

Crystal Structure of the *Escherichia coli* Peptide Deformylase<sup>†,‡</sup>

Michael K. Chan,\* Weimin Gong, P. T. Ravi Rajagopalan, Bing Hao, Chris M. Tsai, and Dehua Pei

Departments of Biochemistry and Chemistry, The Ohio State University, 484 West 12th Avenue, Columbus, Ohio 43210

Received May 16, 1997; Revised Manuscript Received August 26, 1997<sup>§</sup>

**ABSTRACT:** Protein synthesis in bacteria involves the formylation and deformylation of the N-terminal methionine. As eukaryotic organisms differ in their protein biosynthetic mechanisms, peptide deformylase, the bacterial enzyme responsible for deformylation, represents a potential target for antibiotic studies. Here we report the crystallization and 2.9 Å X-ray structure solution of the zinc containing *Escherichia coli* peptide deformylase. While the primary sequence, tertiary structure, and use of coordinated cysteine suggest that *E. coli* deformylase belongs to a new subfamily of metalloproteases, the environment around the metal appears to have strong geometric similarity to the active sites of the thermolysin family. This suggests a possible similarity in their hydrolytic mechanisms. Another important issue is the origin of the enzyme's specificity for N-formylated over N-acetylated substrates. Based on the structure, the specificity appears to result from hydrogen-bonding interactions which orient the substrate for cleavage, and steric factors which physically limit the size of the N-terminal carbonyl group.

Prokaryotic protein synthesis begins with formylmethionine-tRNA<sub>i</sub> resulting in the synthesis of N-terminally formylated polypeptides (Meinzel et al., 1993). Peptide deformylase catalyzes the subsequent removal of the formyl group for the majority of bacterial proteins (Adams, 1968; Livingston & Leder, 1968). Although N-terminal formylation is not essential for bacterial survival, it can stimulate protein synthesis by facilitating the participation of Met-tRNA<sup>Met</sup><sub>f</sub> in translation initiation and by preventing its recognition by the elongation apparatus (Guillon et al., 1993; Varshney & RajBhandary, 1992). Deformylation of peptides after translation is apparently a unique feature of bacterial cells and is essential for bacterial cell survival (Mazel et al., 1994).

Peptide deformylase represents a new subfamily of metalloproteases with interesting physical and catalytic properties. The amino acid sequences of deformylases from *Escherichia coli* (Mazel et al., 1994) *Thermus thermophilus* (Meinzel & Blanquet, 1994), *Haemophilus influenzae* (Fleischmann et al., 1995), *Mycoplasma genitalium* (Fraser et al., 1995), and *Lactococcus lactis*, *Bacillus subtilis*, *Calothrix*, and *Thermotoga maritima* (Mazel et al., 1997) share 28–65% sequence identity, but have no resemblance to other known proteins. The only similarity to other proteins is the presence of an HEXXH motif, characteristic of the family of zinc metallopeptidases (Vallee & Auld, 1990). Purified deformylase has been reported to contain one Zn<sup>2+</sup> atom per polypeptide (Meinzel & Blanquet, 1993; Meinzel et al., 1995). Mutagenesis studies have implicated that two histidines of the HEXXH motif, and Cys-90, are ligands for the metal ion (Meinzel et al., 1995). Another important feature of peptide deformylase is its high specificity for N-formylated substrates. It hydrolyzes corresponding acetylated peptides with rates more than 4 orders of magnitude

lower than the formylated counterparts (Adams, 1968; Meinzel & Blanquet, 1995; Wei & Pei, 1997).

In order to understand the catalytic mechanism of deformylase and its requisite selectivity toward formylated over other acylated substrates, we have undertaken structural studies of this enzyme. In this work, we report the three-dimensional structure of peptide deformylase determined by X-ray crystallography. A mononuclear metal ion is found at the bottom of a cleft, the putative active site for the enzyme. This structure reveals that the metal ion has a tetrahedral geometry with three protein ligands, His-132, His-136, and Cys-90, and a water molecule serving as the fourth ligand. Residues in several conserved motifs converge at the putative active site. Based on the structure, a catalytic mechanism is proposed. While this work was in progress, the structure of a truncated deformylase was solved by NMR (Meinzel et al., 1996a). This X-ray structure reveals significant differences in the tertiary structure and environment around the active site.

