Plant Physiol. (1994) 106: 849–862

The Three-Dimensional Structure of Pectate Lyase E, a Plant Virulence Factor from *Erwinia chrysanthemi*[1]

Susan E. Lietzke, Marilyn D. Yoder, Noel T. Keen, and Frances Jurnak*

Department of Biochemistry (S.E.L., M.D.Y., F.J.), and Department of Plant Pathology (N.T.K.), University of California, Riverside, California 92521

The three-dimensional structure of pectate lyase E (PelE) has been determined by crystallographic techniques at a resolution of 2.2 Å. The model includes all 355 amino acids but no solvent, and refines to a crystallographic refinement factor of 20.6%. The polypeptide backbone folds into a large right-handed cylinder, termed a parallel β helix. Loops of various sizes and conformations protrude from the central helix and probably confer function. A putative Ca^{2+}-binding site as well as two cationic sites have been deduced from the location of heavy atom derivatives. Comparison of the PelE and recently determined pectate lyase C (PelC) structures has led to identification of a putative polygalacturonate-binding region in PelE. Structural differences relevant to differences in the enzymatic mechanism and maceration properties of PelE and PelC have been identified. The comparative analysis also reveals a large degree of structural conservation of surface loops in one region as well as an apparent aromatic specificity pocket in the amino-terminal branch. Also discussed is the sequence and possible functional relationship of the pectate lyases with pollen and style plant proteins.

Pels (EC 4.2.2.2) are extracellular enzymes secreted by pathogenic organisms affecting plants. The enzymes degrade the pectate component of the middle lamella and cell wall of higher plants and have been shown to be causally involved in soft rot diseases. The soft rot *Erwinia* spp., which are gram-negative bacteria, secrete several types of pectic enzymes of which the most important for virulence are the Pels. *Erwinia* spp. typically produce multiple isozymes of independently regulated extracellular Pels. The extracellular isozymes belong to two gene families, *pelADE* and *pelBC*. There is 50 to 85% amino acid sequence identity within a family and 22% or lower identity between the families (Tamaki et al., 1988; Hinton et al., 1989; S. Heffron and F. Jurnak, unpublished observations). The *pelADE* family contains one disulfide bond and the *pelBC* family contains two. The *Erwinia* Pels also share regions of sequence similarity with fungal pectin lyases and pollen and style plant proteins (Wing et al., 1989; Kuster-Van Someren, 1991; Rafnar et al., 1991). A third gene family of intracellular Pels is also produced in some pathogenic species but is unrelated in sequence to the extracellular families (Hinton et al., 1989). Pels are exported through the inner membrane by the secretory-dependent pathway and the extracellular forms are transported across the outer membrane by products of the *out* gene family (He et al., 1991a, 1991b; Condemine et al., 1992; Lindeberg and Collmer, 1992).

Pels cleave α-1,4-linked galacturonic acid residues of the pectate component of the plant cell wall by a β-elimination mechanism that generates an unsaturated galacturonosyl residue at the nonreducing end (Collmer and Keen, 1986). The *Erwinia chrysanthemi* enzymes act preferentially on unmethylated pectin, cleaving internal glycosidic linkages relatively infrequently and then exolytically cleaving exposed termini more frequently (Preston et al., 1992). The different Pel isozymes catalyze the cleavage reaction at pH optima ranging from 8 to 11, yielding primarily either dimer (*pelADE* family) or trimer (*pelBC* family) end products. In addition to mechanistic differences, the Pel enzymes differ considerably in their host range and relative maceration properties, with PelE considered to be the most virulent of the *E. chrysanthemi* EC16 Pel enzymes (Tamaki et al., 1988). Calcium is required for in vitro activity of all Pels, but it has not been established whether calcium binds to the enzyme or simply cross-links the strands of the substrate, PGA, into a structure that is recognized by Pels (Crawford and Kolattukudy, 1987). The mechanism of substrate binding and the molecular basis of catalysis is not yet known.

PelE consists of 355 amino acids and has a mol wt of 38,069, as calculated from the amino acid sequence. The PelE structure folds into a parallel β helix in which the strands of three parallel β sheets fold into a large central coil. The first x-ray structure of a protein in which the parallel β helix topology was observed is *E. chrysanthemi* PelC (Yoder et al., 1993a). Although PelE has only 22% sequence identity with PelC, the core structure of both proteins is strikingly similar (Yoder et al., 1993b). There are significant structural differences in the various loops that protrude from and cover the parallel β helix core of each isozyme. There are also pronounced similarities in the loop conformations in one region

[1] F.J. was supported by the United States Department of Agriculture (award No. MCB9408999).

* Corresponding author; fax 1–909–787–3590.

Abbreviations: B factor, isotropic temperature factor; crystallographic R factor, agreement factor between observed structure factor amplitudes, $|F_o|$, and calculated structure factor amplitudes, $|F_c|$, based on atomic model; F, structure factor; F_c, calculated structure factor; F_o, observed structure factor; MIR, multiple isomorphous replacement; Pel, pectate lyase; PGA, polygalacturonic acid; r.m.s., root mean square; ϕ, rotation angle about the carbon α-nitrogen bond; ψ, rotation angle about the carbon α-carbon bond; χ, rotation angle about the carbon α-carbon β bond; σ, standard deviation.

and in an apparent aromatic specificity pocket. The differences between the structures of the Pels may explain differences in the enzymatic and maceration properties of the isozymes, and the similarities may be related to folding or secretion mechanisms.

MATERIALS AND METHODS

Materials

PEG 4000 was obtained from Koch-Light, Ltd. (Buckinghamshire, UK), and PEG 3350 was from J.T. Baker (Phillipsburg, NJ). The heavy atom compound $LaCl_3$ was purchased from Alfa (Ward Hill, MA), and $UO_2(NO_3)_2$ was from Fisher. All other reagents were obtained from Sigma.

Crystal Preparation

For the present study, PelE was purified in quantity by scaling up the protocol described by Keen and Tamaki (1986). PelE was isolated from the periplasm of *Escherichia coli* cells containing the high-expression construct *pPEL748* of the *Erwinia chrysanthemi* EC16 *pelE* gene. The mature, secreted recombinant protein has the same mol wt, isoelectric point, and maceration properties as PelE isolated from *E. chrysanthemi*. Small crystals of PelE were grown initially from PEG 4000 by vapor diffusion techniques similar to the conditions reported by Kim et al. (1989), with the exception that $CaCl_2$ and the PGA substrate were omitted. PelE crystals for data collection were grown from microseeds using the technique described by Fitzgerald and Madsen (1986). The largest crystals grew in 2 weeks at 4°C from a solution of 18 to 19% PEG 4000, 17 mM Li_2SO_4, and 50 mM Tris-HCl, pH 8.0. The diffraction patterns of PelE crystals were consistent with the space group $P2_12_12_1$, with unit cell parameters of a = 38.79 Å, b = 91.14 Å, and c = 102.99 Å. The solvent content was estimated to be 48% using Matthews' method (1968) with one molecule per asymmetric unit.

