bonding at adjacent atoms. This is due to two factors: first, the well established hyperconjugation between p- or  $\pi$ -orbital electron density and suitably oriented  $\sigma^*$  orbitals of adjacent C–H/D bonds, which causes the C–H/D bonds to have lower frequency IR stretching absorptions (often referred to as Bohlmann bands);<sup>14–16</sup> and second, the reduction of hyperconjugation associated with H-bond formation and the consequent blue shift of the C–H/D absorption.<sup>17–19</sup> Thus, we reasoned that if the N<sub>i+1</sub>–H···N<sub>i</sub> interactions constitute an H-bond, then they might induce blue shifts of the absorptions of suitably positioned C–D bonds.

To explore these ideas, we chose to examine the amino acid proline, which commonly adopts bridge-like conformations in proteins due to restrictions of its  $\phi$  values,<sup>3</sup> and which provides C<sub>α</sub> methine and C<sub>δ</sub> methylene groups that are directly bound to the amide nitrogen, as well as C<sub>β</sub> and C<sub>γ</sub> methylene groups that act as controls by providing “unperturbed” absorptions. Because Bohlmann bands have been characterized most extensively with amines, we first characterized the free amino acid. At pH 13, where the nitrogen is deprotonated, the spectrum of the d<sub>7</sub> labeled proline (fully deuterated at C<sub>α</sub>, C<sub>β</sub>, C<sub>γ</sub>, and C<sub>δ</sub>) consists of three features, a weaker band at 2161 cm<sup>-1</sup>, and two, more intense absorptions around 2112 and 2241 cm<sup>-1</sup> (Figure 1a). Based on previous studies,<sup>20</sup> the weak band is assigned as the C<sub>α</sub>D stretching absorption, while the more intense low and high frequency bands are assigned to overlapping symmetric and asymmetric absorptions, respectively, of the methylene groups. Within both the symmetric and asymmetric bands, overlapping absorptions are apparent, and comparison with the spectrum of the proline variant labeled only at the C<sub>α</sub> and C<sub>δ</sub> positions ((d<sub>3</sub>)proline) both confirms the C<sub>α</sub>D assignment and allows assignment of the lower frequency symmetric and asymmetric absorptions to the C<sub>δ</sub>D<sub>2</sub> group. Consistent with previous studies,<sup>20,21</sup> we ascribe the red shift of the C<sub>δ</sub>D<sub>2</sub> absorptions to hyperconjugation between the  $\sigma^*$  C–D orbitals and the adjacent nitrogen lone pair electrons. To characterize the effects of protonation (as a mimic of H-bond

Received: March 27, 2014

Published: September 16, 2014



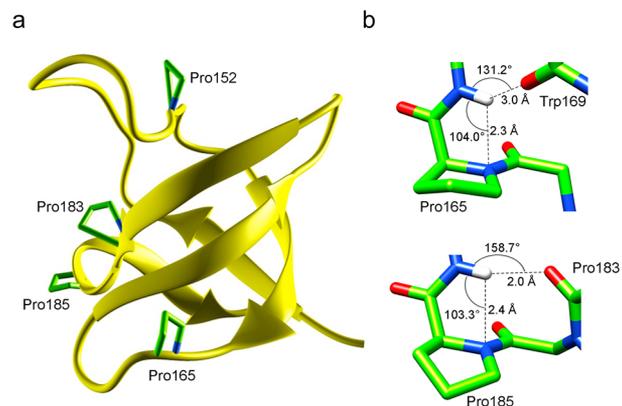


**Figure 1.** (a) IR spectra of ( $d_7$ ) and ( $d_3$ ) proline at pH 13, shown in red and black, respectively. (b) Overlay of ( $d_3$ ) proline spectra as a function of pH, with arrows indicating growth or disappearance of individual absorptions with decreasing pH. All spectra were normalized at  $2235\text{ cm}^{-1}$ . (c) IR spectra of ( $d_7$ ) and ( $d_3$ ) proline carbamate at pH 7.3, shown in red and black, respectively.

formation), we characterized the ( $d_3$ )proline absorptions as a function of pH (Figure 1b). As the pH is decreased, the lower frequency  $C_\delta D_2$  stretches are replaced with blue-shifted  $C_\delta D_2$  stretches, and plotting the amplitudes reveals a simple two-state transition with a midpoint pH of  $\sim 10$  (Supporting Information). Interestingly, the  $C_\alpha D$  absorption also shows a similar pH-dependent blue shift (Supporting Information). We conclude that protonation of the amine reduces hyperconjugation with the  $C_\delta-D$   $\sigma^*$  orbitals, resulting in the expected blue shift.

