VOLUME 122, NUMBER 34
AUGUST 30, 2000

JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Identification of an Essential Backbone Amide Bond in the Folding and Stability of a Multimeric Enzyme

Bassam M. Nakhle, Peter Silinski, and Michael C. Fitzgerald*

Contribution from the Department of Chemistry, Duke University, Durham, North Carolina 27708

Received February 21, 2000

Abstract: Here we utilize a total chemical synthesis strategy and a mass spectrometry-based, combinatorial chemistry approach to identify key molecular interactions that contribute to the folding and stability of a model multimeric enzyme, 4-oxalocrotonate tautomerase (4OT). 4OT is a 41 kDa bacterial enzyme composed of six identical 62 amino acid subunits. A total of 16 different 4OT analogues containing various natural and unnatural N-terminal modifications were prepared by total chemical synthesis and then characterized by catalytic activity, size exclusion chromatography (SEC), and circular dichroism (CD) spectroscopy. The results of our mutational studies indicate that backbone−backbone hydrogen-bonding interactions involving the amide bond between Pro1 and Ile2 in 4OT's 62 amino acid polypeptide chain play an important role in specifying the conformation of this enzyme's native folded state. These results provide the first evidence that backbone−backbone hydrogen-bonding interactions can play such a major role in specifying the native state of a protein.

Introduction

Most mutational studies of protein folding and stability have focused on the role of amino acid side chains and not on the role of the peptide backbone. This is largely because amino acid side chains are more amenable to mutation than the peptide backbone in conventional recombinant DNA-based, site-directed mutagenesis experiments.[1] Recently, two new technologies (one relying on chemical synthesis strategies[2] and one relying on a specialized in vitro translation system[3]) have emerged that permit modification of a protein's peptide backbone. Here we utilize a total chemical synthesis strategy and a mass spectrometry-based, combinatorial chemistry approach to identify amino acid side chains and elements of the peptide backbone that contribute to the folding and stability of a model multimeric enzyme, 4-oxalocrotonate tautomerase (4OT).

4OT is a 41 kDa, hexameric bacterial enzyme composed of six identical 62 amino acid subunits.[4,5] The enzyme is expressed in certain soil bacteria where it catalyzes an isomerization reaction in which β,γ-unsaturated keto acids are converted to their α,β-unsaturated isomers.[6] The catalytic properties of 4OT have been well-studied over the past decade.[6−13] However,

* Corresponding author. Tel: 919-660-1547. Fax: 919-660-1605. E-mail: mfitz@chem.duke.edu.

(1) Smith, M. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1214−1221.
(2) (a) Muir, T. W.; Kent, S. B. H. *Curr. Opin. Biotechnol.* **1993**, *4*, 420−427 (b) Lu, W.; Qasim, M. A.; Laskowski, M.; Kent, S. B. H. *Biochemistry* **1997**, *36*, 673−678.
(3) (a) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314−11322. (b) Chapman, E.; Thorson, J. S.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 7151−7152.
(4) Chen, L. H.; Kenyon, G. L.; Curtin, F.; Harayama, S.; Bembenek, M. E.; Hajipour, G.; Whitman, C. P. *J. Biol. Chem.* **1992**, *267*, 17716−17721.
(5) Subramanya, H. S.; Roper, D. I.; Dauter, Z.; Dodson, E. J.; Davies, G. J.; Wilson, K. S.; Wigley, D. B. *Biochemistry* **1996**, *35*, 792−802.
(6) Whitman, C. P.; Aird, B. A.; Gillespie, W. R.; Stolowich, N. J. *J. Am. Chem. Soc.* **1991**, *113*, 3154−3162.
(7) Whitman, C. P.; Hajipour, G.; Watson, R. J.; Johnson, W. H.; Bembenek, M. E.; Stolwich, N. J. *J. Am. Chem. Soc.* **1992**, *114*, 10104−10110.
(8) Lian, H.; Whitman, C. P. *J. Am. Chem. Soc.* **1993**, *115*, 7978−7984.
(9) Fitzgerald, M. C.; Chernushivech, I.; Standing, K. G.; Kent, S. B. H.; Whitman, C. P. *J. Am. Chem. Soc.* **1995**, *117*, 11075−11080.
(10) Fitzgerald, M. C.; Chernushevich, I.; Standing, K. G.; Whitman, C. P.; Kent, S. B. H. *Proc. Natl. Acad. Sci., USA* **1996**, *93*, 6851−6856.

10.1021/ja0006090 CCC: $19.00
Published on Web 08/11/2000

relatively little is known about the specific noncovalent, inter- and intrasubunit interactions that define this enzyme's homohexameric structure. In earlier studies of 4OT's catalytic properties a number of enzyme analogues were prepared in which Pro1 was replaced with a series of different acyclic amino acids. These constructs which included both coded (e.g. Gly, Ala)[13] and noncoded (e.g. *N*-methylglycine, *N*-ethylglycine, *N*-propylglycine, dimethyl glycine, and aminobutyric acid)[14] amino acid substitutions for Pro1 displayed varying amounts of catalytic activity. However, they all folded into homohexameric structures very similar to that of the wild-type enzyme as judged by far-UV circular dichroism (far-UV−CD) spectroscopy and size exclusion chromatography (SEC). Surprisingly, a truncated construct of the enzyme in which Pro1 was deleted from the enzyme's polypeptide chain *did not* form a native hexameric complex as judged by SEC.[10] This N-terminally truncated 4OT mutant also contained very little secondary structure as judged by far-UV−CD spectroscopy.[10] These observations prompted us to further investigate the molecular contribution that the first residue in 4OT's primary amino acid sequence makes to the enzyme's folding and stability.