## MATERIALS AND METHODS

**Isolation and Crystallization.** *Escherichia coli* deformylase was purified as reported in the following paper (Rajagopalan et al., 1997). The initial crystallization conditions were determined via the hanging drop method using a sparse matrix screen of 50 conditions (Jancarik & Kim, 1991). Droplets containing 2  $\mu$ L of protein solution [20 mg/mL protein, 20 mM NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 7), 10 mM NaCl, and 1.0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>] and 2  $\mu$ L of precipitating solution (0.1 M sodium acetate buffer, pH 4.6, and 2.0 M sodium formate) were mixed on siliconized cover slides, inverted over a greased Linbro tray and sealed above a reservoir containing 0.5 mL of the precipitating solution. Small diamond-shaped crystals could be obtained at 4 °C after 3 days which grew to a size of 0.4 mm  $\times$  0.2 mm  $\times$  0.1 mm.

The crystals diffract to 2.9-Å resolution and have cell dimensions,  $a = b = 55.2806$  Å,  $c = 230.4159$  Å,  $\alpha = \beta = 90.0^\circ$ , and  $\gamma = 120.0^\circ$ . The systematic absence (001 except  $l = 6n$ ) and the Laue symmetry of the diffraction pattern (6/mmm) indicate that the space group is either  $P6_122$  or

<sup>†</sup> This work was supported by funds from The Ohio State University and the Petroleum Research Fund (30850-G3 to M.K.C.).

<sup>‡</sup> The coordinates have been deposited in the Protein Data Bank (PDB file name: 1DFF).

\* To whom correspondence should be addressed. Telephone: 614-292-8375. FAX: 614-292-6773. E-mail: chan@chemistry.ohio-state.edu.

<sup>§</sup> Abstract published in *Advance ACS Abstracts*, October 15, 1997.

Table 1: Summary of Diffraction Data

	native	EMTS
resolution of data (Å)	2.88	3.09
no. of data collected	73018	37985
no. of unique data		
isomorphous	4903	4024
anomalous		3443
completeness of data set (%)	91.6	91.6
average $I/\sigma$ of data set	12.71	14.48
completeness of outermost shell (%)	80.6	86.9
average $I/\sigma$ of outermost shell	4.84	5.82
$R_{\text{sym}}(I)^a$ (%)	9.1	7.3
phasing		
isomorphous $R$ -factor (%)		17
phasing power at 3.1 Å <sup>b</sup>		
isomorphous		3.02
anomalous		1.36
overall figure of merit		0.608
$R_{\text{cullis}}^c$		0.459

<sup>a</sup>  $R_{\text{sym}}(I) = \sum_h \sum_i |I_i - I| / \sum_h \sum_i I_i$ , where  $I$  is the mean intensity of the  $I$  observations of reflection  $h$ . <sup>b</sup> Phasing power =  $\text{rms}(\langle F_H \rangle / E)$ , where  $F_H$  is the calculated structure factor of the heavy atoms and  $E$  is the residual lack of closure. <sup>c</sup>  $R_{\text{cullis}} = \sum ||F_{\text{PH}} - F_{\text{P}}| - F_{\text{H}}(\text{calc})| / |F_{\text{PH}} - F_{\text{P}}|$ , where  $F_{\text{PH}}$  and  $F_{\text{P}}$  denote observed derivative and native crystal structure factors, respectively, and  $F_{\text{H}}$  denotes the calculated heavy atom structure factor.

*P6<sub>3</sub>22*. The calculated value of  $V_m$  (Matthews, 1968) in either of these space groups is 2.29, consistent with one molecule in the asymmetric unit.

**Data Collection and Processing.** X-ray intensity data were collected from single crystals using Cu K $\alpha$  radiation and a Rigaku Raxis II imaging plate system operated at 50 kV and 100 mA. The RAXIS software was used to index images, integrate intensities, and scale the data. The long  $c$ -axis required careful positioning of the  $c^*$  axis along the spindle. While each data collection was sufficient for analysis, a merged native set from two different crystals gave greater completeness of the dataset. For the EMTS (ethyl mercurithiosalicylic acid, sodium salt) derivative, two orientations of the same crystal were utilized.

**Structure Determination and Refinement.** The structure of *E. coli* deformylase was solved by the SIRA method. A single EMTS derivative was obtained with a single high-occupancy site. The EMTS derivative data together with the anomalous diffraction from 12.0 to 3.1 Å were used to obtain the initial SIRA phases. The refinement of heavy atom parameters, solvent flattening, and generation of the initial phases were performed using PHASES (Furey & Swaminathan, 1995). The initial phase was improved by solvent-leveling at 3.1 Å using a solvent content of 0.40 and integrated radii of 10 Å. The refinement procedure above was carried out for both possible space groups, *P6<sub>1</sub>22* and *P6<sub>3</sub>22*, and for each hand within each space group. An electron density map from one of the two *P6<sub>1</sub>22* orientations was of higher quality and more interpretable than any of the others, so this was chosen to build the initial model. The phasing statistics are displayed in Table 1.