With the exception of the La^{3+} derivative, all isomorphous heavy atom derivatives were prepared by diffusing the heavy atom into crystals at 4°C. Prior to diffusion, the crystals were transferred from mother liquor into a solution of 33% PEG 3350 and 50 mM Tris-HCl, pH 8.0, and then the heavy atom was added. For the present studies, a PelE crystal was soaked in 2 mM $UO_2(NO_3)_2$ for 3 d, in 4 mM K_2PtCl_4 for 3 d, or in 5 mM KI and Iodogen for 2 h. The La^{3+} derivative was prepared by adding 0.6 mM $LaCl_3$ to the initial crystallization conditions.

Data Collection and Reduction

X-ray diffraction data were collected at 20°C using a dual-chamber San Diego Multiwire Systems (San Diego, CA) area detector system installed on a Rigaku (Danvers, MA) rotating-anode x-ray generator with a Supper (Natick, MA) graphite monochromator operated at 45 kV × 120 to 150 mA. Each x-ray data set was collected from a single crystal at helium-path distances of 615 mm and 560 mm between the crystal and each detector and processed with the San Diego Multiwire Systems software package (Howard et al., 1985). Derivative x-ray diffraction data sets were scaled, using the Fourier-Bessel program of Terwilliger and Eisenberg (1983), to the native x-ray data set, which had been placed on an absolute scale by Wilson's method (Wilson, 1942). The x-ray data collection statistics are summarized in Table I.

MIR Phase Determination

Four heavy atom derivative data sets were used to determine the phases by the MIR method. All sites were located by difference Pattersons using the program FFT (Ten Eyck et al., 1976) and confirmed by cross-phasing the difference Fourier of each derivative by alternate derivatives. Heavy atom parameters were refined by HEAVY (Terwilliger and Eisenberg, 1983) and the absolute configuration of the molecule was determined based on cross-Fourier calculations as described by Blundell and Johnson (1976). The MIR phases were applied to the native structure factor amplitudes to calculate an initial electron density map at 3.2- to 50.0-Å resolution. Although the parallel β helix topology was very clear in the initial map, the connectivity of the electron density in the polypeptide loop regions was improved by solvent density modification techniques (Wang, 1985). A partial αC backbone was traced on minimaps and a polyal-

Table I. *Summary of x-ray crystallographic data collection*

Parameter	Native	UO_2^{2+}	Pt^{2+}	La^{3+}	I^-
Resolution (Å)	2.2	2.2	2.6	2.2	2.2
Total observations	65,204	50,955	64,239	63,239	64,235
Unique reflections[a]	18,548	30,982	31,328	31,373	31,403
Total possible (%)	96.5	87.7	93.6	96.8	93.1
⟨I/sigma⟩	14.11	11.00	8.49	14.23	11.23
R_{sym} (unedited)[b]	5.39	8.84	7.58	5.82	7.48
R_{sym} (edited)[b]	4.68	6.11	5.58	4.23	6.88
R_{scale}[c]		29.51	12.38	12.36	17.49

[a] Bijvoets merged for native only. [b] $R_{sym} = \Sigma \mid (I_{avg} - I_{obs}) \mid / \Sigma I_{avg}$ where I is the average (avg) or the observed (obs) intensity of the reflection. Unedited is before manual rejection of reflections. Edited is after manual rejection of outlying observations of a reflection. [c] $R_{scale} = 2\left(\frac{\Sigma \mid \mid F_{nat} \mid - \mid F_{der} \mid \mid}{\Sigma \mid F_{nat} \mid + \mid F_{der} \mid}\right)$ where F is the structure factor of the native (nat) or derivative (der) data.

anine model composed of 207 residues was constructed on an Evans and Sutherland (Salt Lake City, UT) PS390 with the program FRODO (Jones, 1985). Partial model phases were combined with heavy atom derivative phases using SIGMAA (Read, 1986) and applied to the structure factors at 2.5 Å; the procedure was reiterated three times. From the series of improved maps, the amino acid side chains were fitted to the electron density and an initial model consisting of 351 of the 355 amino acids was constructed.

Model Refinement

The model was refined with molecular dynamics techniques using the method of simulated annealing (Brünger et al., 1987) with X-PLOR (Brünger, 1993b). After each cycle of crystallographic refinement, the model was checked and manually rebuilt. The final four residues were built into the model from a $2F_o$-F_c map, calculated from model phases, after the first cycle of refinement. Three additional cycles of crystallographic and individual isotropic temperature factor refinement were carried out. Data between 2.2- and 5.0-Å resolution were used in the first two cycles of refinement, and data between 2.2 and 8.0 Å were used in the last two cycles of refinement. Throughout the refinement process, clear density was observed for most of the amino acids in a $2F_o$-F_c electron density map, contoured at 1σ. There were two exceptions, which were located on surface loops with weak density. The entire model was checked by inspection of annealed omit maps calculated by X-PLOR (Bhat and Cohen, 1984; Hodel et al., 1992).

The atomic coordinates are deposited with the Brookhaven Protein Data Bank under the file name 1PCL.

Model Analysis

The program PROCHECK (Laskowski et al., 1993) was used to calculate stereochemical parameters of the refined PelE structure. Secondary structural assignments were initially made with the Kabsch and Sander algorithm (Kabsch and Sander, 1983) in the PROCHECK package, followed by visual inspection of the main-chain hydrogen bonds. The criteria for a hydrogen bond was a donor-acceptor distance of less than 3.5 Å and an angle greater than 120° (Baker and Hubbard, 1984). The PelE and PelC models were compared with the program O (Jones et al., 1991; Jones and Kjeldgaard, 1993) on an Evans and Sutherland ESV graphics system. A least-squares refinement method in the LSQ option was used to superimpose the αC atoms in both structures. The initial alignment was made with the αCs of amino acids in the β strands of the parallel β helix core only.

The distance between the La^{3+} heavy atom site, assumed to be a putative Ca^{2+}-binding site, and an important invariant region, 144-vWiDH-148, was estimated in two ways. First, a minimum distance was estimated by measuring the distance between La^{3+} and the αC of Asp^{147} around the circumference of the parallel β helix. Second, the maximum distance and the number of galacturonic acid residues required to span the distance was estimated by rotating trigalacturonate between the heavy atom site and Asp^{147}, maintaining approximate van der Waals distances to protein atoms. The coordinates for trigalacturonate were obtained from Walkinshaw and Arnott (1981).

Loops were characterized according to the nomenclature proposed by Ring et al. (1992). Linear loops are referred to as strap loops, and nonlinear loops are subdivided into two categories. Omega loops are nonlinear, but planar loops. This definition of omega loops differs from Leszczynski and Rose's (1986) in that the ends of omega loops are not necessarily in close proximity. The second type of nonlinear loops are zeta loops, which are nonplanar.