Here we report the results of mutational studies on 4OT's N-terminus. A total of 16 different analogues containing various natural and unnatural N-terminal modifications were prepared by total chemical synthesis in order to investigate the role of Pro1 in 4OT's folding and stability. The results of our mutational studies indicate that backbone−backbone hydrogen-bonding interactions involving the amide bond between Pro1 and Ile2 in 4OT's 62 amino acid polypeptide chain play an important role in specifying the conformation of this enzyme's native folded state. Our results provide the first evidence that backbone−backbone hydrogen-bonding interactions can play a major role in specifying the native state of a protein.

Results and Discusion

Combinatorial Synthesis and Characterization of 4OT Analogues. We initially prepared three small combinatorial libraries of 4OT analogues, which included analogues **1**−**11** in Scheme 1. Library I contained wild-type 4OT and three analogues of the enzyme with one, two, and three amino acid residues deleted from the N-terminus. Library II was similar to library I with two exceptions: (1) the N-terminal residues of the truncated analogues in library I were mutated to proline, leaving a secondary amine instead of a primary amine at 4OT's N-terminus, and (2) an additional analogue, **4**, was included in order to determine what effect an N-terminal extension of 4OT's polypeptide chain would have on the enzyme's folding and stability. Library III contained the same analogues as in library I except that the N-termini of these analogues were acetylated. This acetyl group precluded the accumulation of charge (e.g. protonation) at the N-termini of these mutants during the folding reaction. It also added an additional N-terminal amide bond to the peptide backbones of the deletion analogues in library I. For example, N-terminal acetylation of **1** (to give **9**) restored the first amide bond in 4OT's covalent structure.

Scheme 1

	Residue 1	Residue 2	Residue 3	Residue 4	Residue 39	Analogue #	SEC Results
Library I	Pro	Ile	Ala	Gln	Arg〰	Wt	9 min Peak
		Ile	Ala	Gln	Arg〰	1	> 11 min
			Ala	Gln	Arg〰	2	> 11 min
				Gln	Arg〰	3	> 11 min
Library II	Pro	Ala	Ile	Ala	Gln Arg〰	4	9 min Peak
	Pro	Ile	Ala	Gln	Arg〰	Wt	9 min Peak
		Pro	Ala	Gln	Arg〰	5	> 11 min
			Pro	Gln	Arg〰	6	> 11 min
				Pro	Arg〰	7	> 11 min
Library III	Ac	Pro	Ile	Ala	Gln Arg〰	8	9 min Peak
	Ac	Ile	Ala	Gln	Arg〰	9	9 min Peak
		Ac	Ala	Gln	Arg〰	10	> 11 min
			Ac	Gln	Arg〰	11	> 11 min
		OIle	Ala	Gln	Arg〰	12	12 min Peak
	Valr		Ala	Gln	Arg〰	13	12 min Peak
	Pro	OIle	Ala	Gln	Arg〰	14	9/12 min Peak
	Pro	Ile	Ala	Gln	OAla〰	15	9/12 min Peak
	Pro	OIle	Ala	Gln	OAla〰	16	9/12 min Peak

Abbreviations: Ac = Acetylated N-terminus; Wt = Wild-type; Valr = Valeric acid ($CH_3(CH_2)_3CO$-); OIle = α-hydroxy acid of isoleucine (L-2-hydroxy-3-methyl-valeric acid); OAla = α-hydroxy acid of alanine (L-lactic acid).

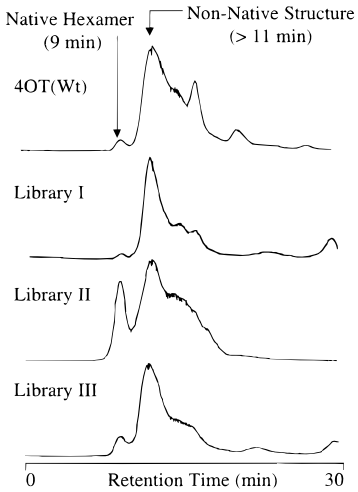

Figure 1. SEC chromatograms of wild-type 4OT and libraries I−III. The wild-type enzyme and the analogue libraries were prepared by total chemical synthesis. SEC analysis of the protein constructs were performed after the unpurified synthetic material from each synthesis was subjected to a folding protocol.

The unpurified, synthetic products obtained from each library synthesis were subjected to a folding protocol and then analyzed by SEC. The SEC chromatograms obtained for libraries I−III are shown in Figure 1. Analogues with an elution time in the SEC experiment consistent with the elution time of the native hexamer (9 min) were identified by their mass using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. This was possible because the analogues in each library had unique masses. MALDI mass spectra of the material in libraries I−III obtained before and after hexamer isolation by SEC are shown in Figure 2. The folding selection and the MALDI analysis employed in these combinatorial studies identified polypeptide chains capable of forming stable, nativelike mul-

(11) Stivers, J. T.; Abeygunawadana, C.; Mildvan, A. S.; Hajipour, G.; Whitman, C. P.; Chen, L. H. *Biochemistry* **1996**, *35*, 803−813.
(12) Czerwinski, R. M.; Johnson, W. H., Jr.; Whitman, C. P.; Harris, T. K.; Abeygunawardana, C.; Mildvan, A. S. *Biochemistry* **1997**, *36*, 14551−14560.
(13) Czerwinski, R. M.; Harris, T. K.; Johnson, W. H., Jr.; Legler, P. M.; Stivers, J. T.; Mildvan, A. S.; Whitman, C. P. *Biochemistry* **1999**, *38*, 12358−12366.
(14) Fitzgerald, M. C.; Czerwinski, R. M.; Whitman, C. P. Unpublished Results.