The initial model was built using the package O (Jones et al., 1991), and refined to an initial  $R$ -factor of 23.61% using the program X-PLOR (Brünger, 1988). The deformylase model was then rebuilt using a new map generated from combined SIRA and model phases to give the present structure, which has been refined to 2.9 Å resolution with 1277 protein atoms, 47 solvent molecules, and individual  $B$  values. No sigma cutoff was used for either phasing or

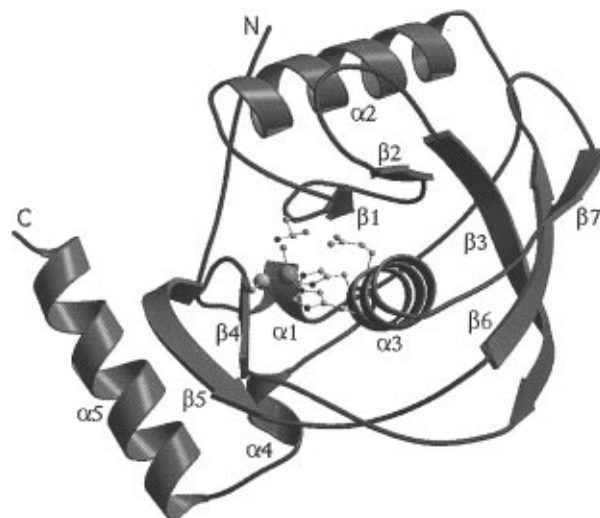


FIGURE 1: Ribbons diagram of the *E. coli* deformylase viewed toward the substrate cleft into the metal center. The secondary structures are color-coded with  $\alpha$ -helical regions as light blue,  $\beta$ -sheet regions as red, and the remainder as green. Atoms in the metal center are colored by elements, with carbon as gray, nitrogen as blue, oxygen as red, sulfur as yellow, phosphate as magenta, and the metal ion as green. This figure was prepared with the programs MOLSCRIPT and Raster 3D (Kraulis, 1991; Merritt & Murphy, 1994).

refinement. The  $R$ -factor of the current model is 18.4% ( $R_{\text{free}} = 28.9\%$ ), with reasonable bond distances and angles (distances rms, 0.014 Å; bond angle rms, 1.84°). The protein model was assessed using the program PROCHECK (Laskowski et al., 1993). Based on analysis of the Ramachandran plot, the structure is acceptable; 82% of the residues are in the most favored regions, 16% are in the allowed regions, and only one residue, Glu-95, is in the generously allowed region. All stereochemical parameters are better than or within the expected cutoffs. The programs VERIFY-3D (Luethy et al., 1992) and ERRAT (Colovos & Yeates, 1993) also gave acceptable values.

## RESULTS AND DISCUSSION

**The Protein.** *E. coli* peptide deformylase consists of a single domain of 168 amino acid residues. The protein, shown in Figure 1, contains three major  $\alpha$ -helices, three  $\beta$ -sheet regions, and a potentially critical 3-10 helix. The secondary structure assignments determined using PROCHECK (Laskowski et al., 1993) and DSSP (Kabsch & Sander, 1983) are as follows:  $\alpha$ -helix, (I) 11–14, (II) 25–40, (III) 124–136, (IV) 142–145, (V) 148–163;  $\beta$ -strands, (I) 45–47, (II) 57–60, (III) 70–81, (IV) 87–90, (V) 93–99, (VI) 105–111, (VII) 117–120; and 3-10 helix, 49–51. In some sense, the overall fold of the protein resembles a hand which wraps around the central  $\alpha$ -helix, helix III. The N-terminal residues, including helix I and helix II, and the antiparallel  $\beta$ -sheet formed from  $\beta$ -strands I, II, and III make up the fingers; the  $\beta$ -sheet formed from  $\beta$ -strands III, VI, and VII forms the palm; and helices IV and V and strands IV and V make up the thumb. While two of the major  $\alpha$ -helices, helix II and helix V, appear to play structural roles, the third major helix, helix III (124–136), appears to be involved in both structure and catalysis. Located on helix III is a conserved HEXXH motif, which is critical for metal coordination and substrate activation. The structure also indicates the presence of a 3-10 helix (residues 49–51) which