Figure Preparation

The Ramachandran plot in Figure 1 was prepared with the program PROCHECK. In Figure 2, the $2F_o$-F_c map was calculated by X-PLOR and displayed on O; the x-utility program Oplot (G.J. Klegwegt, unpublished data) was then used to create the illustration. Figures 5, 6, 7, and 8 were produced with the program MOLSCRIPT (Kraulis, 1991). The space-filling models in Figure 9 were created with InSight (Biosym Technologies, Inc., San Diego, CA).

Table II. *Refined heavy atom derivative parameters for PelE in the resolution range of 2.2 to 50.0 Å*

Derivative	Occupancy	Atomic Coordinates			B (Å²)
		X	Y	Z	
UO_2^{2+}	0.411	0.371	0.357	0.259	13.3
	0.541	0.436	0.367	0.380	28.7
	0.289	0.599	0.427	0.145	31.9
Pt^{2+a}	0.153	0.379	0.024	0.327	31.2
La^{3+}	0.456	0.385	0.357	0.260	31.2
I^-	0.351	0.805	0.309	0.288	24.5
	0.234	0.350	0.701	0.162	25.4
	0.310	0.761	0.220	0.314	85.3
	0.176	0.399	0.154	0.266	17.9

[a] Heavy atom derivative was refined only to 3.2 Å.

RESULTS

Structural Determination

Four heavy atom derivative data sets produced nine heavy atom sites, eight of which were unique. The refined heavy atom parameters are listed in Table II. Although none of the individual sites were strongly substituted, the phasing power, F_h/E, was sufficient for an unambiguous interpretation of the electron density. The overall figure of merit for the MIR refinement was 0.54 for 12,589 reflections in the 2.5- to 50.0-Å resolution range. The MIR phasing statistics are presented in Table III. Each heavy atom position substituted at a chemically reasonable site. The three uranyl (UO_2^{2+}) sites substituted near carboxylic acid residues: uranyl1 at Asp^{134}, Asp^{173}, and Asp^{177}; uranyl2 at Glu^{229}; and uranyl3 at Glu^3. La^{3+} occupied the same site as uranyl1, and the four iodine sites substituted at Tyr^{123} (two sites), Tyr^{169}, and Tyr^{352}. The platinum derivative substituted at the amino terminus with low occupancy and was used only for phasing to a resolution of 3.2 Å.

The PelE model presented herein consists of 2,693 non-hydrogen atoms in all 355 amino acids, but no solvent molecules. With the exception of the tips of two extended loops, the main chain and most side chains are unambiguously assigned. The R factor for 15,182 reflections in the 2.2- to 8.0-Å resolution shell with $F > 2\sigma$ is 20.6%. The corresponding free R is 26.8% for the 10% of the data used as a test set (Brünger, 1993a). The geometry of the model is good, with r.m.s. deviations from ideality of 0.006 Å in bond length and 1.56° in bond angle. A summary of model refinement statistics is presented in Table IV. A Ramachandran plot of the backbone dihedral angles in Figure 1 illustrates that 81.9% of non-Gly residues lie in the most favored region and 17.7% lie in allowed regions. Only one non-Gly residue lies in a generously allowed region, Phe^{236} with ϕ and ψ angles of 62.5° and −27.7°. Phe^{236} is located in a short turn between two β strands and its electron density is very well resolved. Equally clear is the electron density for the single *cis*-Pro, Pro^{232}, which is analogous to *cis*-Pro^{220} at the same position in PelC. Stereochemical parameter analysis indicates that the side chains of the PelE model also have good geometry. The observed χ_1 torsion angles agree well with the preferred conformers g^- (+60°), t (+180°), and g^+ (−60°). The average χ_1 angles for 289 residues are 61.1° ± 10.6°, 183.5° ± 10.2°, and −62.9° ± 9.9°, respectively, which correlate well with the values calculated by Morris et al. (1992) for well-refined structures at high resolution.

Table IV. *PelE refinement statistics*

Resolution range 2.2–8.0 Å
Number of reflections ($F > 2\sigma$) 16,846
Number of non-H atoms/asymmetric unit 2,693
Crystallographic R factor ($F > 2\sigma$, 90% data), with no solvent molecules, 20.6%
Free R factor 26.9%
R factor ($F > 2\sigma$, 100% data) 21.0%
r.m.s. deviations from ideality
Bond length 0.006 Å
Bond angle 1.56°
Improper angle 1.090°
Average refined B factor (isotropic)
Main chain atoms 11.5 Å^2
All non-H atoms 12.5 Å^2

With the exception of the ends of two surface loops previously mentioned, the final $2F_o$-F_c electron density map, contoured at 1σ, shows continuous density for all main-chain atoms. A representative section of the $2F_o$-F_c map is shown in Figure 2. The density of most amino acid side chains, with the exception of some surface side chains, mostly Lys's, is well resolved. The ambiguous regions of the map are the tips of two extended loops, region 1 consisting of residues 124 to 127 and region 2 consisting of residues 164 to 167. There is consistently weak density in the two regions in all maps, indicating that these regions may have considerable flexibility. The main-chain atoms for residues 164 to 167 are reasonably clear in a σ_A weighted $2F_o$-F_c electron density map (Read, 1986). In contrast, there is minimal density for Ser^{125} and Gly^{126}, making it difficult to determine the orientations of the main chain as well as the side chains of Glu^{124} and Asp^{127}. Ser^{125} and Gly^{126} have ultimately been modeled on the basis of proper backbone dihedral angles and the minimization of potential close contacts with a symmetry-related molecule. The lack of density may be a result of multiple conformations for the tip of the loop. Confirmation of the correct direction for both loops is provided by the substitution of iodine at

Table III. *PelE MIR phasing statistics*

F_h/E, Ratio of heavy atom structure factor and the residual lack of closure. R_{cullis}, Cullis R for centric reflections = $\Sigma | F_{pH} - F_p - F_H | / \Sigma | F_{pH} - F_p |$ where F_p, F_{pH}, and F_H are the structure factors of the protein, derivative, and heavy atom, respectively. FOM, Figure of merit.