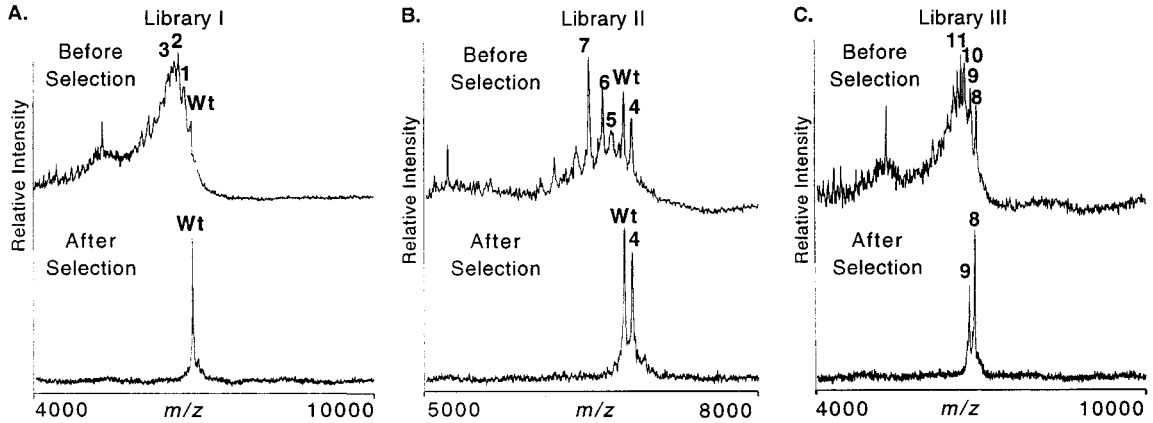

Figure 2. MALDI-MS analysis of libraries I (A), II (B), and III (C) before and after hexamer selection by SEC.

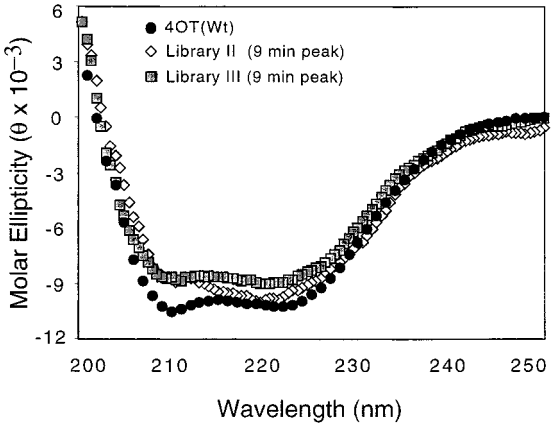

Figure 3. Far-UV−CD spectra of a wild-type 4OT control and material isolated from the 9 min peaks in the SEC elution profiles of libraries II and III.

timeric complexes. The wild-type 62 amino acid polypeptide chain of 4OT served as a positive control in our combinatorial experiments: it was efficiently selected in our folding selections of libraries I and II. Analogue **1**, a 4OT construct that displayed no higher order structure in other studies,[10] served as a negative control in our combinatorial experiments: it was not selected in our folding selection of library II.

The results of our SEC and MALDI analysis of libraries I−III indicate that none of the deletion analogues, **1**−**3**, **5**−**7**, **10** or **11**, were competent to form stable, nativelike multimeric complexes. The only 4OT analogues in these three libraries capable of forming nativelike hexameric structures were **4**, **8**, and **9**. The structural properties of these folding analogues were further characterized by far-UV−CD spectroscopy. The far-UV−CD spectra of material isolated from the 9 min peaks in the SEC chromatograms of libraries II and III were very similar to a spectrum of the wild-type enzyme (see Figure 3). These results indicate that the secondary structures adopted by analogues **4**, **8**, and **9** in our folding selection were not substantially different from that of the wild-type enzyme.

Polarity Effects at 4OT's N-Terminus. Analogue **9** was present in our hexamer selection of library III, and **1** was not present in our hexamer selection of library I. These two analogues differ by only a few backbone atoms at the N-termini of their polypeptide chains. Initially, we reasoned that differences in the folding behavior of **1** and **9** might be due to differences in the polarity of the functional groups at the N-termini of their polypeptide chains. The primary amino group at **1**'s N-terminus is significantly more polar than the acetamide group at **9**'s N-terminus. In an attempt to restore a nativelike

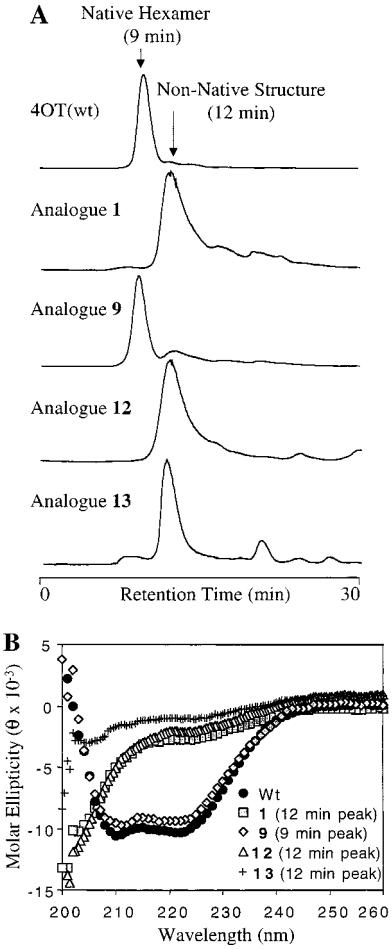

Figure 4. (A) SEC chromatogram obtained for a wild-type 4OT control and analogues **1**, **9**, **12**, and **13**. (B) Far-UV−CD spectra of a wild-type 4OT control and analogues **1**, **9**, **12**, and **13**.

hexamer fold to **1**, we altered the polarity of this analogue's N-terminus. We prepared two analogues, **12** and **13**, that were very similar in structure to **1**, except that the functional groups at their N-termini (HO and $CH_3(CH_2)_3C$, respectively) were less polar than the primary amino group at **1**'s N-terminus.