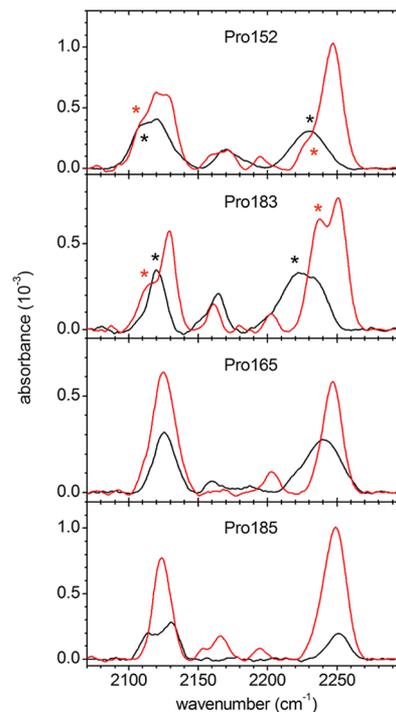
To explore the spectra when the C–D bonds are adjacent to a nitrogen involved in conjugation, we next characterized the IR spectra of both ( $d_3$ ) and ( $d_7$ )proline *tert*-butyl carbamate (Figure 1c). The  $d_7$  spectrum is similar to that of the free amino acid and indicates the presence of a  $C_\alpha D$  absorption and two symmetric and two asymmetric stretching bands, with the  $d_3$  spectrum again confirming the assignments and allowing assignment of the lower energy symmetric and asymmetric absorptions to the  $C_\delta D_2$  group. Thus, we conclude that the  $C_\delta D_2$  group absorbs at lower frequencies due to hyperconjugation between the  $\sigma^*$  orbital of one or both  $C_\delta D$  bonds and the  $\pi$ -electron density of the carbamate.

To characterize  $N_{i+1}-H\cdots N_i$  interactions in a protein, we turned to the N-terminal Src homology 3 (nSH3) domain from the murine Crk-II adaptor protein (Figure 2a).<sup>22</sup> We selected this protein because it has four Pro residues, two of which engage in  $N_{i+1}-H\cdots N_i$  interactions and two of which do not (Figure 2b and Supporting Information). Using Boc-solid phase peptide synthesis, we synthesized each variant wherein one of the Pro residues was replaced with ( $d_7$ ) or ( $d_3$ )Pro (Supporting Information).



**Figure 2.** (a) Structure of nSH3 (PDB ID: 1CKA). Side chains of deuterated Pro residues are shown in green. (b) Structure of putative three-center  $N_{i+1}-H\cdots N_i$  H-bonds at Pro165 and Pro185 with bond lengths and angles indicated.

Pro152 is part of the conserved RT loop, Pro183 is at the end of the fourth strand of the protein's single  $\beta$ -sheet, and the crystal structure reveals that neither participate in  $N_{i+1}-H\cdots N_i$  interactions.<sup>22</sup> The C–D stretching region of ( $d_7$ )Pro152 and ( $d_7$ )Pro183 are similar to that of the small molecule model systems (Figure 3). One exception is the appearance of a weak



**Figure 3.** IR spectra of site-specifically labeled nSH3. The  $d_7$  labeled variants are shown in red, and the  $d_3$  variants in black. The red-shifted absorptions assigned as  $C_\delta D_2$  Bohlmann bands are indicated with stars.