Resolution Range (Å)	UO_2^{2+}		La^{3+}		I^-		FOM
	F_h/E	R_{cullis}	F_h/E	R_{cullis}	F_h/E	R_{cullis}	
9.04	1.52	0.55	0.79	0.55	1.38	0.61	0.75
5.69	1.96	0.52	1.75	0.47	1.18	0.69	0.82
4.45	1.23	0.61	1.18	0.67	0.73	0.73	0.69
3.77	1.05	0.62	1.06	0.64	0.60	0.72	0.61
3.33	1.01	0.63	0.91	0.63	0.51	0.76	0.51
3.01	1.09	0.66	0.79	0.74	0.54	0.72	0.45
2.77	1.09	0.69	0.70	0.72	0.53	0.70	0.42
2.58	1.06	0.69	0.58	0.73	0.51	0.74	0.37
Total	1.20	0.60	0.94	0.61	0.67	0.71	0.54

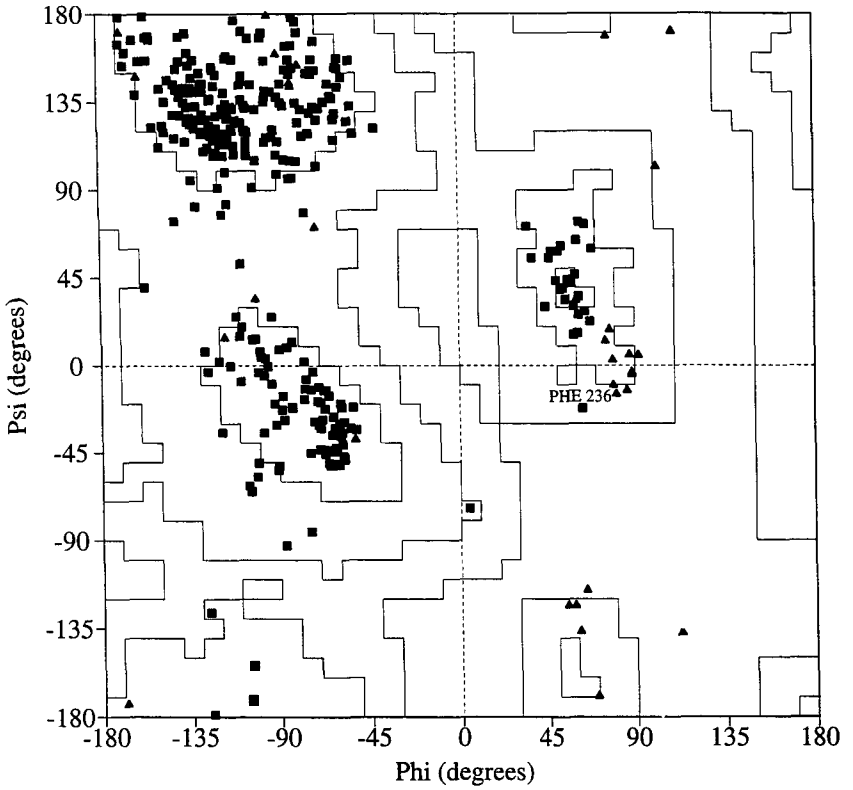

Figure 1. Ramachandran plot of the ϕ and ψ dihedral angles of the PelE polypeptide backbone. Gly residues are shown as triangles; all other residues are shown as squares. Phe^{236}, the only non-Gly residue outside of most favored or allowed regions, has ϕ and ψ angles of 62.5° and −27.7°. The plot was prepared with the program PROCHECK.

Tyr^{123} and Tyr^{169}, both near the most poorly defined regions of electron density. Two iodine positions were substituted near Tyr^{123}, possibly reflecting alternate conformations of the residue due to flexibility of the side chain and/or region 1.

Because the topology for PelE is unusual, additional measures were considered to substantiate the validity of the model. The average real space correlation coefficient, calculated by O, for main-chain atoms is 0.905. Figure 3 shows that the real space correlation coefficient is good for all parts of the model except the two flexible loop regions, residues 124 to 127 and 164 to 167. The average temperature factor for main-chain atoms is 12.5 $Å^2$; a plot of the average B factor for each residue is shown in Figure 3. The highest peaks in the plot correspond to the loop regions, and the valleys correspond to the core of the parallel β helix. Amino acids within the two surface loops with ill-defined electron density have the highest average main-chain B factors, 55 and 45 $Å^2$. Finally, the model was characterized by a three-dimensional window profile analysis (Lüthy et al., 1992). Figure 4 shows that the PelE model has fairly high three-dimensional/one-dimensional profile scores. The overall three-dimensional/one-dimensional window profile score per number of residues was 167.3, which is comparable to the value for correct models reported by Lüthy et al. (1992).

Structural Results

The PelE molecule is a single structural domain with approximate dimensions of 53 × 52 × 41 Å. The core of the protein consists of three parallel β sheets whose individual parallel β strands coil into a large right-handed cylinder (Fig. 5). The folding topology is similar to that first observed for PelC (Yoder et al., 1993a) and is termed a parallel β helix.

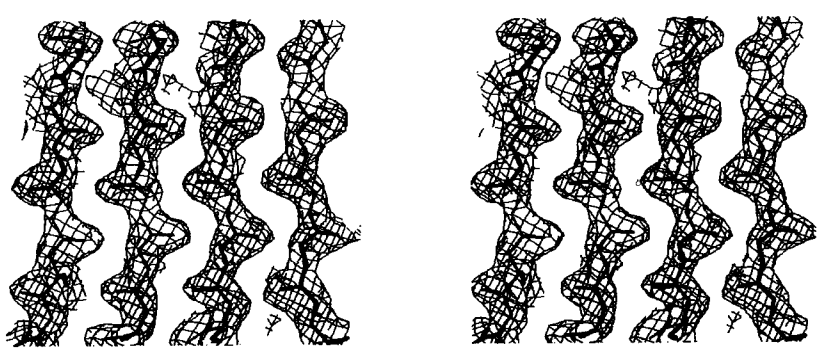

Figure 2. Representative portion of the $2F_o$-F_c electron density map of PelE contoured at 1σ. The backbone of several parallel β strands is superimposed on the electron density map. All carbonyl oxygens are visible in the final electron density map. In this section, the carbonyl groups lie in the plane of the paper and the αC-βC vector is parallel to the view direction.

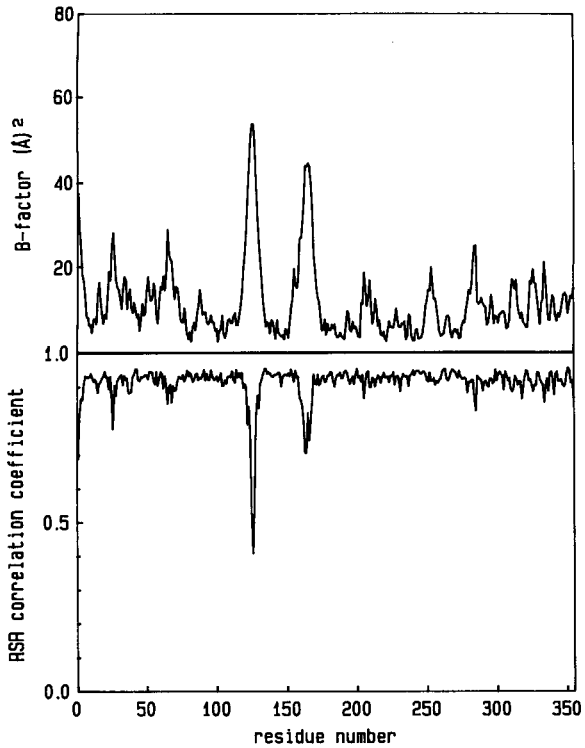

Figure 3. Average temperature factors of the main-chain atoms of PelE, plotted as a function of residue number. The residues in flexible loop regions have the highest average temperature factors. In the lower plot, the real-space correlation coefficients of main-chain atoms of PelE are plotted as a function of residue number. Correlation coefficients were calculated with the program O. The dips in the plot correspond to two loop regions with poorly defined density.