The polypeptide chains of **1**, **9**, **12**, and **13** were individually prepared by total chemical synthesis, purified by reversed-phase HPLC, and then characterized by SEC and by far-UV−CD spectroscopy. The SEC elution profiles of **1**, **9**, **12**, and **13** after they were subjected to identical folding protocols are shown in Figure 4A. A major peak at 9 min corresponding to nativelike hexamer was only detected in the SEC analysis of **9**. A minor

signal at ~9 min is discernible in the SEC elution profile of **13**; however, we attribute this signal to chemical and/or instrumental noise, since no peptide signal was observed in the MALDI analysis of material isolated from this peak. The results of our SEC analyses of **1**, **9**, **12**, and **13** indicated that **9** was the only one of these analogues to fold into a nativelike hexamer structure (see Figure 4A). As expected, analogue **9** displayed no enzymatic activity because it lacked a catalytically significant N-terminal amine functionality. However, a far-UV−CD spectrum of material isolated from the 9 min hexamer peak in the SEC chromatogram of **9** was similar to a far-UV−CD spectrum recorded for the wild-type enzyme (see Figure 4B). Moreover, the SEC and far-UV−CD data we collected on this reversed-phase purified version of **9** was consistent with our data collected on **9** when it was characterized in the library III mixture, as described above. It is noteworthy that in both cases the molar ellipticity values we recorded for the nativelike hexamer complexes of **9** were approximately 10% lower than those recorded for the wild-type enzyme. This small spectral difference is likely due to a minor variation in the secondary structural features of **9** as compared to the wild-type enzyme.

In contrast to our results with **9**, there were no nativelike hexamer structures detected in the SEC analysis of the final folded states of **1**, **12**, and **13** (see Figure 4A). The SEC elution profiles for these analogues were very similar. A major peak at approximately 12 min appeared in the SEC elution profile of each analogue. Unfortunately, it is not possible to precisely characterize the oligomeric state of material in these 12 min peaks solely on the basis of their retention time in these SEC experiments because of 4OT's aberrant behavior on conventional sizing columns.[4,10] Our efforts to characterize the oligomerization state of analogues **1**, **12**, and **13** using analytical ultracentrifugation methods have also met with limited success. The lack of aromatic residues in 4OT's primary amino acid sequence and the decreased solubility of these analogues made their detection difficult in the analytical ultracentrifugation experiment. However, it is clear from the SEC elution profiles obtained for analogues **1**, **12**, and **13** that the oligomerization states of these analogues are similar and that they are significantly reduced in size from the hexameric state of the wild-type enzyme.

Far-UV−CD spectra collected on analogues **1**, **12**, and **13** were also consistent with these analogues only partially folding to a non-native state(s) (see Figure 4B). The amount of structural information that can be extracted from far-UV−CD spectra is limited; however, it is noteworthy that the far-UV−CD spectra recorded for **1**, **12**, and **13** are not indicative of polypeptide chains with a completely random coil conformation.[15] The final folded states of analogues **1** and **12** appear to be similar on the basis of the spectral features in their far-UV−CD spectra. The far-UV−CD spectra obtained for these analogues are similar to spectra obtained for polypeptide chains containing some helical and random coil content.[15] It is also possible that the partially folded states of analogues **1**, **12**, and **13** are not composed of a single, unique conformation. There may be multiple protein conformations contributing to the observed CD signal. This could possibly explain the far-UV−CD results with **13**.

Our SEC and far-UV−CD results with **1**, **12**, and **13** show that altering the polarity of **1**'s N-terminus does not restore a nativelike fold to this analogue. These results suggest that the hydrophobic character of 4OT's N-terminal residue is not an essential determinant of the enzyme's folding and stability.

(15) Greenfield, N.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108−4116.

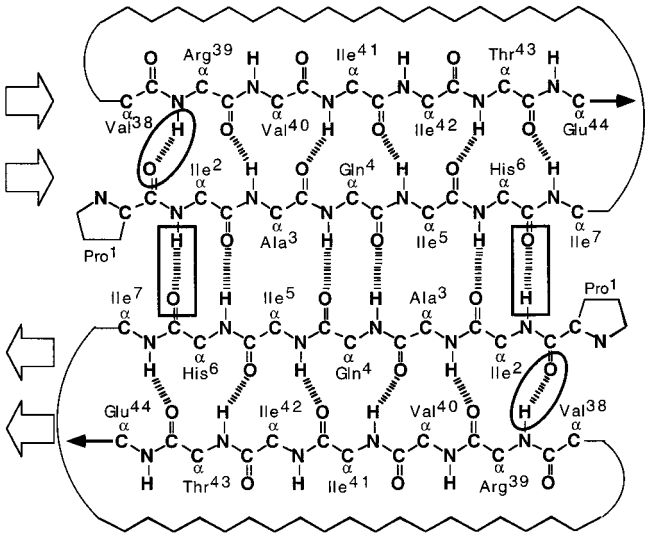

Figure 5. Schematic representation of the putative backbone−backbone hydrogen-bonding interactions involving 4OT's N-terminus. The inter- and intrasubunit hydrogen bonds involving the N-terminal amide bond between Pro1 and Ile2 are highlighted in the two subunits shown here with rectangles and ovals, respectively.