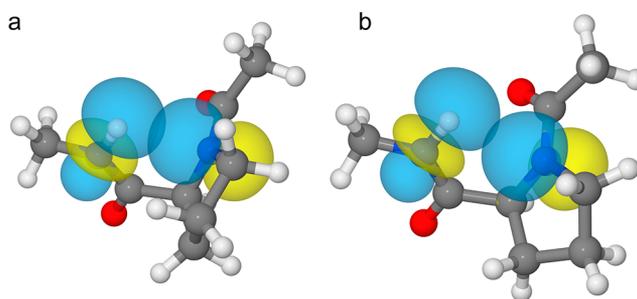
absorption around  $2200\text{ cm}^{-1}$ , which may result from backbone heterogeneity or a Fermi resonance. As with the model systems, both the symmetric and asymmetric bands of ( $d_7$ )Pro152 and ( $d_7$ )Pro183 show overlapping absorptions (although the lower frequency asymmetric absorption is more intense at Pro183), and comparison with the  $d_3$  proline spectra again clearly allows for the assignment of the lower energy absorptions to the  $C_\delta D_2$  stretches. The lower energy absorptions of the  $C_\delta D_2$  moieties,

relative to the  $C_{\beta}D_2$  and  $C_{\gamma}D_2$  absorptions, indicate that, as with the small molecule model systems, significant electron density is available at the peptide amide nitrogen for hyperconjugation with adjacent C–D bonds.

Pro165 is the second residue of a type VIII turn within the n-Src loop, and Pro185 forms part of a type I turn within a  $3_{10}$ -helix. The crystal structure reveals that both engage in  $N_{i+1}-H\cdots N_i$  interactions,<sup>22</sup> with Pro165 forming a bifurcated H-bond with the backbone carbonyl oxygen of Trp169 and Pro185 with the backbone carbonyl oxygen of Pro183 (Figure 2b). The IR spectra of both ( $d_7$ )Pro165 and ( $d_7$ )Pro185 show a single symmetric and a single asymmetric absorption around 2125 and 2247  $\text{cm}^{-1}$ , respectively (as well as weaker absorptions around 2165  $\text{cm}^{-1}$  that are again assigned to  $C_{\alpha}D$  stretches) (Figure 3). This clearly reveals that the symmetric and asymmetric  $C_{\delta}D_2$  absorptions are not differentiated from those of the  $C_{\beta}D_2$  and  $C_{\gamma}D_2$  absorptions, which is confirmed by the ( $d_3$ )Pro165 and ( $d_3$ )Pro185 spectra.

Based on the deconvoluted asymmetric absorptions of the  $d_7$ -labeled proteins (Supporting Information), the frequencies of the  $C_{\beta}D_2$  and  $C_{\gamma}D_2$  absorptions at the different Pro residues are all within 5  $\text{cm}^{-1}$ . In contrast, relative to ( $d_3$ )Pro152 and ( $d_3$ )Pro183, the  $C_{\delta}D_2$  absorptions of ( $d_3$ )Pro165 are blue-shifted by 10 and 17  $\text{cm}^{-1}$ , and those of ( $d_3$ )Pro185 are blue-shifted by 21 and 28  $\text{cm}^{-1}$ . While a similar analysis for the symmetric absorptions is not straightforward due to the more complicated band structure, the same trend is again apparent, with the  $C_{\delta}D_2$  absorptions at Pro165 and Pro185 uniquely blue-shifted relative to those at Pro152 and Pro183 as well as relative to the  $C_{\beta}D_2$  or  $C_{\gamma}D_2$  absorptions. The observed blue shifts are unlikely to result from the conformation of the pyrrolidine ring, as the conformation and shift are not correlated (the ring conformation is  $C_{\gamma}$ -exo in Pro152 and Pro185, and  $C_{\gamma}$ -endo in Pro165 and Pro183). In addition, the blue shift does not appear to be correlated with solvent exposure (Pro183 is buried, while Pro152, Pro165, and Pro185 are solvent exposed). Moreover, based on the invariance of the proximal  $C_{\beta}D_2$  and  $C_{\gamma}D_2$  frequencies, the blue shifts do not result from local electrostatics. In contrast, the blue shift is correlated with the presence of the  $N_{i+1}-H\cdots N_i$  interaction. Moreover, the blue shift results in  $C_{\delta}D_2$  absorptions that are indistinguishable from the  $C_{\beta}D_2$  and  $C_{\gamma}D_2$  absorptions, consistent with it originating from the ablation of the perturbation caused by hyperconjugation, just as was observed with protonation of free proline. Similar blue shifts have been observed with C–H stretching absorptions in small molecules when H-bonds with adjacent p or  $\pi$  electron donors are introduced.<sup>17–19</sup> Thus, the observed blue shifts constitute evidence that the  $N_{i+1}-H\cdots N_i$  interactions observed in the structure of nSH3 are indeed H-bonds.