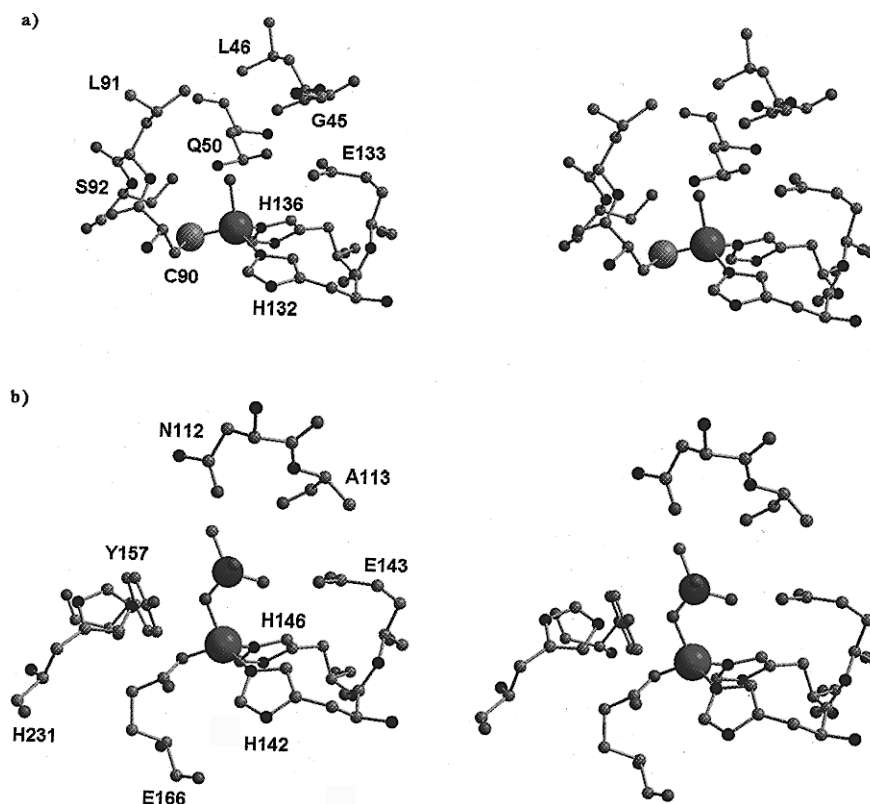


FIGURE 2: Metal site of deformylase and a comparison of its structure to the zinc site of thermolysin (Tronrud et al., 1987). (a) Ball-and-stick stereodiagram of deformylase viewed in the same orientation as Figure 1. (b) Ball-and-stick diagram of thermolysin viewed in a similar orientation to (a). The atoms are colored by element as in Figure 1. The thermolysin coordinates were obtained from the Protein Data Bank (PDB ID code: 6TMN). These figures were prepared with the programs MOLSCRIPT and Raster 3D (Kraulis, 1991; Merritt & Murphy, 1994).

orients a conserved glutamine, Gln-50, and which may be important for catalysis. The C-terminal helix, helix V (148–163), which is truncated in the NMR structure, forms an  $\alpha\beta$  interaction with  $\beta$ -strand V (93–99). Helix V appears to buttress the loop containing the coordinated cysteine ligand perhaps to stabilize the positioning of this domain. Despite having crystallized the intact protein, the last four amino acids are not observed, and are disordered.

**The Metal Site.** The *E. coli* deformylase has previously been reported to contain one  $\text{Zn}^{2+}$  ion per polypeptide (Meinzel & Blanquet, 1993; Meinzel et al., 1995). Indeed, a tetrahedral metal ion is found in the three-dimensional structure, with side chains of Cys-90, His-132, and His-136 from the protein, and a bound water molecule as its ligands (Figure 2a). The water is positioned based on its electron density (Figure 3). The lack of anomalous scattering exhibited by the metal center is consistent with its assignment as zinc. There is some concern, however, whether zinc is the metal ion in the native enzyme (Rajagopalan et al., 1997). The surroundings of the metal ion are formed by four motifs that are distant from each other in their primary sequence. Three of these motifs, residues 43–52, 88–92, and 130–139, are highly conserved among all deformylase proteins (Mazel et al., 1997). The protein ligands originate from two of the motifs, Cys-90 from motif E88–S92 and His-132 and His-136 from motif Q131–G139. This latter region contains the signature HEXXH motif, which is found in many zinc metalloproteases including thermolysin (Blundell, 1994; Jongeneel et al., 1989). Together with results from biochemical studies (Rajagopalan et al., 1997), the structural features around the metal suggest that it is the active site, and that the metal ion directly participates in catalysis.

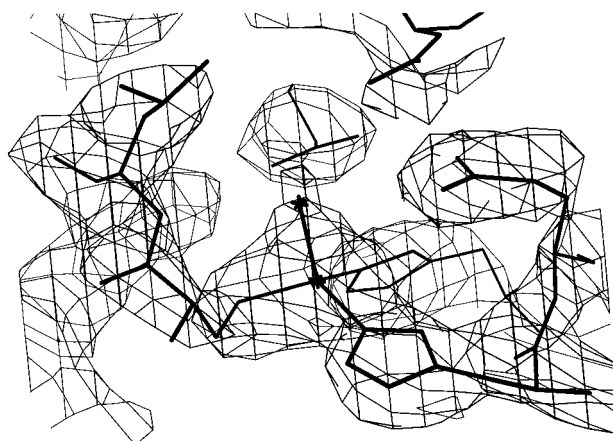


FIGURE 3:  $2F_o - F_c$  map of the metal active site. The contour level is  $1\sigma$ . This figure was generated using the program O (Jones et al., 1991).