Backbone−Backbone Hydrogen Bonds at 4OT's N-Terminus. Next, we reasoned that the nativelike folding of **9** might be due to the additional N-terminal amide bond in **9** and that the non-native folding of **1** might be due to the lack of this amide bond in **1**. An inspection of the X-ray crystallographic data available for 4OT indicates that in the enzyme's native folded state this N-terminal amide bond is involved in both inter- and intrasubunit interactions involving putative backbone−backbone hydrogen bonds (Figure 5).[5] To evaluate the role of this N-terminal amide bond and these backbone−backbone hydrogen-bonding interactions in 4OT's folding and stability, we investigated the folding behavior of three ester bond containing analogues of 4OT, **14**−**16**. The ester linkage is a useful amide bond isostere that can be used to modulate the hydrogen-bonding properties of the peptide bond.[2b,3] Like the amide bond, it favors the *trans* configuration, and there is a significant energy barrier to *cis*−*trans* isomerization.[16] However, the ester bond substitution effectively deletes the hydrogen-bond-donating properties of the peptide bond and decreases the basicity of the carbonyl oxygen.[17,18]

Analogues **14** and **15** contained a single amide (−CONH−) to ester (−COO−) bond mutation in their polypeptide backbones between residues 1 and 2 and between residues 38 and 39, respectively. The single ester bond mutations in these analogues effectively deleted one of the two putative backbone−backbone hydrogen bonds involving the N-terminal amide bond between Pro1 and Ile2 in 4OT's polypeptide chain. Both of the putative backbone−backbone hydrogen bonds involving 4OT's N-terminal amide bond were simultaneously deleted in analogue **16** by incorporating two −CONH− to −COO− mutations into 4OT's polypeptide backbone (one between residues 1 and 2 and one between residues 38 and 39). Analogues **15** and **16** also contained an arginine to alanine mutation at position 39 in 4OT's polypeptide chain. Incorporating the ester bond between residues 38 and 39 was synthetically more convenient with this mutation due to the commercial availability of the α-hydroxy acid of alanine. In studies of 4OT's catalytic properties it was established that a (R39A)4OT construct had similar folding

(16) Wiberg, K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1987**, *109*, 5935−5943.

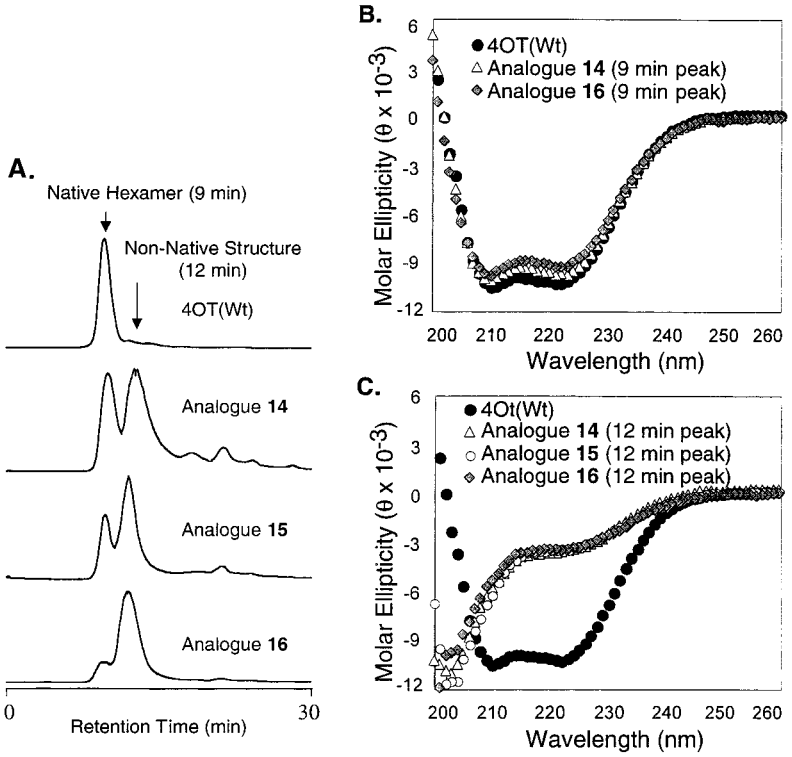

Figure 6. (A) SEC elution profiles obtained for a synthetic, wild-type construct of 4OT and three ester bond containing analogues of the enzyme, **14**−**16**. (B) Far-UV−CD analysis of material isolated from 9 min peaks in the SEC elution profiles of **14** and **16**. The spectra are very similar to that obtained for wild-type 4OT. (C) Far-UV−CD analysis of material isolated from 12 min peaks in SEC elution profiles of **14** and **16**. The spectra are significantly different from that obtained for wild-type 4OT.

properties to the wild-type enzyme, despite its catalytic deficiencies, which included a 100-fold reduction in k_{cat}/K_m.[13]

