

trile (including 15 mM ammonium acetate) in 60 min was carried out at a flow rate of 1 ml/min.

31. <sup>35</sup>S-labeled RNA derived from *Tetrahymena* rRNA wild-type intron (pBFSN-1) and mutant G264: C311G:G414A [(6); the wild-type and mutant introns were provided by M. D. Been] were prepared as described (12). The wild-type or the mutant intron RNA was incubated for 15 min at 37°C under splicing conditions [40 mM tris-HCl (pH 7.4), 10 mM MgCl<sub>2</sub>, 200 mM NaCl, 0.4 mM spermidine]. Cofactors were used at their concentrations (6) as follows: guanosine, 10 μM, and purine riboside, 300 μM (both from Sigma).

32. We thank R. Schroeder for discussions and comments and our colleagues in the Noller Laboratory for advice on experimental methods and for useful discussions. E. Westhof provided helpful information in explaining the perturbations in terms of the three-dimensional model for the group I intron. This research was sponsored by a long-term fellowship from the European Molecular Biology Organization (U.v.A.), NIH grant GM17129 (H.F.N.), and a grant to the CMB/RNA from the Lucille P. Markey Charitable Trust.

22 December 1992; accepted 10 March 1993

## New Domain Motif: The Structure of Pectate Lyase C, a Secreted Plant Virulence Factor

Marilyn D. Yoder, Noel T. Keen, Frances Journak\*

Pectate lyases are secreted by pathogens and initiate soft-rot diseases in plants by cleaving polygalacturonate, a major component of the plant cell wall. The three-dimensional structure of pectate lyase C from *Erwinia chrysanthemi* has been solved and refined to a resolution of 2.2 angstroms. The enzyme folds into a unique motif of parallel β strands coiled into a large helix. Within the core, the amino acids form linear stacks and include a novel asparagine ladder. The sequence similarities that pectate lyases share with pectin lyases, pollen and style proteins, and tubulins suggest that the parallel β helix motif may occur in a broad spectrum of proteins.

Pectate lyases (E.C. 4.2.2.2) are microbial extracellular enzymes that are important during plant pathogenesis (1). The enzymes randomly cleave α-1,4 linked galacturonosyl residues of the pectate component found in the primary cell wall and intercellular regions of higher plants. Enzymatic cleavage of the glycosidic bond occurs at a pH optimum of 8 to 11 through a β elimination mechanism, resulting in an unsaturated C-4-C-5 bond in the galacturonosyl moiety at the nonreducing end of the polysaccharide (1). Calcium is essential for enzymatic activity, but it is unclear whether Ca<sup>2+</sup> binds to the protein or to the substrate (2). Pectate lyases usually exist as multiple, independently regulated isozymes that are 27 to 80% identical in amino acid sequence (3, 4).

Pectate lyase C (PelC) was isolated from the periplasm of *Escherichia coli* cells containing a high-expression plasmid construct of the *pelC* gene from *Erwinia chrysanthemi* EC16 (5, 6). The periplasmic form of recombinant PelC has the identical molecular weight, isoelectric point, and maceration properties as the *E. chrysanthemi* isolate (7). The enzyme has 353 amino acids and a molecular weight of 37,676 (6). Crystals of PelC were grown from ammonium sulfate and belong to space group *P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>* with

unit-cell parameters of  $a = 73.38 \text{ \AA}$ ,  $b = 80.26 \text{ \AA}$ , and  $c = 95.12 \text{ \AA}$  and one molecule per asymmetric unit (8). The structure was solved with multiple isomorphous replacement (MIR) techniques (9). The model, which includes the first 352 residues but no solvent, has a crystallographic refinement factor of 20.2% for all measured reflections with structure-factor amplitude

$F > 2\sigma$  in the 2.2 to 5.0 Å resolution shell. The root-mean-square deviations are 0.006 Å from bond length ideality and 1.38° from bond angle ideality.

Because the structural motif of PelC is unusual, the correctness of the model was of particular concern throughout the analysis. In the initial stages of interpretation, there were few ambiguous regions in either the polypeptide tracing or in the amino acid sequence assignment in the MIR electron density maps. Furthermore, all heavy atom derivatives substituted at chemically reasonable sites: platinum at Met<sup>26</sup>, osmium at His<sup>153</sup>, and uranyl and lead at a common site in a pocket formed by Asp<sup>131</sup>, Glu<sup>166</sup>, and Asp<sup>170</sup>. After refinement, all carbonyl oxygens were visible in the final  $2F_0 - F_c$  electron density map (Fig. 1). All connections and polypeptide directions found in  $2F_0 - F_c$  maps, in which 10% of the total residues were sequentially omitted from the model, were consistent with those of the refined model. Only three residues were outside the normal range of the φ and ψ angles for standard secondary structural elements (Fig. 2A). Two additional measures supported the validity of the model (Fig. 2B).