To support the hypothesis that these interactions constitute H-bonds, we examined methyl-terminated proline dipeptide mimics with conformations corresponding to those observed for Pro165 and Pro185 via density functional theory (DFT) as well as natural bond orbital (NBO) analysis,<sup>23,24</sup> which converts the full electron density from the DFT calculation into a set of localized natural atomic and bonding orbitals. The NBO  $N_{i+1}-H\sigma^*$  and  $N_i$  orbitals are shown in Figure 4, where clear overlap, indicative of  $n \rightarrow \sigma^*$  charge transfer is apparent and supports the hypothesis that the  $N_{i+1}-H\cdots N_i$  interactions indeed constitute H-bonds. The calculated stabilization energies due to the  $n \rightarrow \sigma^*$  charge transfer are 0.6 and 0.1 kcal/mol for Pro165 and Pro185, respectively. While these values are likely



**Figure 4.** NBO  $N_{i+1}-H\sigma^*$  and  $N_i$  orbitals showing overlap indicative of  $n \rightarrow \sigma^*$  charge transfer for (a) Pro165 and (b) Pro185.

underestimated due to the effects of other electrostatic and dispersive contributions, as well as the effects of the protein environment (the charge transfer results in increased charge separation, and the calculations were run in the gas phase), they suggest that the interactions are stabilizing.

While the results provide strong evidence that the  $N_{i+1}-H\cdots N_i$  interactions constitute H-bonds, a more quantitative analysis of the strength of the H-bonds awaits further analysis, such as more realistic calculations and the deconvolution of the effects of the hyperconjugation from other factors such as structure-dependent coupling between the C–D bonds. However, both the observation that proteins appear to structurally adjust to accommodate such H-bonds<sup>3,4</sup> and the observation of similar H-bonds in small molecule peptide mimics<sup>6,25–28</sup> suggest that they can be stabilizing. It is also interesting to note that many  $\alpha$ -helix and turn residues adopt conformations near the bridge region with apparent  $N_{i+1}-H\cdots N_i$  H-bonds, suggesting that they are common. Regardless of their effects on stability, such H-bonds are likely to increase the single bond character of the N–C bond and thus increase backbone flexibility. Indeed, the formation of identical  $N_{i+1}-H\cdots N_i$  H-bonds are thought to underlie some forms of catalyzed *cis*–*trans* proline isomerization<sup>6,29,30</sup> and similar side-chain-mediated amide H-bonds have been suggested to facilitate “intramolecular catalysis” during the folding of dihydrofolate reductase.<sup>31</sup> Thus,  $N_{i+1}-H\cdots N_i$  H-bonds may make important but previously unrecognized contributions not only to stability and secondary structure formation, but also to the folding, dynamics, and function of proteins in general.

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

Methods, characterization of synthesized proteins, data analysis, additional tables and figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

floyd@scripps.edu

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This research was supported by the National Science Foundation Grants MCB-0346967 (F.E.R.) and CHE-0845736 (S.A.C.).