Within PelE, there are seven complete turns in the parallel β helix. A schematic of the PelE polypeptide backbone, highlighting the β structural features, is shown in Figure 5B. Only those amino acids with repetitive ϕ and ψ angles of classical β structure and maximal hydrogen bond formation between the amide and carbonyl groups on adjacent strands are considered to be β structure. The three parallel β sheets consist of 8, 10, and 6 strands. Within each β sheet, the β strands are relatively short, ranging from two to five residues in length and generating only a small, right-handed twist. As a consequence of the packing arrangement of the parallel β sheets, the cross-section of the parallel β helix is not circular but L-shaped. Two of the β sheets form a parallel β sandwich, with the third parallel β sheet oriented approximately 110° to the parallel β sandwich.

Polar and hydrophobic side chains fill the interior of the protein core and many are involved in extensive stacking interactions. Short stacks of two or three aliphatic amino acids are frequently found on the interior of the parallel β helix. One long stack of Ile's, Ile^{100}, Ile^{138}, Ile^{178}, and Ile^{201}, is found in the interior (fig. 3c, Yoder et al., 1993b). Another extended interior stack is composed of aromatic residues, Phe^{192}, Phe^{224}, Tyr^{246}, and Phe^{276} (fig. 9a, Yoder et al., 1993b). The aromatic planes are oriented in a manner that resembles base-pair stacking interactions in double-stranded DNA structures. The inter-ring distance is 3.6 Å and the planes of the aromatic side chains are nearly parallel. A third type of stack involves the polar residues Ser and Asn. One example is a short, interior stack composed of Ser^{141} and Ser^{182}. The second example is the ladder composed of Ser^{190}, Asn^{222}, Asn^{244}, Asn^{274}, and Ser^{307} (fig. 7b, Yoder et al., 1993b). Both of the stacks form extensive networks of hydrogen bonds that stabilize tight turns connecting β strands. The latter stack is analogous to the six-residue Asn ladder observed in PelC and described in more detail by Yoder et al. (1993b).

Although the structural core of the protein is very regular, no repetition in the primary sequence is detected as a consequence of the variation in the size of the loops that protrude from the central core in the peptide connections between two β strands. One of the three peptide segments connecting the β strands in each helical turn forms a regular secondary structural element consisting of two amino acids. There are seven examples of the unique bend in PelE, in which the first amino acid has average ϕ and ψ angles of 55.7° and 39.5° and the second amino acid has angles of −106.09° and 160.75°, respectively. The bend resembles a distorted $\gamma\beta_E$ turn, with the exception that the αC_{i-1}-αC_{i+2} is 9.15 Å in the PelE structure and 7 Å or less in the rare $\gamma\beta_E$ turn categorized by Wilmot and Thornton (1990). The remaining peptide connections between the β strands in each β helical turn form loops of various sizes and conformations. The loops range from 4 to 23 residues in length. Some loops contain short segments of α or 3_{10} helices and others fold into β arches that form many main- and side-chain hydrogen bonds with the main-chain atoms in the core regions of the parallel β helix. The loops cover approximately 25% of the surface of the parallel β helix and cap the ends of the parallel β helix. The N-terminal end of the parallel β helix is covered by an α helix and the C-terminal end is capped by two loops that are connected by a disulfide bond between Cys^{291} and Cys^{320}. The disulfide bridge has a left-hand conformation with a χ_3 angle of −87.7°.

The structurally conserved N- and C-terminal loops run parallel to the axis of the parallel β helix and meet in a

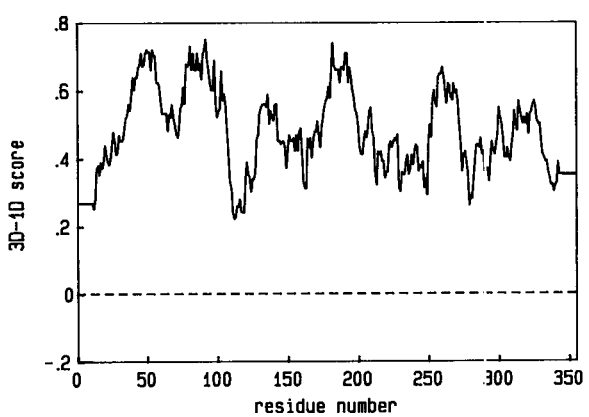

Figure 4. Three-dimensional window profile plot. The average three-dimensional/one-dimensional scores for residues in a 21-residue sliding window are plotted by residue number.