The polypeptide backbone of PelC folds into a single structural domain (Fig. 3). The predominant secondary structural elements are parallel β strands that are coiled into a large right-handed cylinder that we call a parallel β helix. The NH<sub>2</sub>-terminal end of the cylinder is covered by a short α helix of three turns, and the COOH-terminal end is covered by polypeptide loops. Polypeptide loops also protrude from the

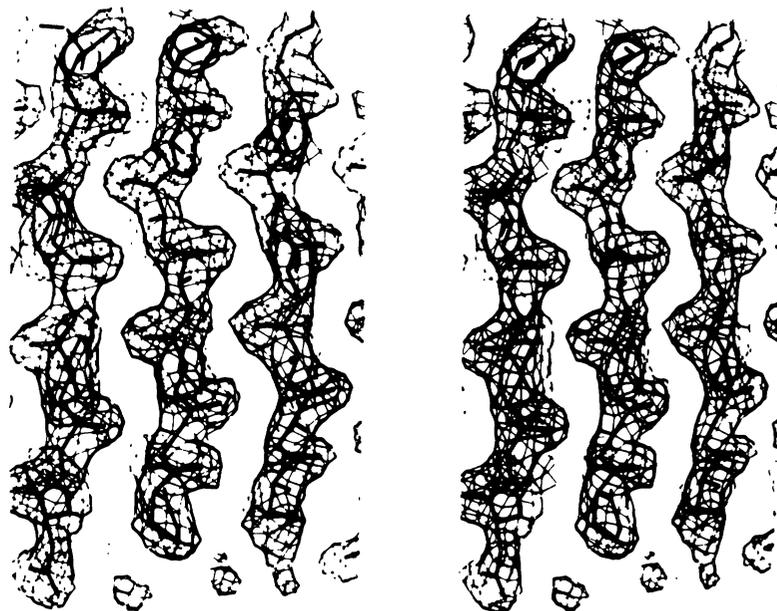


Fig. 1. Stereo view of the superposition of the backbone of several parallel β strands on the final  $2F_0 - F_c$  electron density map of PelC contoured at 1.0 σ. The model is shown in black lines, and the electron density map is in gray lines. All carbonyl oxygen atoms are apparent in the final map.

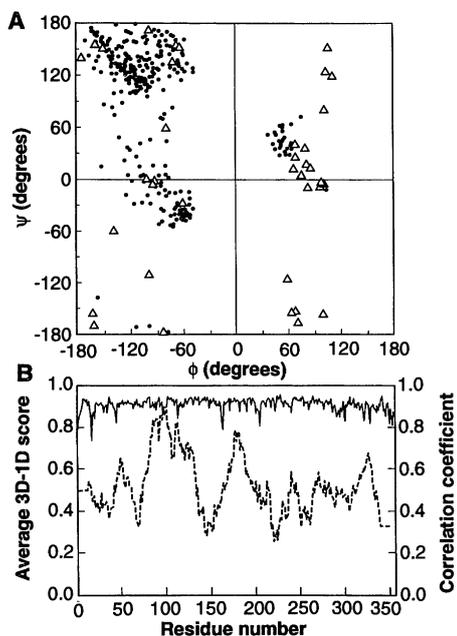
M. D. Yoder and F. Journak, Department of Biochemistry, University of California, Riverside, CA 92521.

N. T. Keen, Department of Plant Pathology, University of California, Riverside, CA 92521.

\*To whom correspondence should be addressed.

central parallel  $\beta$  helix at irregular intervals in the sequence. The loops vary in size and secondary structure and cover approximately 43% of the outer surface of the central cylinder. There are two disulfide bonds, Cys<sup>55</sup>-Cys<sup>172</sup> and Cys<sup>329</sup>-Cys<sup>352</sup>, both of which stabilize loops (Fig. 3) on the exterior of the parallel  $\beta$  helix.

The cylinder consists of seven complete turns and has dimensions of 34 Å in length and 17 to 27 Å in diameter. There is no periodicity of residues per turn of the parallel  $\beta$  helix because of the variation in position and size of the external loops in PelC. If there were no loops, then the parallel  $\beta$  helix in PelC would have approximately 22 residues per turn and a rise per residue of 0.22 Å. No repetitive pattern in the sequence or in any amino acid property is apparent. The inner core of the parallel  $\beta$