## ■ REFERENCES

- (1) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283–438.
- (2) Swindells, M. B.; MacArthur, M. W.; Thornton, J. M. *Nat. Struct. Biol.* **1995**, *2*, 596–603.
- (3) Karplus, P. A. *Protein Sci.* **1996**, *5*, 1406–1420.
- (4) Scarsdale, J. N.; Van Alsenoy, C.; Klimkowski, V. J.; Schaefer, L.; Momany, F. A. *J. Am. Chem. Soc.* **1983**, *105*, 3438–3445.
- (5) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: New York, NY, 1997.
- (6) Cox, C.; Lectka, T. *J. Am. Chem. Soc.* **1998**, *120*, 10660–10668.
- (7) Adhikary, R.; Zimmermann, J.; Liu, J.; Dawson, P. E.; Romesberg, F. E. *J. Phys. Chem. B* **2013**, *117*, 13082–13089.
- (8) Chin, J. K.; Jimenez, R.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 2426–2427.
- (9) Cremeens, M. E.; Zimmermann, J.; Yu, W.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2009**, *131*, 5726–5727.
- (10) Groff, D.; Thielges, M. C.; Cellitti, S.; Schultz, P. G.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2009**, *48*, 3478–3481.
- (11) Sagle, L. B.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 3384–3385.
- (12) Weinkam, P.; Zimmermann, J.; Sagle, L. B.; Matsuda, S.; Dawson, P. E.; Wolyne, P. G.; Romesberg, F. E. *Biochemistry* **2008**, *47*, 13470–13480.
- (13) Zimmermann, J.; Thielges, M. C.; Seo, Y. J.; Dawson, P. E.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2011**, *50*, 8333–8337.
- (14) Bykov, S. V.; Myshakina, N. S.; Asher, S. A. *J. Phys. Chem. B* **2008**, *112*, 5803–5812.
- (15) Alabugin, I. V.; Gilmore, K. M.; Peterson, P. W. *WIREs Comput. Mol. Sci.* **2011**, *1*, 109–141.
- (16) Wolfe, S.; Schlegel, H. B.; Whangbo, M.-H.; Bernardi, F. *Can. J. Chem.* **1974**, *52*, 3787–3792.
- (17) Chandra, A. K.; Parveen, S.; Das, S.; Zeegers-Huyskens, T. *J. Comput. Chem.* **2008**, *29*, 1490–1496.
- (18) Xu, Z.; Li, H.; Wang, C.; Wu, T.; Han, S. *Chem. Phys. Lett.* **2004**, *394*, 405–409.
- (19) Xu, Z.; Li, H.; Wang, C.; Pan, H.; Han, S. *J. Chem. Phys.* **2006**, *124*, 244502.
- (20) Sheena Mary, Y.; Ushakumari, L.; Harikumar, B.; Varghese, H. T.; Panicker, C. Y. *J. Iran. Chem. Soc.* **2009**, *6*, 138–144.
- (21) Krueger, P. J.; Jan, J. *Can. J. Chem.* **1970**, *48*, 3229–3235.
- (22) Wu, X.; Knudsen, B.; Feller, S. M.; Zheng, J.; Sali, A.; Cowburn, D.; Hanafusa, H.; Kuriyan, J. *Structure* **1995**, *3*, 215–226.
- (23) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.
- (24) Bartlett, G. J.; Choudhary, A.; Raines, R. T.; Woolfson, D. N. *Nat. Chem. Biol.* **2010**, *6*, 615–620.
- (25) Gieren, A.; Dederer, B.; George, G.; Marquarding, D.; Ugi, I. *Tetrahedron Lett.* **1977**, *18*, 1503–1506.
- (26) Gieren, A.; Dederer, B.; Schanda, F. *Z. Naturforsch. C* **1980**, *35*, 741–746.
- (27) Hossain, M. B.; Van der Helm, D. *J. Am. Chem. Soc.* **1978**, *100*, 5191–5198.
- (28) Cieplak, A. S. *J. Am. Chem. Soc.* **1985**, *107*, 271–273.
- (29) Fischer, S.; Dunbrack, R. L.; Karplus, M. *J. Am. Chem. Soc.* **1994**, *116*, 11931–11937.
- (30) Cox, C.; Lectka, T. *Org. Lett.* **1999**, *1*, 749–752.
- (31) Texter, F. L.; Spencer, D. B.; Rosenstein, R.; Matthews, C. R. *Biochemistry* **1992**, *31*, 5687–5691.