Glutamate, E133, on the HEXXH motif is positioned near the metal-bound water in an orientation similar to that of the conserved glutamate in thermolysin (Figure 2b) (Matthews, 1988). It may play a similar role in the protonation and deprotonation of reaction intermediates during catalysis.

Among the other noncoordinating residues near the metal site, one of the most important may be the conserved Gln-50, the central residue of the 3-10 helix (residues 49–51). This residue is hydrogen-bonded to the water, which is coordinated to the metal ion, and hence is likely involved in the protonation and deprotonation of reaction intermediates. It may also play an important role in substrate specificity. In support of the importance of the orientation of this residue, the side chain amide nitrogen and carbonyl are held in place

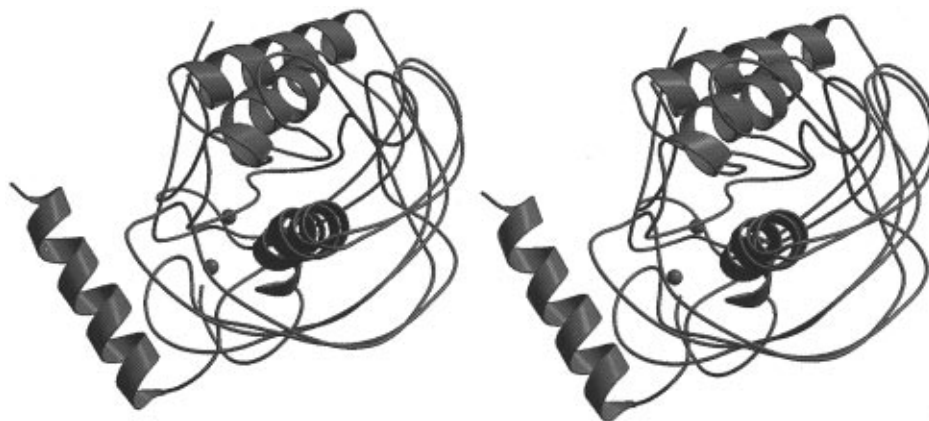


FIGURE 4: Comparison of the X-ray structure of *E. coli* deformylase (red) with one conformation of the NMR structure (blue) (PDB ID code: 1DEF). This figure was prepared with the programs MOLSCRIPT and Raster 3D (Kraulis, 1991; Merritt & Murphy, 1994).

by hydrogen bonds, the first to Ser-91 and the second to the amide nitrogen of Ala-47. The amide hydrogen of Leu-91 is also oriented to hydrogen-bond the coordinated water.

**Comparison to the Published NMR Structure of a Truncated Deformylase.** While this work was being completed, Meinnel et al. (1996a) reported the solution structure of an *E. coli* deformylase variant, in which the last 21 amino acids residues were truncated. They described the secondary and tertiary structures of deformylase and its interactions with potential substrates, and reported that the orientation of the Cys-90 relative to the two histidine ligands is unique, when compared to the orientation of functionally analogous ligands found in other zinc metalloproteases.

One possible advantage that an X-ray structure may have over an NMR structure is better resolution. This allows for better assignment of backbone and side chains which may be important in defining the location of residues around an active site. This is critical to the fundamental understanding of the mechanism at the metal center, and for the design of potential antibiotics. The X-ray structure also has the added advantage of being the full-length native protein, while the NMR structure determined was of a fragment (Meinnel et al., 1996a). Interestingly, a comparison of the NMR structure and our X-ray structure, shown in Figure 4, reveals significant differences in their tertiary structures.

Overall, the secondary structure assignments are similar, although there are differences in several  $\beta$ -strands. The largest discrepancy is in the assignment of the first  $\beta$ -strand determined from the NMR structure (residues 16–19). This is assigned as random coil in the X-ray structure. The second  $\beta$ -strand from the NMR structure (residues 53–61) is much shorter in the X-ray structure (residues 57–60). The third and fourth NMR  $\beta$ -strands (residues 69–73, 76–81) appear to be single strand (residues 70–81) based on the X-ray structure. These differences could represent differences in the solution versus solid state structures, or reflect a difference in the algorithm in which these secondary structure assignments were made. The  $\alpha$ -helix assignments from the NMR and X-ray models are similar, although the X-ray structure of the complete protein reveals an additional  $\alpha$ -helix (helix V) which was truncated in the NMR study.