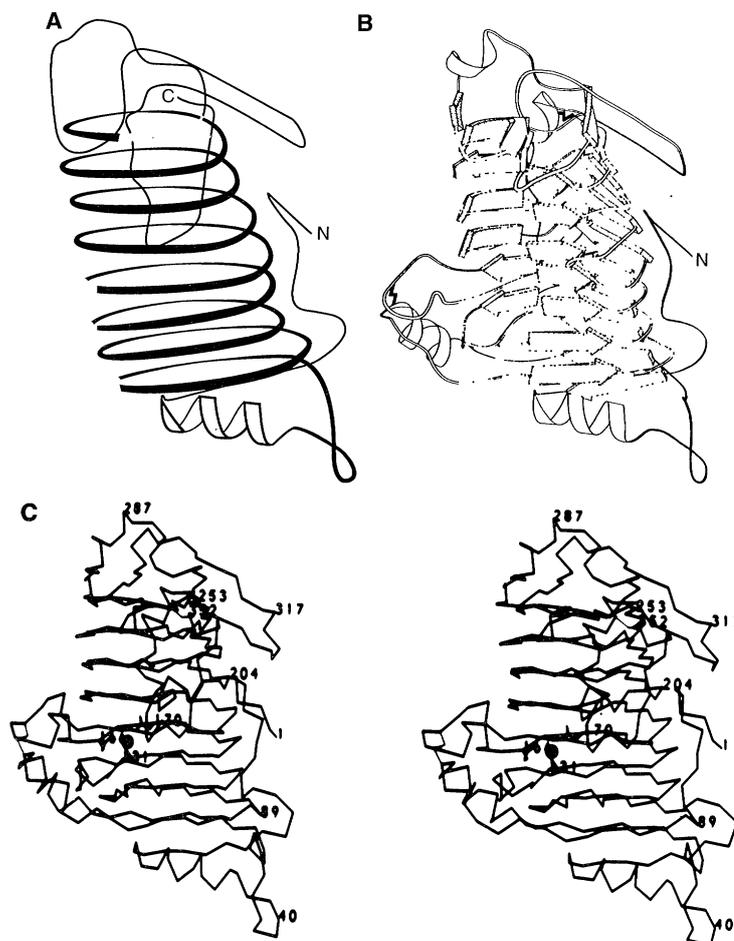


**Fig. 2.** (A) Ramachandran plot of dihedral angles of backbone atoms. Triangles represent glycine residues; the dots are remaining residues. The amino acids with the greatest deviation in  $\phi$  and  $\psi$  values include Tyr<sup>268</sup> ( $\phi = 103.8^\circ$ ,  $\psi = -13.3^\circ$ ), Arg<sup>218</sup> ( $\phi = 75.4^\circ$ ,  $\psi = 71.0^\circ$ ), and Ala<sup>150</sup> ( $\phi = -157.4^\circ$ ,  $\psi = -136.7^\circ$ ). (B) Main chain, real space correlation coefficient [solid trace (19)] on the right vertical axis and a three-dimensional (3D) window profile plot (20) with the average 3D-1D score for residues in a 21-residue sliding window (dashed trace) on the left vertical axis. The sequence position of the correlation coefficient and of the center of the sliding window is indicated on the horizontal axis. The validity of the PelC model is supported by the relatively smooth plot of the main chain, real space correlation coefficient and the high 3D-1D scores of the 3D window profile analysis. In the latter analysis, the overall score per number of residues, 183.05 for 352 residues, compares favorably with those of correct models (20).

helix is completely filled with hydrophobic and polar amino acids. Most unusual is the packing arrangement of side chains, which form prominent stacks (Fig. 4).

The present structural results do not resolve the controversy over the role of Ca<sup>2+</sup> because this ion could not be included in the (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> crystallization solutions. Nevertheless, the heavy atom derivative results are informative. Uranyl compounds are known to bind to carboxyl clusters of glutamic and aspartic acids, residues frequently involved in Ca<sup>2+</sup> binding (10). Lutetium has an ionic radius of 0.86 Å, close to the 1.0 Å radius of Ca<sup>2+</sup>,

and is used in substitutions for Ca<sup>2+</sup> in proteins (11). The Lu<sup>3+</sup> ion substitutes at a common site shared by the uranyl and lead derivatives and is located 2.1 to 2.9 Å from three acidic residues, Asp<sup>131</sup>, Glu<sup>166</sup>, and Asp<sup>170</sup>. In addition, three water molecules complete an octahedral coordination sphere, and each forms a bridge between the heavy atom and the side chains of Asp<sup>129</sup>, Lys<sup>172</sup>, and Arg<sup>218</sup>. Five of the amino acids are invariant among all extracellular pectate lyases, and the sixth, Glu<sup>166</sup>, is either glutamic acid or glutamine. The crystallographic results suggest that the common heavy atom site is the



**Fig. 3.** (A) Idealized schematic of the PelC backbone illustrating the parallel  $\beta$  helix topology. The gray regions indicate the positions where several external loops have been removed for clarity. (B) Ribbon diagram of the PelC polypeptide backbone. The arrows represent  $\beta$  strands, the coils represent  $\alpha$  helices, and the thick black lines represent disulfide bonds. The predominant supersecondary structural elements are three parallel  $\beta$  sheets, each marked by a distinctive arrow design. Each parallel  $\beta$  sheet is composed of seven to ten parallel  $\beta$  strands, with an average length of four to five residues in each strand. Each sheet has typical  $\beta$  properties including maximum interstrand hydrogen bonds between amide and carbonyl groups, alternating side chain orientations,  $\phi$  and  $\psi$  values within the standard  $\beta$  range, and a right-handed twist when viewed along the polypeptide chain direction. In the regions connecting the parallel  $\beta$  sheets, the geometry is atypical but recurrent and may represent novel aperiodic secondary structures: The most frequent connections have  $\phi$  and  $\psi$  angles similar to rare  $\gamma$ - $\beta_E$  turns (21) but an average C<sub>i</sub>-C<sub>i+3</sub> distance of 9.1 Å rather than 7.0 Å or less. Consequently, the peptide strand bends less than do typical  $\beta$  turns. (C) Stereo view of the  $\alpha$  carbon tracing of PelC. The disulfide bonds are indicated by gray lines, the putative Ca<sup>2+</sup> site is marked by a sphere, and three amino acids (Asp<sup>131</sup>, Glu<sup>166</sup>, and Asp<sup>170</sup>) that participate in its coordination are superimposed on the backbone.

putative, but weak, Ca<sup>2+</sup> binding site on the protein.

The putative Ca<sup>2+</sup> site is located on the exterior of the parallel  $\beta$  helix (Fig. 3) and in a narrow groove of surface charges (Fig. 5). The groove approximates the shape and length of a dodecameric galacturonate substrate. The results of computer graphic experiments, in which polygalacturonate was docked to PelC, are compatible with a shared Ca<sup>2+</sup> coordination between the enzyme and the substrate. Although no amino acids in the active site have been identified, the catalytic requirement for Ca<sup>2+</sup> and findings from the computer graphic experiments suggest that the putative Ca<sup>2+</sup> site on PelC may be the location of the active site.