Analogues **14**−**16** were subjected to identical folding protocols and then analyzed by SEC. The SEC elution profiles obtained for these ester bond containing analogues were significantly different from a SEC elution profile of the wild-type enzyme (see Figure 6A). A single peak at approximately 9 min was observed in the SEC chromatogram of a synthetic wild-type construct of the enzyme; whereas, two distinct peaks (one at approximately 9 min and one at approximately 12 min) were observed in the SEC chromatograms of **14**−**16**. One concern in our studies of **14**−**16** was the lability of the ester bonds in these analogues. Characterization of these analogues by reversed-phase high-performance liquid chromatography (RP-HPLC) and electrospray ionization mass spectrometry (ESI-MS) prior to their folding confirmed the initial integrity of these analogues in our experiments (see the Supporting Information). Subsequent ESI-MS analyses of the material isolated in the SEC experiments were used to check the integrity of these analogues after their folding (data not shown). No hydrolysis was detected during the folding and SEC analysis of **15**, and less than 25% of the material used in the folding protocol for **14** and **16** was hydrolyzed during the time course of folding and SEC analysis. The ester bond between Pro1 and Ile2 was most susceptible to hydrolysis in **14** and **16**. However, the 61 amino acid peptide product of this hydrolysis reaction appeared only in material isolated from the 12 min SEC peak. There was no evidence of hydrolyzed material in the 9 min peak, even 24 h after its isolation. These observations are consistent with material from the 12 min peak being in a partially unfolded, non-native state in which the ester bond between Pro1 and Ile2 is more solvent exposed than it is in the enzyme's native state.

The material that eluted at 9 and 12 min in the SEC chromatograms of **14**−**16** was isolated and further characterized by its catalytic activity and by far-UV−CD spectroscopy. Material from the 12 min peaks of these ester bond-containing analogues was not catalytically active. The far-UV−CD spectra that were recorded on material from these 12 min peaks were also substantially different from a spectrum recorded for the wild-type enzyme (i.e. molar ellipticity values at 222 nm were approximately 40% of those observed for the wild-type enzyme) (see Figure 6C). It is also noteworthy that the far-UV−CD spectra obtained for material in the 12 min peaks of **14**−**16** were very similar to those recorded for analogues **12** and **13** (see Figure 4B), each of which also eluted at 12 min in the SEC experiment.

Attempts to isolate material from the 9 min peaks in the SEC chromatograms of **14**−**16** were only successful in the case of **14** and **16**. It was not possible to obtain material exclusively from the 9 min SEC peak of **15**. Once it was isolated, material from the 9 min peak of **15** rapidly equilibrated to a mixture of species with a SEC elution profile equivalent to that initially obtained with this analogue (i.e., a 9 and 12 min peak were observed). Our observation that material in the 9 and 12 min peaks interconvert after purification for **15**, but not for **14** and **16**, suggests that the amide bond between Pro1 and Ile2 may be critical for the folding pathway, rather than (or in addition to) the stability of the folded enzyme.

The material that we isolated in the 9 min SEC peaks of **14** and **16** was further characterized by far-UV−CD spectroscopy and by its catalytic activity. Far-UV−CD spectra recorded for material isolated from the 9 min SEC peaks of **14** and **16** were nearly identical to a UV−CD spectrum of the wild-type hexamer (Figure 6B). The catalytic efficiencies, k_{cat}/K_M values, that we calculated for material isolated in these 9 min peaks were 5.4×10^6 and 8.5×10^5 M^{-1} s^{-1} for **14** and **16**, respectively. These catalytic efficiency measurements constitute a sensitive structural probe of the nativelike, hexamer material in these 9 min peaks.

Structural and mechanistic studies have previously established that the side chains of Pro1 and Arg39 play a major role in 4OT catalysis; however, there is no evidence to suggest that backbone atoms involving Pro1 or Arg39 are involved in the catalytic reaction. Therefore, the backbone modifications in **14** and **16** should not influence the catalytic properties of 4OT unless these modifications significantly perturb the structure of the enzyme's active site in which they are positioned. The catalytic efficiency of **14** can be compared to that of the wild-type enzyme which has a reported k_{cat}/K_M of 1.9×10^7 M^{-1} s^{-1}, and the catalytic efficiency of **16** can be compared to that of (R39A)4OT construct of the enzyme which has a reported k_{cat}/K_M of 9.7×10^4 M^{-1} s^{-1}. Our catalytic data on **14** and **16** were not dramatically different from data obtained on the appropriate amide bond containing controls (the wild-type enzyme and an (R39A)4OT mutant, respectively). Our catalytic activity measurements, CD analysis, and SEC experiments on **14** and **16** indicate that the material in the 9 min peak of the SEC elution profile of these analogues is folded to a very similar hexameric state as that of the wild-type enzyme.

The relative peak areas of the 9 and 12 min peaks in the SEC elution profiles of **14**, **15**, and **16** were approximately 1:1, 1:2, and 1:10, respectively. We attribute these relative peak areas to the relative magnitude of the destabilizing effects of the hydrogen-bond deletions in these analogues. The equilibrium between the nativelike hexameric state in the 9 min peak and the partially folded non-native state(s) in the 12 min peak is almost completely shifted in favor of the partially folded non-native state(s) in the case of analogue **16**. Our detection of a minor 9 min peak in the SEC analysis of **16** suggests that hydrogen bonds involving the amide bond between Pro1 and Ile2 do not constitute the only stabilizing molecular interactions at 4OT's N-terminus. However, our qualitative analysis of **16** and the other 4OT mutants in this study suggest that these hydrogen-bonding interactions do play a major role in stabilizing the native hexameric structure of 4OT. Future biophysical studies comparing the thermodynamic properties of wild-type 4OT to the thermodynamic properties of our folding mutants will provide a more quantitative assessment of the relative energetic contributions of these backbone−backbone hydrogen bonds in 4OT's folding and stability.