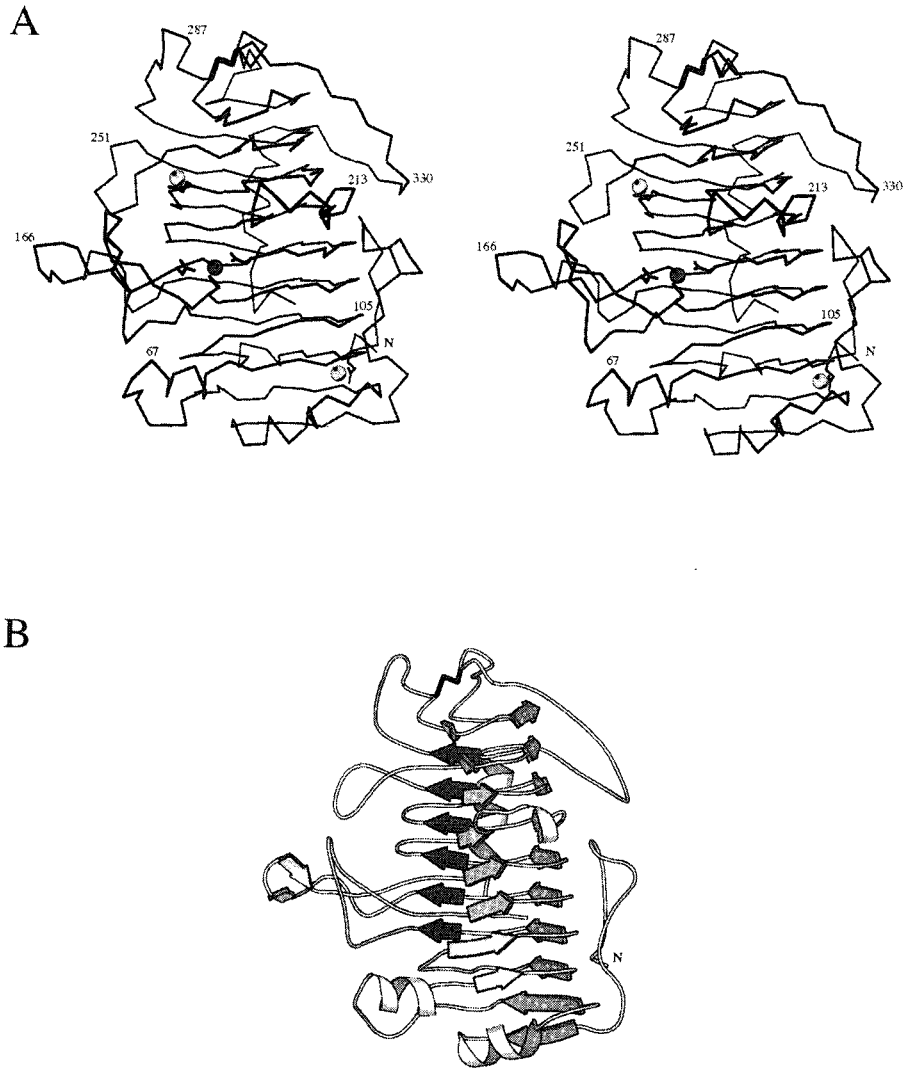

Figure 5. A, Stereo view of the αC backbone of PelE. The disulfide bond between Cys^{291} and Cys^{320} is indicated in black. The putative Ca^{2+} ion is shown by the black sphere, and the two uranyl sites are shown by light gray spheres. The coordinating ligands Asp^{134}, Asp^{173}, and Asp^{177} (Ca^{2+}); Glu^{229} (uranyl2); and Glu^{3} (uranyl3) are illustrated in black. B, Schematic ribbons diagram of the PelE polypeptide backbone. The three parallel β sheets are shown as light, medium, and dark gray, respectively.

midregion section. The N terminus, residues 5 to 22, is a zeta-type loop. The C terminus folds into a large omega loop, residues 323 to 339, followed by a 10-residue α helix and a short hook composed of the last 6 residues. Not only do the surface loops protect hydrophobic regions from solvent interactions, but the loops appear to form the functional regions. For example, the most distinctive feature of the PelE surface is a large groove (Fig. 5) formed by five loops, residues 114 to 135, 153 to 175, 201 to 216, 248 to 260, and 297 to 301. The groove contains two cationic binding sites and is analogous to the putative saccharide-binding region in PelC (Yoder et al., 1993a).

Ca^{2+} is required for in vitro activity of all extracellular Pels, but whether Ca^{2+} binds the substrate or the protein has not been established. The locations of heavy atom sites on PelE suggest several potential cationic sites, including a possible Ca^{2+} site. Although Ca^{2+} is not present in the PelE crystals, La^{3+} and uranyl were selected as derivatives because they frequently substitute at Ca^{2+} sites. The single La^{3+} site substituted at the same position as one of the uranyl sites and at a location analogous to the putative Ca^{2+}-binding site in PelC. The site lies in the major surface groove and is closely coordinated to the carboxylic acid groups of three Asp residues, Asp^{134}, Asp^{173}, and Asp^{177}, at distances ranging from 2.5 to 2.8 Å. The first two are invariant in all Pels. The third, Asp^{177} is invariant in the *pelADE* subfamily but is a conserved Glu in the *pelBC* subfamily. The second, strongly substituted uranyl site also lies in the major surface groove, approximately 12.8 Å from the putative Ca^{2+} site. Because only one ligand, Oε2 of Glu^{229}, directly coordinates the second uranyl

site at a distance of 2.4 Å, it is more likely to be a general cationic site than a Ca^{2+} site. The third uranyl site is only weakly substituted, with an occupancy of 28.9%. The site is located 2.5 Å from the coordinating atom, Oε2, of Glu^{3}, suggesting another potential cationic binding site.

Comparison of PelE and PelC

The three-dimensional structure of PelE was superimposed upon the PelC structure by first aligning the αC of the parallel β strands within the parallel β helix core and then minimizing the r.m.s. distances between αC in both models. The results of the superposition, illustrating the alignment of secondary structural elements and αC backbones, are shown in Figures 6 and 7. As evident in Figure 6, the core parallel β helix aligns very well, within an r.m.s. deviation of 1.02 Å for the αC of 81 amino acids (Yoder et al., 1993b). The r.m.s. deviation of the αC is 2° or less for 197 residues and 3° or less for 232 residues. Sixty eight of 198 residues with αC r.m.s. deviations of 2° or less are invariant between PelE and PelC. Both Pels form seven complete turns of the parallel β helix. The number of amino acids per β strand and the number of parallel β strands per β sheet vary only slightly between the two enzymes. As noted previously, most of the amino acids found on the interior of the core of PelE and PelC are involved in extensive side-chain stacking interactions. However, the length of the homologous stacks differs: the Asn ladder of six residues in PelC is only three in PelE, Asn^{222}, Asn^{244}, and Asn^{274}; the Ser stack of three residues in PelC is only two in PelE, Ser^{141} and Ser^{182}; and the aromatic stack of three residues in PelC is four in PelE, Phe^{192}, Phe^{224}, Tyr^{246}, and Phe^{276}. The heterologous hydrophobic stack of Ile^{107}, Val^{135}, and Ile^{171} in PelC is a homologous Ile stack of four residues in PelE, Ile^{100}, Ile^{138}, Ile^{178}, and Ile^{201}.

As illustrated in Figures 6 and 7, many of the loops that protrude from and cover the central parallel β helix occur at analogous positions in PelE and PelC. Eight loops are structurally conserved. Four of these loops are found in the N- or C-terminal regions, and include the N-terminal zeta loop, residues 9 to 22; the α helix covering the amino end of the parallel β helix, residues 35 to 41; a large omega loop, residues 328 to 339; and the C-terminal α helix and hook, residues 345 to 355. The remaining structurally conserved loops are short loops that protrude from the parallel β helix core. The degree of structural conservation is somewhat surprising because both the N-terminal and C-terminal loops have unusual conformations. Three of these four conserved loops are spatially adjacent to one another and meet in a midregion that is located on the opposite external surface of the parallel β helix as the putative Ca^{2+} site. In the N-terminal loop, the similarities between PelE and PelC extend beyond the peptide conformation to a structural cavity that appears to be a specificity pocket for an aromatic residue. In contrast to the results of primary sequence alignments (Hinton et al., 1989), Trp^{11} in PelE is structurally aligned with Tyr^{7} in PelC. The side chains of both the aromatic residues are located in pockets surrounded by a network of seven ringed residues, five of which are identical in PelE and PelC as well as highly conserved in the *pelADE* and *pelBC* subfamilies. The conserved amino acids in the network include His^{220}, His^{240}, Tyr^{242}, Tyr^{331}, and Tyr^{333}. Tyr^{325} is similar to a Tyr in PelC; however, its centroid is positioned 3 Å farther away from the centroid of the Tyr in PelC. Tyr^{184} is not conserved in PelC. As shown in Figure 8, most of the ringed residues interact in the edge-to-face manner that is common among proteins (Burley and Petsko, 1985; Singh and Thornton, 1985) rather