Proteins fold into structural domains

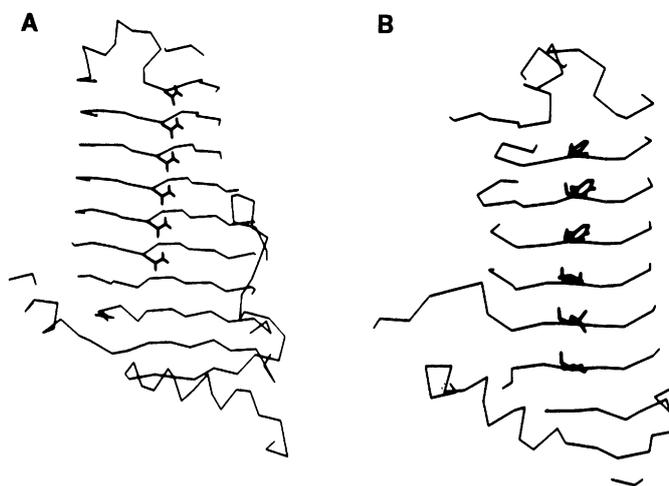
classified according to the predominant secondary structural feature. With the exception of the domains of proteins smaller than 10 kD, all structural domains belong to one of three categories: all  $\alpha$  helix,  $\alpha$  helix alternating with parallel  $\beta$  strands, and antiparallel  $\beta$  strands (12). The domain motif observed for PelC is unusual because the predominant structural feature is parallel  $\beta$  strands, but without helices between the  $\beta$  strands. A few short helices are located at the periphery of the parallel  $\beta$  helix, but there is no repetitive pattern in their arrangement. Unlike the structure of other enzymes with parallel  $\beta$  strands, polar groups are found on both surfaces of the parallel  $\beta$  helix. Moreover, most enzymes with parallel  $\beta$  strands are biochemically

less stable than other types of structural domains. In contrast, PelC is stable for prolonged periods at room temperature. The extensive network of interstrand hydrogen bonds and the multiple side chain stacks in the core probably contribute to its stability. One prominent feature of the motif is the simplicity of the probable folding mechanism—coiling of the peptide into a large cylinder much like winding rope into a tidy coil. The structure of the motif complies with all accepted structural rules, albeit in an unusual manner.

The structural results raise the possibility that the novel motif is related to functional properties of PelC. For example, the parallel  $\beta$  helix may satisfy the structural requirements for external secretion. Pectate lyases use two conserved secretion pathways: the Sec-dependent pathway to pass through the inner bacterial membrane to the periplasm (13) and an Out pathway to pass through the outer membrane in *E. chrysanthemi* (14). In addition to cleavage of an NH<sub>2</sub>-terminal signal peptide, all proteins exported by the Sec-dependent pathway appear to fold into their native conformation in the periplasm (15). The topology of PelC is compatible with the export of an unfolded polypeptide, followed by a simple but highly cooperative folding into the parallel  $\beta$  helix. No sequence signal has been identified in the Out pathway, which suggests that outer membrane secretion may be dependent on a structural feature (13, 15). The possibility that the unique topology of PelC might contribute to outer membrane secretion is compatible with observations that PelC leaks through dialysis membranes with an average molecular weight retention of 8 kD or greater. Similarly, the PelC motif may facilitate its pathogenic function of binding to and cleaving buried galacturonate polymers in plant cell walls.

Sequences of 12 extracellular and three periplasmic pectate lyases from four organisms have been reported (3). The periplasmic family, although highly homologous, shares little sequence homology with the extracellular enzymes. In contrast, all extracellular pec-

**Fig. 4.** Examples of stacking motifs within the core of the parallel  $\beta$  helix. The amino acids, including hydrogen atoms, involved in the interior stacks are superimposed on portions of the  $\alpha$  carbons. The amino acids are shown in black, and the backbone is in gray. (A) The stacking of six asparagines in consecutive turns of the parallel  $\beta$  helix. The side chains are oriented so that the number of hydrogen bonds between the asparagines is maximized



and the overall stacking resembles a ladder. One hydrogen of each Asn NH<sub>2</sub> group forms a hydrogen bond with an adjacent Asn side chain carbonyl oxygen, and the second hydrogen bonds to a neighboring main chain carbonyl oxygen. The Asn ladder is located at the inner juncture of two  $\beta$  sheets; three of the Asns are invariant in all pectate lyases, pectin lyases, and plant pollen and style proteins. (B) An aromatic ring and an aliphatic stack in the interior of the parallel  $\beta$  helix. The aromatic rings are tilted 50° relative to the polypeptide backbone but maintain an interplanar ring distance of 3.6 Å. Within the parallel  $\beta$  helix core, there are six examples of aliphatic stacks that include either Val, Ala, Ile, or Leu. The aliphatic stacks range in size from three to eight side chains. Another stacking motif not shown is a stack of Ser side chains that form the maximum number of hydrogen bonds. In all stacking motifs, the side chains form a linear arrangement parallel to the axis of the parallel  $\beta$  helix.