Conclusion. The results of our mutational studies on 4OT indicate that the non-native folding properties of an N-terminally truncated analogue of this enzyme (*des*Pro1)4OT are primarily due to the deletion of an essential backbone amide bond between Pro1 and Ile2. Our chemical synthesis and qualitative characterization of a series of backbone-modified 4OT analogues indicate that the backbone−backbone hydrogen-bonding interactions involving this N-terminal amide bond in 4OT play a major role in specifying the conformation of this enzyme's native folded state. It has recently been shown that hydrogen bonds involving amino acid side chains can specify the conformation of a protein's folded state.[19] Our results with the 4OT analogues in this study provide the first evidence that hydrogen bonds involving backbone−backbone interactions can play a major role in specifying the native state of a protein.

Experimental Section

Materials and Methods. The *tert*-butoxycarbonyl (Boc) amino acids and 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Peptide Institute, Inc. and Quantum Biotechnologies, respectively. RP-HPLC purifications were performed on a Rainin Dynamax system. Semipreparative RP-HPLC was performed on a C_{18} Vydac column (1.0 × 25.0 cm, 300 Å) using a flow rate of 3 mL/min, 230 nm detection, and linear mobile phase gradients of buffer B in A (A = 0.1% trifluoroacetic acid (TFA) in water, B = 90% acetonitrile and 0.09% TFA in water). Analytical RP-HPLC was performed on a C_{18} Vydac column (0.46 × 15.0 cm, 300 Å) using flow rates of 1 mL/min, 214 nm detection, and linear mobile phase gradients of buffer B in A. SEC was performed on a Rainin Dynamax system using a Superdex 75 HR 10/30 column (Pharmacia Biotech) at a flow rate of 1 mL/min and a mobile phase of 20 mM sodium phosphate buffer (pH 7.4). ESI-MS was performed on a Perkin-Elmer SCIEX API 150EX. Samples were infused into the mass spectrometer at a flow rate of 5 μL/min using a Harvard syringe pump. ESI mass spectra were acquired in the positive ion mode. Observed masses were calculated from the ion signals observed for all the protonation charged states of all molecular species, using the program *BioMultiview* (SCIEX). MALDI-MS mass spectra were obtained on a Voyager DE Biospectrometry Workstation (PerSeptive Biosystems, Inc.). All MALDI spectra were collected in the positive-ion mode using an acceleration voltage of 20 kV and a delay of 200 ns. Each spectrum obtained was the sum of approximately 50 laser shots. MALDI samples were prepared using approximately 5 μM of sample in 45% acetonitrile/55% water/0.1% TFA and an equal volume of saturated α-cyano-4-hydroxycinnamic acid (HCCA) matrix solution. MALDI spectra were calibrated using an external calibrant. CD spectra were recorded at 25 °C using a quartz cuvette with a 1-cm path length and 20 mM sodium phosphate (pH 7.4) buffer on a Jasco J-710 spectropolarimeter. Each CD spectrum was recorded using a bandwidth of 1.0 nm. Catalytic activity measurement on the 4OT analogues in this study were performed using the substrate 2-hydroxymuconate as described elsewhere.[6] All protein concentrations in this work were determined by ultraviolet/visible (UV/vis) spectroscopy using the Waddell method.[20]

General Chemical Synthesis Strategies. The 4OT analogues in this study were prepared by total chemical synthesis using manual solid-phase peptide synthesis (SPPS) methods and *in situ* neutralization protocols for Boc chemistry as described elsewhere.[21] Syntheses were initiated on Boc-N^{γ}-tosyl-L-arginine-4-(oxymethylphenylacetamidomethyl) resin and the following side chain protection was used: Arg(Tos), Asp(OcHxl), Glu(OcHxl), Lys(2ClZ), Ser(Bzl), His(Dnp), Thr(Bzl), where OcHxl is cyclohexyl, 2ClZ is 2-chlorobenzyloxycarbonyl, Dnp is 2,4-dinitrophenyl, and Bzl is benzyl. After chain assembly was complete, the Dnp group on His residues was removed by two 10-min treatments with a solution of 20% 2-mercaptoethanol and 10% *N*,*N*-diisopropylethylamine (DIEA) in dimethylformamide (DMF). Removal of side chain protecting groups and cleavage of peptide from the resin was achieved by treatment with anhydrous HF. After removal of HF under reduced pressure, the crude peptide products were precipitated and washed with anhydrous ether, dissolved in a 70% acetonitrile water solution containing 0.1% TFA, diluted with water, and lyophilized.

Combinatorial Syntheses. Analogues **1**−**3**, **4**−**7**, and **8**−**11** were respectively synthesized in three, small combinatorial libraries, library I, II, and III. Library I was synthesized on the solid phase in a single synthetic procedure that involved splitting and pooling the resin at specific points in the synthesis. This procedure involved the use of two vessels where the first vessel comprised the wild-type (Wt) polypeptide synthesis and the second vessel stored the polypeptide analogues of the library. After each of the three N-terminal residues were incorporated into the polypeptide chain of 4OT in the first vessel, one-fourth of the original total peptide−resin volume was transferred to the second vessel until all the analogues were completely synthesized. Library II was synthesized on the solid phase using a split-resin strategy similar to that described above for library I. Library III was derived from library I. It was prepared on the solid phase by acetylating the N-termini of the side-chain-protected, resin-bound peptides in library 1. The acetylation reaction involved reacting the primary amino group

(17) Arnett, E. M.; Mitchell, E. J.; Murty, T. S. S. R. *J. Am. Chem. Soc.* **1974**, *96*, 3875−3891.
(18) Jaffe, H. H.; Freedman, L. D.; Doak, G. O. *J. Am. Chem. Soc.* **1953**, *75*, 2209−2221.
(19) Hill, R. B.; Hong, J.-K.; DeGrado, W. F. *J. Am. Chem. Soc.* **2000**, *122*, 746−747.
(20) Waddell, W. J. *J. Lab. Clin. Med.* **1956**, *48*, 311−314.
(21) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int J. Pept. Protein Res.* 1992, *40*, 180−193.

at the N-termini of these polypeptide chains with 0.5 M acetic anhydride and 0.5 M pyridine in DMF for 10 min.