Comparison of the tertiary structures reveals significant differences. Relative to the X-ray structure, the protein loop in the NMR structure appears more compact, with the two sides of the protein loop being shifted toward each other at the link, the metal–cysteine bond. Due to steric factors,

one side is shifted upward, and the other is shifted downward. This conformational difference has a major impact on the region around the metal site. The location of Cys-90, the third metal ligand, is several angstroms shifted between the NMR and X-ray structures. This results from a difference in the positioning of the entire loop containing Cys-90, and a bending of the central  $\alpha$ -helix (helix III). The consequence of these changes is a very different picture for the active site. In the NMR structure, Cys-90 is positioned in a different orientation relative to analogous residues found in thermolysin (glutamate) (Tronrud et al., 1987) and stromelysin (histidine) (Gooley et al., 1994). In the X-ray structure, however, this orientation is similar to that found in these same metalloenzymes. The NMR structure also leads to the observation that the metal center is buried. In the X-ray structure, there is an opening which is filled by a solvent molecule. The X-ray structure also allows for identification of specific residues, Gln-50 and Glu-133, which may be critical to the mechanism of hydrolytic cleavage. These same residues are pointed away from the active site in the majority of the conformations reported for the NMR structure. This ability of the X-ray structure to identify residues which may be critical to the catalytic mechanism of deformylation would support it being the active enzyme conformation. Additionally, preliminary studies reveal that crystals of deformylase have activity.

The origin of these structural discrepancies is unclear. One possibility is that this difference results from the C-terminal 21 amino acid truncation. This idea is supported by the fact that the C-terminal helix (helix V), which is truncated in the NMR structure, forms an  $\alpha\beta$  interaction with  $\beta$ -strand V on the protein loop. The role of this helix may be to stabilize the orientation of the loop containing Cys-90, the third metal ligand. It is possible that the loss of this helix allows the loop to shift to the positions observed in the NMR structure. Evidence against this, however, comes from the observation that truncation of the C-terminal 21 amino acid residues has little effect on the catalytic activity of *E. coli* deformylase (Meinnel et al., 1996b; P.T.R.R. and D.P., unpublished results). Whether this alternative conformation determined by NMR has functional significance or is an artifact of the 21 amino acid truncation will require additional study.

**Comparison to Thermolysin.** While the tertiary and secondary structures of peptide deformylase differ from other zinc metalloproteases (Meinnel et al., 1996a), the active site

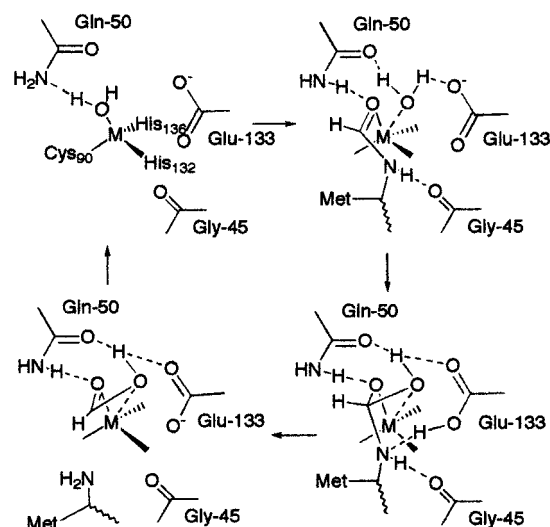


FIGURE 5: Proposed mechanism of deformylation by *E. coli* peptide deformylase.

environment has strong homology to other zinc metalloenzyme active sites. Most notable is its similarity to proteins of the thermolysin family, all of which contain the characteristic HEXXH motif (Blundell, 1994; Matthews, 1988). A comparison of the metal sites of deformylase and thermolysin is shown in Figure 2. The thermolysin molecule has been rotated into the same orientation as deformylase using the program O (Jones et al., 1991) using a least-squares alignment of the metal ion and the conserved histidines and glutamate of the HEXXH motif found in both structures. Both contain a tetrahedral metal ion coordinated by three protein ligands and a water molecule. While deformylase contains a cysteine residue and thermolysin contains a glutamate, the remaining two protein ligands are histidine, and the overall positioning of the coordinated ligands is similar. Perhaps most intriguing is the similarity in the potential hydrogen-bonding interactions. As expected, the conserved Glu from the HEXXH motif is in a similar location, but perhaps somewhat surprising is the number of other hydrogen-bonding interactions which are also conserved, although using different residues. While in thermolysin these interactions come from a tyrosine and histidine, in deformylase the coordinated water is in hydrogen-bonding distance to a glutamine and a backbone amide nitrogen. Also conserved in both structures is the backbone carbonyl group (from Gly-45 in deformylase, and Ala-113 in thermolysin) which has been implicated in hydrogen-bonding the amide nitrogen. In fact, the  $\beta$ -sheet region which contains this conserved backbone carbonyl and which is formed from  $\beta$ -strands I, II, and III in deformylase overlaps with a similar motif in thermolysin. These features may indicate a general similarity in the mechanism of deformylation by deformylase and peptide bond cleavage by thermolysin.