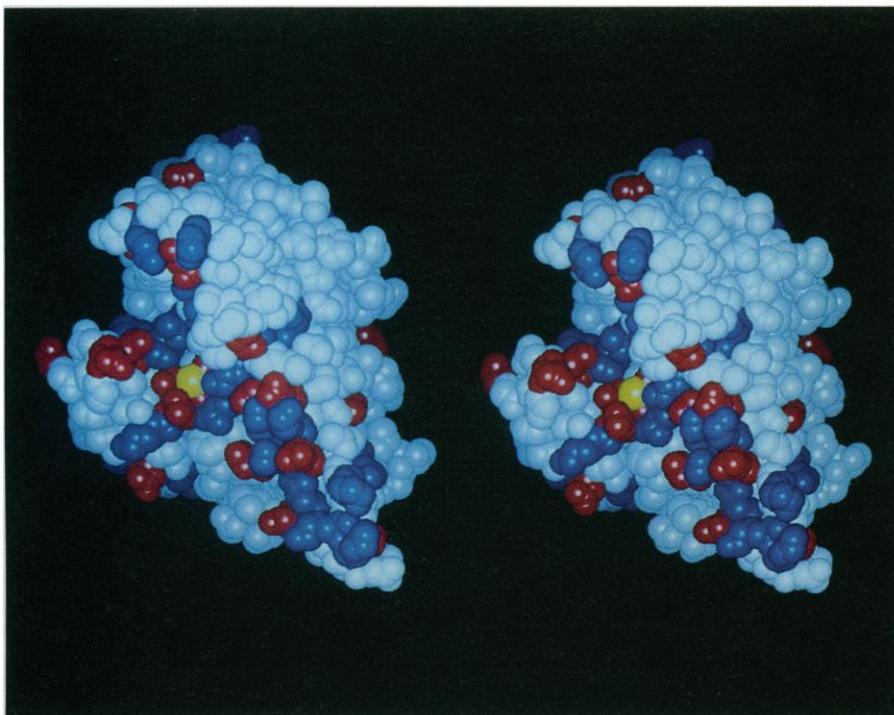
**Table 1.** Crystallographic data collection and phasing statistics (9). The figure of merit for 18,406 reflections from 2.5 Å to 50.0 Å is 0.57.

Data set	Native	Pt1	Pt2	U	Pb	Os
Resolution (Å)	2.2	2.5	2.5	2.5	2.5	2.5
Total observations	98,070	74,304	81,160	79,339	81,208	81,932
Unique reflections	27,769	34,025	32,616	33,956	34,826	33,549
Total percent possible	95.1	90.4	86.2	97.5	92.0	88.7
$I/\sigma$	11.4	11.3	10.4	6.6	11.4	13.5
$R_{\text{merge}}^*$ (unedited)	6.6	8.9	8.9	12.0	8.2	8.2
$R_{\text{merge}}^*$ (edited)	6.0	5.2	6.5	6.4	6.0	6.1
$R_{\text{scale}}^\dagger$		13.9	16.4	14.7	14.7	21.8
Centric $R_{\text{f}}^\ddagger$		0.71	0.67	0.65	0.75	0.69
$F_{\text{H}}/E^\S$		0.72	1.18	1.55	1.19	0.96

\* $R_{\text{merge}} = \sum (F_{\text{avg}} - F_{\text{obs}}) / \sum F_{\text{avg}}$ .  $\dagger R_{\text{scale}} = \sum |F_{\text{nat}} - F_{\text{der}}| / \sum (F_{\text{nat}} + F_{\text{der}})$ .  $\ddagger$ Centric  $R =$  Cullis  $R$  for centric reflections =  $\sum |F_{\text{pH}} - F_{\text{p}} - F_{\text{H}}| / \sum |F_{\text{pH}} - F_{\text{p}}|$ .  $\S F_{\text{H}}/E =$  ratio of heavy atom structure factor to the residual lack of closure, which equals the phasing power.

**Table 2.** Refined heavy atom derivative parameters. Occ., occupancy; Temp., temperature.

Heavy atom	Occ.	Atomic coordinates			Temp. factor
		X	Y	Z	
Pt1	0.262	0.073	0.234	0.193	22.13*
	0.040	0.288	0.032	0.884	29.23
Pt2	0.546	0.074	0.234	0.194	22.00
	0.038	0.060	0.472	0.161	23.63
U	0.023	0.091	0.325	0.183	17.63
	0.417	0.378	0.258	0.942	14.64
Pb	0.330	0.380	0.257	0.941	8.20
	0.024	0.352	0.525	0.661	27.97
Os	0.975	0.347	0.001	0.904	28.31



**Fig. 5.** Stereo view of a space-filling model PelC coded by electrostatic charge. The basic residues are shown in dark blue, the acidic residues in red, the neutral groups in light blue, and the putative  $\text{Ca}^{2+}$  site in yellow. The surface not shown is neutral. The narrow groove of surface charges approximates the shape and length of a dodecamer of a galacturonate polymer. From computer graphics experiments with a model substrate, the putative  $\text{Ca}^{2+}$  site is located between the third and fourth saccharide units from the nonreducing end. If the putative  $\text{Ca}^{2+}$  site approximates the active site region, then its location helps to explain the trimer as the limit product of enzymatic cleavage (22).

tate lyases share 49 invariant and an additional 29 highly conserved residues (3). Two well-known regions of strong homology correspond to 168-Ala-X-Asp-Ile-Lys-Gly-4X-Val-Thr-X-Ser-181 (region I) and 215-Val-2X-Arg-X-Pro-2X-Arg-X-Gly-2X-His-3X-Asn-232 (region II) in PelC (4). Both regions are spatially located around the putative  $\text{Ca}^{2+}$  binding site, with Asp<sup>170</sup>, Lys<sup>172</sup>, and Arg<sup>218</sup> found within its octahedral coordination sphere. Other invariant residues are broadly distributed throughout the structure. Because of their sequence and functional similarities, all extracellular pectate lyases are likely to differ only in the size and conformation of the protruding loops.