Synthesis of Analogue 9. Side-chain- protected peptide−resin Ile^{2}-Arg^{62}-PAM (approximately 300 mg, of 0.089 mmol/g of peptide−resin) was treated with 0.5 M acetic anhydride (2.5 mmol) in 0.5 M pyridine (2.5 mmol) in DMF. This N-terminal acetylation reaction was allowed to proceed for 10 min. Ultimately, the peptide was deprotected and cleaved from the resin by treatment with HF. Subsequent RP-HPLC purification of the crude peptide product (150 mg) yielded approximately 6 mg of pure peptide.

Synthesis of Analogue 12. Analogue **12** was generated by hydrolysis of **14** (the synthesis of which is described below). Crude synthetic product (approximately 100 mg) obtained from the synthesis of **14** was quantitatively hydrolyzed after 5 h at room temperature in 5 mL of a 200 mM NH_4HCO_3 buffer (pH 6) containing 1 M hydroxylamine. After lyophilization, the crude peptide product was purified by RP-HPLC to yield the 61 amino acid polypeptide chain of **12**.

Synthesis of Analogue 13. Side-chain-protected peptide−resin Ala^{3}-Arg^{62}-PAM (approximately 300 mg, of 0.089 mmol/g of peptide−resin) was treated with 1.1 mmol of valeric acid $CH_3(CH_2)_3CO_2H$ (Aldrich) that was preactivated in a DMF solution of 0.5 M HBTU (1.0 mmol) and DIEA (2.8 mmol). The valeric acid coupling reaction was allowed to proceed for 15 min at room temperature. Ultimately the peptide was deprotected and cleaved from the resin by treatment with HF. Subsequent RP-HPLC purification of the crude peptide product (150 mg) yielded approximately 6 mg of pure peptide.

Synthesis of Analogue 14. Analogue **14** was manually synthesized in stepwise fashion using standard SPPS protocols for Boc chemistry as described above. Ester bonds were incorporated into the polypeptide chains of these analogues by coupling an appropriate α-hydroxy acid at the site of the mutation. In **14** this involved coupling isoleucic acid (L-2-hydroxy-3-methylvaleric acid) at position 2 in 4OT's 62 amino acid polypeptide chain. In **15** this involved coupling L-lactic acid ((*S*)-2-hydroxypropionic acid) at position 39 in 4OT's 62 amino acid polypeptide chain. In **16**, the double mutant, both isoleucic acid and L-lactic acid were incorporated into 4OT's polypeptide chain at positions 2 and 39, respectively. The protocols used to couple these α-hydroxy acids to the growing peptide chain were identical to those described by Lu and co-workers.[2b] Subsequent coupling of appropriate Boc-protected amino acid residue to these α-hydroxy acids was also performed according to the protocol described by Lu and co-workers.[2b] After chain elongation was complete, the Dnp protecting groups on the two His residues were removed by two 15-min treatments of the resin with thiophenol (0.72 mmol) solution in DMF (1 mL). The remaining side chain protecting groups were removed and the peptide was cleaved from the resin by treatment with anhydrous HF. Subsequent RP-HPLC purification of the crude peptide product obtained from each synthesis gave good yields (approximately 3−5 wt %) of high-purity 62 amino acid polypeptide chains of 4OT.

Protein Folding Protocols for Libraries I−III. Libraries I−III were folded by dissolving 20 mg of the crude, lyophilized product from each synthesis in 100 μL of a 20 mM sodium phosphate buffer (pH 7.4) containing 6 M guanidine hydrochloride (GdnHCl). To initiate folding, this concentrated protein solution was diluted with 15 mL of 20 mM sodium phosphate buffer (pH 7.4). The resulting protein solution was concentrated to 1.5 mL using an ultrafiltration membrane (Amicron Diaflo, YM3) with a molecular weight cutoff of 3500 and allowed to sit at room temperature for 1−3 days. Ultimately, the mixture was centrifuged to remove any precipitate and the solution was analyzed by SEC.

Protein Folding Protocols for Analogues 1, 9, and 12−16. Each analogue was folded in a manner similar to that described above for libraries I−III, with the exception that RP-HPLC-purified samples were used in the folding protocol instead of the crude peptide product. After HPLC purification and lyophilization, peptide analogues **1**, **9**, and **12**−**16** (0.5 mg) were each dissolved in 20 μL of a 20 mM sodium phosphate buffer (pH 7.4) containing 6 M GdnHCl. This concentrated peptide solution was then diluted to 1.0 mL with 20 mM sodium phosphate buffer at pH 7.4 and allowed to sit for 1 h at room temperature. After centrifugation to remove any precipitate, the resulting protein solution was analyzed by SEC.

Acknowledgment. This work was supported in part by Duke University. Acknowledgment is also made to donors of the Petroleum Research Fund, administered by the ACS (M.C.F.), for partial support of this research. P.S. was the recipient of a Burroughs Wellcome Graduate Fellowship.

Supporting Information Available: RP-HPLC and ESI-MS analysis of synthetically derived wild-type 4OT and analogues **1**, **9**, and **12**−**16** (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0006090