**A Proposed Mechanism.** The deformylase structure provides a picture on which to base a mechanistic hypothesis on how the protein functions. One plausible mechanism is depicted in Figure 5. This proposal has some similarity to the mechanism suggested for thermolysin (Matthews, 1988), and more distantly carboxypeptidase A (Christianson et al., 1987). In this model, coordination of the formyl carbonyl to the zinc site results in formation of a five-coordinate center in which the water molecule shifts to hydrogen-bond both Gln-50 and Glu-133. The water proton hydrogen-bonded

to Glu-133 is proposed to shift to protonate the amide of the substrate making it a better leaving group. This would also form a metal hydroxide which is a better nucleophile. The combined effect of having a good nucleophilic oxygen for attack on the carbonyl and a good nitrogen leaving group results in facile formamide bond cleavage. Consistent with this mechanism, mutations at Glu-133 (to Ala, Asp, or Gln) eliminate its catalytic activity (Meinzel et al., 1995).

It is important to note that other variations of this mechanism are possible. The nucleophilic water, for instance, may not be bound to the metal ion. Another subtlety which remains to be understood is why deformylases use cysteine as a third ligand, while all other amide hydrolases utilize carboxylate or histidine residues. Some of these detailed questions regarding the mechanism of peptide deformylation could be best addressed via the structure determination of site-directed mutants, and protein/inhibitor complexes. Such studies are presently underway.

**The Origin of Specificity.** One of the important characteristics of the deformylase enzyme is its selectivity for peptides containing N-terminal formylmethionine. *E. coli* deformylase has a high specificity for formylated substrates over acetylated substrates (Adams, 1968; Wei & Pei, 1997). Understanding the origin of this specificity is one of the intriguing aspects of the deformylase protein, particularly since these new findings could be used to engineer new types of hydrolytic enzymes.

Ultimately the issue of specificity would be best addressed through the structure determination of a high-affinity substrate/inhibitor bound to the deformylase protein. This would allow for direct determination of the substrate binding mode and potentially the structures of reaction intermediates. In the absence of such a structure, however, one alternative way to gain insight is from comparisons to the binding mode of ligands to thermolysin, a protein with a similar active site structure and perhaps a similar mechanism.

Structures of inhibitors bound to thermolysin have been solved, and a comparison of the binding mode of these substrates with the active site of deformylase is shown in Figure 2. The position and orientation of the phosphoramidite group coordinated to the zinc in thermolysin are consistent with what one would expect for the tetrahedral intermediate of deformylase based on potential hydrogen-bonding interactions. It may be that these hydrogen-bonding interactions act to orient the various groups of the transition state. Based on this comparison, in the tetrahedral intermediate the formyl hydrogen should be directed upward toward a cleft formed between residues Leu-46, and Leu-91. These groups may form a physical barrier against larger substrates such as a methyl from an acetyl group. When the structures of native deformylase and the thermolysin inhibitor complex are overlayed, Leu-91 of peptide deformylase is in van der Waals distance of one of the phosphoramidite oxygens.

Three residues from the conserved EGCLS motif, Cys-90, Leu-91, and Ser-92, and the conserved residue Gln-50 are particularly interesting in how they contribute to orient the tetrahedral intermediate, and assist in inducing the observed specificity. One residue orients the tetrahedral intermediate on the metal, another forms part of the physical barrier which limits the substrate size, and the last two residues anchor the first two residues in place. The residue which may play a pivotal role in orienting the position of the coordinated substrate is Gln-50. It is positioned to

hydrogen-bond both the nucleophilic water and the carbonyl group. The physical barrier is formed from residue Leu-91. It is located just above the metal ion where it may limit the access of larger substrates to the active site. Perhaps the most subtle contributor may be Ser-92. This residue along with Cys-90 acts to fix Leu-91 in place. While Cys-90 itself is anchored via direct coordination to the zinc atom, Ser-92 hydrogen-bonds to Glu-50 and may also affect the orientation of this group. The importance of these particular residues in deformylation is supported by the fact they are conserved within the entire deformylase family. Assessment of the function of the above residues by site-directed mutagenesis is already underway in our laboratories.