The extracellular pectate lyases also share regions of sequence similarity with fungal pectin lyases (16), plant pollen and style proteins (17), and the putative guanine nucleotide binding sequence in tubulins (18). Only the pectin lyases have a known related function: cleavage of the methylated form of polygalacturonate. The sequence similarities among the pathogenic factors and plant homologs include but are not limited to the conserved regions I and II. Other invariant and conserved amino acids cluster in a region on the external parallel  $\beta$  helix surface, which is located approximately  $180^\circ$  from the putative  $\text{Ca}^{2+}$

site in PelC and is covered by a large external loop. Most homologous amino acids are found on five consecutive turns of the parallel  $\beta$  helix, preceding and including the PelC asparagine ladder. No biochemical data indicate whether the region represents an essential structural feature of the parallel  $\beta$  helix, a secretion signal, or some unidentified function shared by the pathogenic factors and plant homologs. The similarities shared by all pectate lyases and the tubulins include two regions rich in glycine, Lys-Gly-4X-Gly and Gly-Gly-3X-Gly; an Asp-2X-Gly(Ala) region; and an Asp-X-Lys-Asn region that has the reverse pattern of the guanine nucleotide motifs in guanosine triphosphatase proteins (18). The results of computer graphics experiments, in which a guanine nucleotide was docked to PelC, suggest that a reasonable binding pocket is formed by the 125-Gly-Gly-3X-Gly-130, 170-Asp-2X-Gly-173, and 269-Asp-X-Lys-Asn-272 sequences of PelC. It is not known whether any pectate lyases bind nucleotides during secretion or pathogenic function.

#### REFERENCES AND NOTES

1. A. Collmer and N. T. Keen, *Annu. Rev. Phytopathol.* **24**, 383 (1986); A. Kotoujansky, *ibid.* **25**, 405 (1987); M. C. M. Perombelon and A. Kel-

- man, *ibid.* **18**, 361 (1980).
2. M. S. Crawford and P. E. Kolattukudy, *Arch. Biochem. Biophys.* **258**, 196 (1987).
3. J. C. D. Hinton, J. M. Sidebotham, D. R. Gill, G. P. C. Salmond, *Mol. Microbiol.* **3**, 1785 (1989).
4. S. J. Tamaki, S. Gold, M. Robeson, S. Manulis, N. T. Keen, *J. Bacteriol.* **170**, 3468 (1988).
5. N. T. Keen *et al.*, *ibid.* **159**, 825 (1984).
6. N. T. Keen and S. Tamaki, *ibid.* **168**, 595 (1986).
7. F. Barras, K. K. Thurn, A. K. Chatterjee, *Mol. Gen. Genet.* **209**, 319 (1987).
8. M. D. Yoder, D. A. DeChaine, F. Journak, *J. Biol. Chem.* **265**, 11429 (1990).
9. The platinum (Pt in Tables 1 and 2) and osmium (Os) derivatives were made by transfer of crystals to 1.80 M  $(\text{Li})_2\text{SO}_4$ , 0.09 M Hepes (pH 6.9), and several grains of  $\text{Pt}(\text{NH}_3)_2\text{Br}_2$  for 1 day (Pt1) and for 7 days (Pt2) or 5 mM  $\text{K}_2[\text{OsCl}_6]$  for 1 day. The uranyl (U) derivative was made by the soaking of a crystal in 2 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.10 M Hepes (pH 6.9), and 0.20 mM  $\text{U}(\text{NO}_3)_2$  for 1 day. The lead (Pb) derivative was prepared by the soaking of a crystal in 40%  $(\text{NH}_4)_2\text{SO}_4$ , 0.10 M Hepes (pH 6.9), and 0.01 M  $\text{Pb}(\text{C}_2\text{H}_3\text{O}_2)_2$  for 1 day. We collected x-ray diffraction data at room temperature using a San Diego Multiwire Systems (SDMS) area detector installed on a rotating-anode x-ray generator with a graphite monochromator. Each data set was collected from a single crystal and processed with the SDMS software package [A. J. Howard, C. Nielsen, N.-H. Xuong, in *Methods of Enzymology*, H. W. Wyckoff, C. H. W. Hirs, S. N. Timasheff, Eds. (Academic Press, New York, 1985), vol. 114, pp. 452–472]. The structure was solved by MIR techniques with four heavy atom derivatives, three of which substituted at unusual sites. A fifth heavy atom derivative, lutetium, was solved but not used because the single site was essentially the same as the uranyl and lead derivatives and did not contribute significantly to the phasing information. Heavy atom derivative parameters were refined with HEAVY [T. C. Terwilliger and D. E. Eisenberg, *Acta Crystallogr.* **A 39**, 813 (1983)]. The initial electron density map was solvent-flattened [B. C. Wang, in *Methods of Enzymology*, H. W. Wyckoff, C. H. W. Hirs, S. N. Timasheff, Eds. (Academic Press, New York, 1985), vol. 115, pp. 90–112]. A partial  $\alpha$  carbon backbone was traced on minimaps and a polyalanine model composed of 151 residues was constructed with the use of FRODO [T. A. Jones, *ibid.*, p. 157]. Model phases were combined with heavy atom derivative phases with SIGMAA [R. Read, *Acta Crystallogr.* **A 42**, 140 (1986)], and a model of 254 polyalanines was constructed and used to generate an improved map, from which the final model was built. The model was improved by cycles of crystallographic refinement and manual model rebuilding. Refinement consisted of molecular dynamics refinement by the method of simulated annealing [A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987)] with X-PLOR [A. T. Brünger, *X-PLOR Manual, Version 3.0* (Yale University, New Haven, CT, 1992)].
10. C. A. McPhalen, N. C. J. Strynadka, M. N. G. James, *Adv. Protein Chem.* **142**, 77 (1991).
11. C. H. Evans, in *Biochemistry of the Elements*, E. Frieden, Ed. (Plenum, New York, 1990) vol. 8, p. 11.
12. J. S. Richardson and D. C. Richardson, in *Prediction of Protein Structure and the Principles of Protein Conformation*, G. D. Fasman, Ed. (Plenum, New York, 1989), pp. 1–98.
13. S. Y. He *et al.*, *J. Bacteriol.* **173**, 4310 (1991); S. Y. He, M. Lindeberg, A. Collmer, in *Biotechnology in Plant Disease Control*, I. Chet, Ed. (Wiley-Liss, New York, 1993), pp. 39–64.
14. S. Y. He *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1079 (1991); S. Y. He and A. Collmer, *J. Bacteriol.* **172**, 4988 (1990); G. Condemine *et al.*, *Mol. Microbiol.* **6**, 3199 (1992); M. Lindberg and A. Collmer, *J. Bacteriol.* **174**, 7385 (1992).
15. A. P. Pugsley, *Microbiol. Rev.* **57**, 50 (1993).
16. M. Kuster-Van Someren, thesis, University of Utrecht, the Netherlands (1991); C. Gysler *et al.*, *Gene* **89**, 101 (1990).