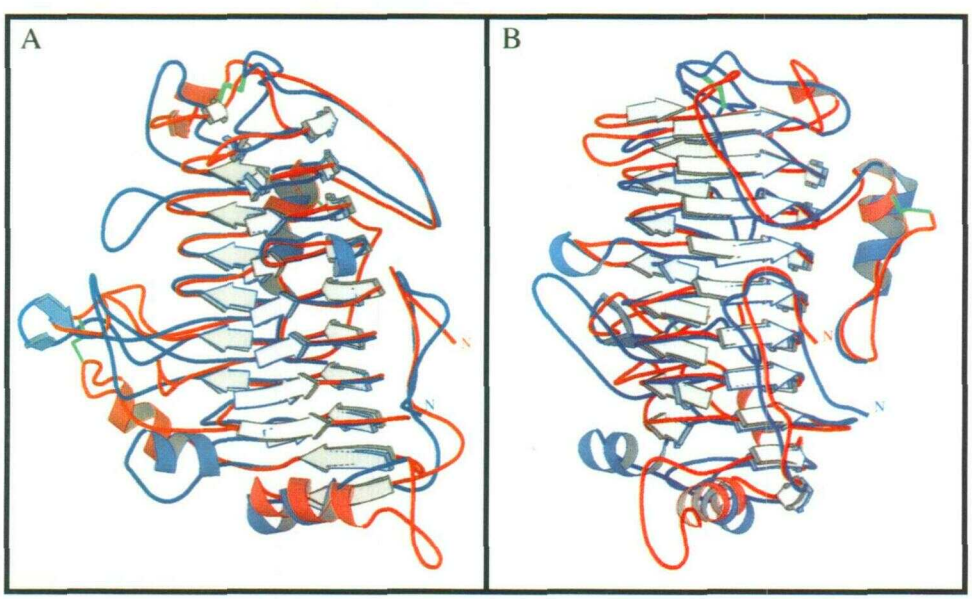

Figure 6. A, Superposition of the β ribbons diagrams of PelE (in blue) and PelC (in red). The parallel β strands of the parallel β helix core of both proteins are shown by gray arrows; the PelE β strands are outlined in blue. The single disulfide bond in PelE is located at the top of the figure, and the two disulfide bonds in PelC are illustrated by green lines. B, The same figure as in A, rotated 80° in the vertical axis.

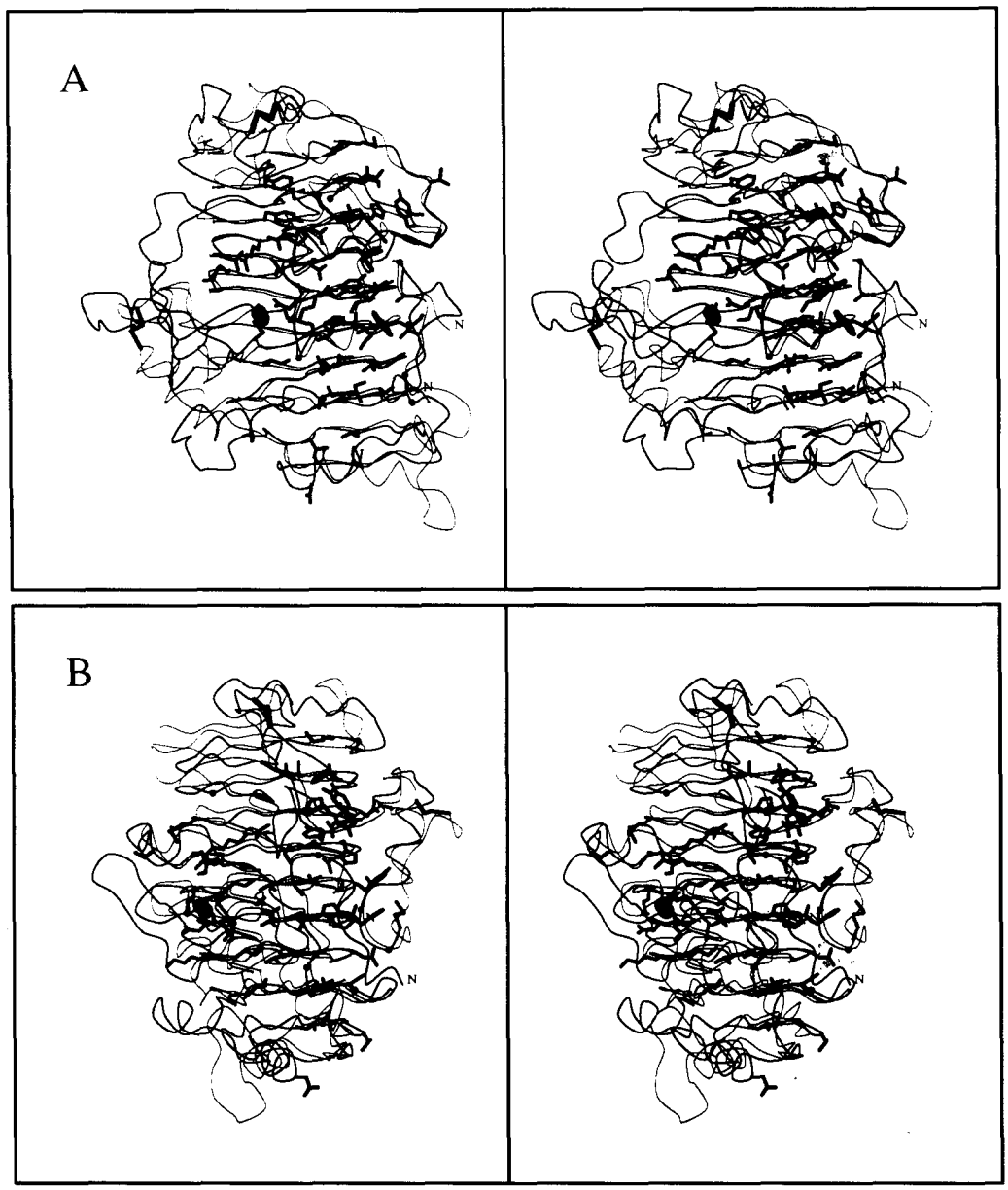

Figure 7. A, Stereo view of the superposition of the ribbon backbones of PelE (in dark gray) and PelC (in light gray). The single PelE disulfide bond is illustrated by the thickest black line, and the two disulfide bonds in PelC are illustrated by thick dark gray lines. The putative Ca^{2+} ions are marked by spheres, PelE in dark gray and PelC in light gray. The side chains of 59 invariant residues between PelE and PelC are shown in black for PelE only. The nine invariant Gly's are illustrated with small dark gray spheres at their αC position. B, The same figure as in A, rotated 80° in the vertical axis.

than in the base stacking mode observed within the core of the parallel β helix.