## CONCLUDING REMARKS

We have determined the X-ray structure of the deformylase protein to 2.9 Å resolution. The structure allows us to propose a mechanism for how the deformylation reaction takes place and suggests possible roles for certain residues in the specificity of the protein. This is a major step toward the design of inhibitors, which may have utility as antibiotics and for site-directed engineering experiments to alter the substrate specificity.

## ACKNOWLEDGMENT

We thank Professor Sundralingam and his co-workers for allowing us to collect data on their Rigaku RAXIS II imaging plate system.

## REFERENCES

- Adams, J. M. (1968) *J. Mol. Biol.* 33, 571–589.  
 Blundell, T. L. (1994) *Nat. Struct. Biol.* 1, 73–75.  
 Brünger, A. T. (1988) *J. Mol. Biol.* 203, 803–816.  
 Christianson, D. W., David, P. R., & Libscomb, W. N. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 1512–1515.  
 Colovos, C., & Yeates, T. O. (1993) *Protein Sci.* 2, 1511–1519.  
 Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., Mckenney, K., Sutton, G., Fitzhugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L. I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., & Nguyen, D. T. (1995) *Science* 269, 496–512.  
 Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J.-F., Dougherty, B. A., Bott, K. F., Hu, P.-C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchinson, C. A., III, & Venter, C. (1995) *Science* 270, 397–403.  
 Furey, W., & Swaminathan, S. (1995) in *Macromolecular Crystallography, a volume of Methods in Enzymology* (Carter, C., & Sweet, R., Eds.) Academic Press, Orlando, FL.  
 Gooley, P. R., O'Connell, J. F., Marcy, A. I., Cuca, G. C., Salowe, S. P., Bush, B. L., Hermes, J. D., Esser, C. K., Hagmann, W. K., Springer, J. P., & Johnson, B. A. (1994) *Nat. Struct. Biol.* 1, 111–118.  
 Guillon, J. M., Mechulam, Y., Blanquet, S., & Fayat, G. (1993) *J. Bacteriol.* 175, 4507–4514.  
 Jancarik, J., & Kim, S. H. (1991) *J. Appl. Crystallogr.* 24, 409–411.  
 Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr.* A47, 110–119.  
 Jongeneel, C. V., Bouvier, J., & Bairoch, A. (1989) *FEBS Lett.* 213, 110–114.  
 Kabsch, W., & Sander, C. (1983) *Biopolymers* 22, 2577–2637.  
 Kraulis, P. (1991) *J. Appl. Crystallogr.* 24, 946–950.  
 Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291.  
 Livingston, D. M., & Leder, P. (1968) *Biochemistry* 8, 435–443.  
 Luethy, R., Bowie, J. U., & Eisenberg, D. (1992) *Nature* 356, 83–85.  
 Matthews, B. W. (1968) *J. Mol. Biol.* 33, 491–497.  
 Matthews, B. W. (1988) *Acc. Chem. Res.* 21, 333–340.  
 Mazel, D., Pochet, S., & Marliere, P. (1994) *EMBO J.* 13, 914–923.  
 Mazel, D., Coic, E., Blanchard, S., Saurin, W., & Marliere, P. (1997) *J. Mol. Biol.* 266, 939–949.  
 Meinnel, T., & Blanquet, S. (1993) *J. Bacteriol.* 175, 7737–7740.  
 Meinnel, T., & Blanquet, S. (1994) *J. Bacteriol.* 176, 7387–7390.  
 Meinnel, T., & Blanquet, S. (1995) *J. Bacteriol.* 177, 1883–1887.  
 Meinnel, T., Mechulam, Y., & Blanquet, S. (1993) *Biochimie* 75, 1061–1075.  
 Meinnel, T., Lazennec, C., & Blanquet, S. (1995) *J. Mol. Biol.* 254, 175–183.  
 Meinnel, T., Blanquet, S., & Dardel, F. (1996a) *J. Mol. Biol.* 262, 375–386.  
 Meinnel, T., Lazennec, C., Dardel, F., Schmitter, J., & Blanquet, S. (1996b) *FEBS Lett.* 385, 91–95.  
 Merritt, E., & Murphy, M. (1994) *Acta Crystallogr. Sect. D* 50, 869–873.  
 Rajagopalan, P. T. R., Datta, A., & Pei, D. (1997) *Biochemistry* (following paper in this issue).  
 Tronrud, D., Holden, H., & Matthews, B. (1987) *Science* 235, 571–574.  
 Vallee, B. L., & Auld, D. S. (1990) *Biochemistry* 29, 5647–5659.  
 Varshney, U., & RajBhandary, U. L. (1992) *J. Bacteriol.* 174, 7819–7826.  
 Wei, Y., & Pei, D. (1997) *Anal. Biochem.* 250, 29–34.

BI9711543