17. R. A. Wing *et al.*, S. McCormick, *Plant Mol. Biol.* 14, 17 (1989); K. A. Budelier *et al.*, *Mol. Gen. Genet.* 224, 183 (1990); T. Rafnar *et al.*, *J. Biol. Chem.* 266, 1229 (1991); H. J. Rogers *et al.*, *Plant Mol. Biol.* 20, 493 (1992).
18. E.-M. Mandelkow, K. Linse, E. Mandelkow, in *The Guanine-Nucleotide Binding Proteins*, L. Bosch, B. Kraal, A. Parmeggiani, Eds. (Plenum, New York, 1989), vol. 165, pp. 385-390; K. Farrell and F. Journak, unpublished observations.
19. T. A. Jones, J.-Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr. A* 47, 110 (1991).
20. R. Luthy *et al.*, *Nature* 356, 83 (1992).
21. C. M. Wilmot and J. M. Thornton, *Protein Eng.* 3, 479 (1990).
22. J. F. Preston, J. D. Rice, L. O. Ingram, N. T. Keen, *J. Bacteriol.* 174, 2039 (1992).
23. We thank K. Farrell for discussions regarding the PeIC similarities with tubulin and the Academic Computing Graphics and Visual Imaging Lab, University of California, Riverside, for help generating Fig. 5. Supported by U.S. Department of Agriculture award 91-37303-6645 and the San Diego Supercomputer Center.

24 December 1992; accepted 16 April 1993

## Feedback Regulation Mechanisms for the Control of GTP Cyclohydrolase I Activity

Toshie Harada, Hiroyuki Kagamiyama, Kazuyuki Hatakeyama\*

Guanosine triphosphate (GTP) cyclohydrolase I, the rate-limiting enzyme in the biosynthesis of tetrahydrobiopterin (BH<sub>4</sub>), is subject to feedback inhibition by BH<sub>4</sub>, a cofactor for phenylalanine hydroxylase. Inhibition was found to depend specifically on BH<sub>4</sub> and the presence of another protein (p35). The inhibition occurred through BH<sub>4</sub>-dependent complex formation between p35 protein and GTP cyclohydrolase I. Furthermore, the inhibition was specifically reversed by phenylalanine, and, in conjunction with p35, phenylalanine reduced the cooperativity of GTP cyclohydrolase I. These findings also provide a molecular basis for high plasma BH<sub>4</sub> concentrations observed in patients with hyperphenylalaninemia caused by phenylalanine hydroxylase deficiency.

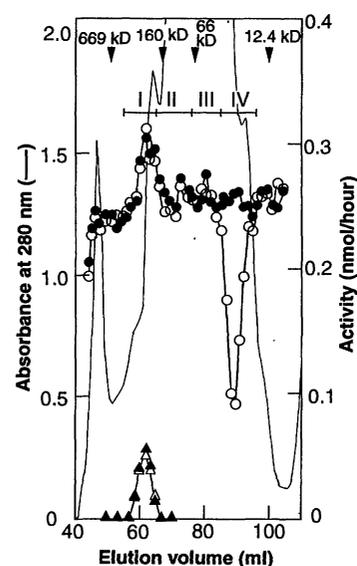
The enzyme GTP cyclohydrolase I (E.C. 3.5.4.16) catalyzes the conversion of GTP to dihydroneopterin triphosphate, the first step of BH<sub>4</sub> biosynthesis in animals (1). One function of BH<sub>4</sub> is as a cofactor for phenylalanine hydroxylase as well as for tyrosine hydroxylase, tryptophan hydroxylase, the O-alkylglycerolipid cleavage enzyme (1), and nitric oxide synthases (2). In contrast to the other cofactors, BH<sub>4</sub> is one of the regulators of these enzymic reactions (3) and is thus implicated in the control of phenylalanine catabolism and of neural and immune functions. The intracellular BH<sub>4</sub> concentration is controlled by the rate of its de novo biosynthesis (1), mainly at the step catalyzed by GTP cyclohydrolase I. This enzyme is the rate-limiting enzyme in the synthetic pathway, and it has been suggested that its activity is regulated at transcriptional (4) and substrate (5, 6) levels. The activity of GTP cyclohydrolase I in the liver is thought to be regulated by feedback inhibition. The enzyme activity in crude rat liver extracts is inhibited by BH<sub>4</sub> (7), and the in vivo occurrence of such feedback inhibition is strongly suggested by clinical observation of patients with an inherited disorder of BH<sub>4</sub> metabolism (8). However, the biochemical mechanisms underlying the feedback inhibition of this enzyme by BH<sub>4</sub> remain obscure.

In the course of our studies on GTP cyclohydrolase I (5, 6, 9), we noticed that rat GTP cyclohydrolase I expressed in *Escherichia coli* (10) was not inhibited by BH<sub>4</sub>. The recombinant rat enzyme could not be distinguished from the enzyme purified from rat liver on the basis of native molecular weight, subunit molecular weight, or ion spray mass spectrometric analysis of the molecular weights of peptides obtained from lysyl-endopeptidase digests of the two enzymes. Consistent with this finding, we determined that GTP cyclohydrolase I purified from rat liver was also not inhibited by BH<sub>4</sub>, although the enzyme activity in crude liver extracts was inhibited, as reported (7). In order to determine if a dissociable inhibitory factor existed, we fractionated the 100,000g supernatant obtained from rat liver homogenates by gel filtration. Endogenous GTP cyclohydrolase I was no longer inhibited by BH<sub>4</sub> (Fig. 1, triangles), in sharp contrast to the inhibition of enzyme activity noted in the crude extracts. We found that the fractions that were eluted around a volume of about 88 ml inhibited GTP cyclohydrolase I activity in a BH<sub>4</sub>-dependent manner (Fig. 1). The inhibitory factor was sensitive to digestion with proteinase K and trypsin and was insensitive to deoxyribonuclease I and ribonuclease, suggesting that it was a protein. We combined the fractions containing the inhibitory protein and used them for further characterization of this protein, which was designated as the p35 fraction according to its elution

volume corresponding to a molecular size of 35 kD on Superdex 75 HR column chromatography (11).

The p35 fraction was incubated with recombinant GTP cyclohydrolase I. The reaction was linear up to 60 min both in the presence and in the absence of BH<sub>4</sub>. In the presence of BH<sub>4</sub>, the p35 fraction inhibited GTP cyclohydrolase I activity in a dose-dependent manner, but no inhibition was observed in the absence of BH<sub>4</sub>. On the basis of these findings, we estimated that the total inhibitory activity contained in the p35 fraction derived from 1.0 g of liver was capable of inhibiting GTP cyclohydrolase I activity by 59 nmol/hour. Because the total GTP cyclohydrolase I activity recovered from the same amount of liver tissue was 10 nmol/hour, rat liver contained enough inhibitory activity to totally inhibit the activity of GTP cyclohydrolase I.

We examined the specificity of the inhibition of GTP cyclohydrolase I activity by p35 (12). A median effective concentration (EC<sub>50</sub>) of BH<sub>4</sub> for the inhibition of GTP cyclohydrolase I by p35 was 2 μM at a GTP concentration of 0.1 mM, and the effect of BH<sub>4</sub> reached its maximum at about 7 μM. This range of effective concentrations is very similar to the hepatic concentration of BH<sub>4</sub>, which was estimated to be approximately 6 μM from the reported value of 1.6 ± 0.4 μg per gram (wet weight) of rat liver (13), and



**Fig. 1.** Gel filtration chromatography data of the 100,000g supernatant of rat liver homogenates (28). Plotted are endogenous GTP cyclohydrolase I activity in the presence ( $\Delta$ ) or absence ( $\blacktriangle$ ) of 20  $\mu$ M BH<sub>4</sub> and inhibitory activity in the presence ( $\circ$ ) or absence ( $\bullet$ ) of 20  $\mu$ M BH<sub>4</sub>. The fractions were combined as indicated by the Roman numerals. Fraction IV (p35 fraction) was concentrated nearly tenfold. Calibration standards: thyroglobulin, 669 kD;  $\gamma$ -globulin, 160 kD; bovine serum albumin, 66 kD; cytochrome c, 12.4 kD.

Department of Biochemistry, Osaka Medical College, Takatsuki, Osaka 569, Japan.

\*To whom correspondence should be addressed.