The remaining loops that protrude from the parallel β helix are conserved neither in length nor in conformation. Most of the nonconserved loops are located around the putative Ca^{2+} site and form a cleft on the surface of PelE or PelC. The difference in the number, size, and conformation of the loops that constitute the cleft is the most significant structural difference between the two proteins. Five loops form the cleft region in PelE and consist of residues 114 to 135, 153 to 175, 201 to 216, 248 to 260, and 297 to 301. The cleft in PelC is formed by residues in six loops, 36 to 51, 57 to 80, 126 to 130, 150 to 168, 267 to 274, and 292 to 306. The shape of the PelC cleft differs substantially from PelE. The PelC groove is longer as a consequence of the loop composed of residues 36 to 51 at the amino end of the parallel β helix. Also, the Cys^{72}-Cys^{155} bond, which connects two loops in PelC, residues 57 to 80 and 150 to 168, presumably constrains the loops into a more rigid conformation with close interaction. In contrast, the analogous loops in PelE are conforma-

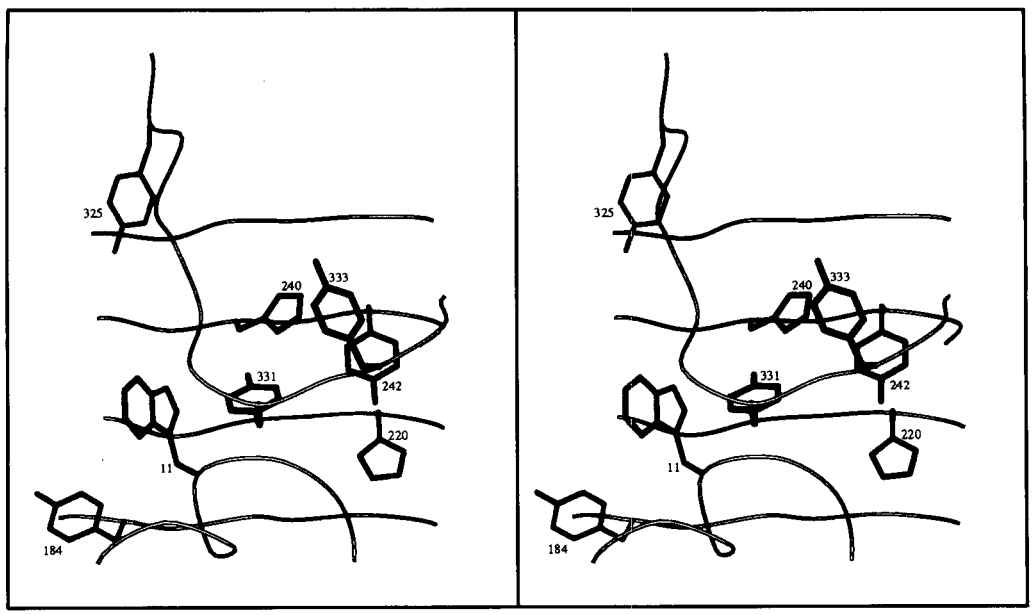

Figure 8. Stereo view of the apparent aromatic specificity pocket in PelE. The ringed residues that are identical in PelE and PelC are illustrated in black, and the conserved aromatic residue, Trp^{11}, is illustrated in gray. Also illustrated in gray is Tyr^{325}, whose centroid is positioned 3 Å away from the centroid of a comparable Tyr in PelC. The nonconserverd Tyr^{184} is illustrated in light gray.

tionally flexible. The disulfide bond in the PelC cleft is one of two that are invariant in the *pelBC* subfamily; the second invariant disulfide bond is Cys^{329}-Cys^{352}. Neither of the two invariant disulfide bonds in the PelC subfamily are analogous to the single invariant disulfide bond in the PelE subfamily. In PelE, the Cys^{291}-Cys^{320} bond connects two loops that cap the C-terminal end of the parallel β helix. In PelC, the second disulfide bond appears to stabilize a C-terminal loop, which is five residues longer than in PelE and extends back up to the carboxy end of the parallel β helix.

Two views of space-filling models of PelE and PelC are compared in Figure 9, with the charged amino acids coded by color. Surprisingly, the distribution of the surface charge and the overall shape of the two proteins are very different. In PelE, the charged groups are more uniformly distributed over the entire surface of the protein than is observed in PelC. One exception is a small region around the putative Ca^{2+} site. The majority of charged amino acids in PelC cluster in a narrow groove that includes the putative Ca^{2+} site and extends 51 Å. The size and shape of the groove in PelC approximate the dimensions of a dodecamer of galacturonate, leading Yoder et al. (1993a) to speculate that the region is the saccharide-binding site in PelC. The remainder of the PelC surface is predominantly neutral. Given the surface charge distribution in PelE, a potential saccharide-binding site is ambiguous. In addition to the charge distribution, another pronounced difference is the overall shape. PelE is quite globular, whereas PelC is more elongated.

DISCUSSION

The three-dimensional structure of *E. chrysanthemi* PelE is the most recent example of the newest structural class of proteins in which the predominant tertiary feature is an all-parallel β motif (Cohen, 1993; Sprang, 1993; Yoder et al., 1993a). The structural family also includes the related Pels *E. chrysanthemi* PelC (Yoder et al., 1993a) and *Bacillus subtilus* PelE (Pickersgill et al., 1994), as well as a partial domain of the *Pseudomonas aeruginosa* alkaline protease (Baumann et al., 1993). In all the structures, the parallel β strands are folded into a right-hand coil. In addition to the novel tertiary structure, the proteins have special interactions that contribute to their stability. In the Pels, the side chains of the amino acids on the interior of the core form extensive linear arrays of stacking interactions; in the protease structure, five interior Ca^{2+} ions stabilize short polypeptide bends between β strands through coordination to the protein. A major question left unanswered is the functional role, if any, of the novel topology. It is quite clear that the extensive network of intrastrand H bonding between parallel β sheets and the special types of side-chain or Ca^{2+} interactions are responsible for the observed stability of the proteins in solution. Such stability is, no doubt, required of extracellular proteins secreted into a hostile environment. What is more intriguing is the possibility that the novel parallel β motifs may aid in the secretion of each protein through the outer membrane of its respective organism or in its subsequent pathogenic function. Answers to these questions must await further investigation.

E. chrysanthemi PelE and PelC are isozymes sharing 22% sequence identity. Their genes are independently regulated (Collmer and Keen, 1986), yet both have similar enzymatic, secretory, and pathogenic properties. Although there are limited biochemical data regarding the mode of action of Pels, some details are known to differ. PelE and PelC cleave α-1,4-PGA in vitro in the presence of Ca^{2+} by a β-elimination reaction, but the end product of cleavage is a dimer for PelE

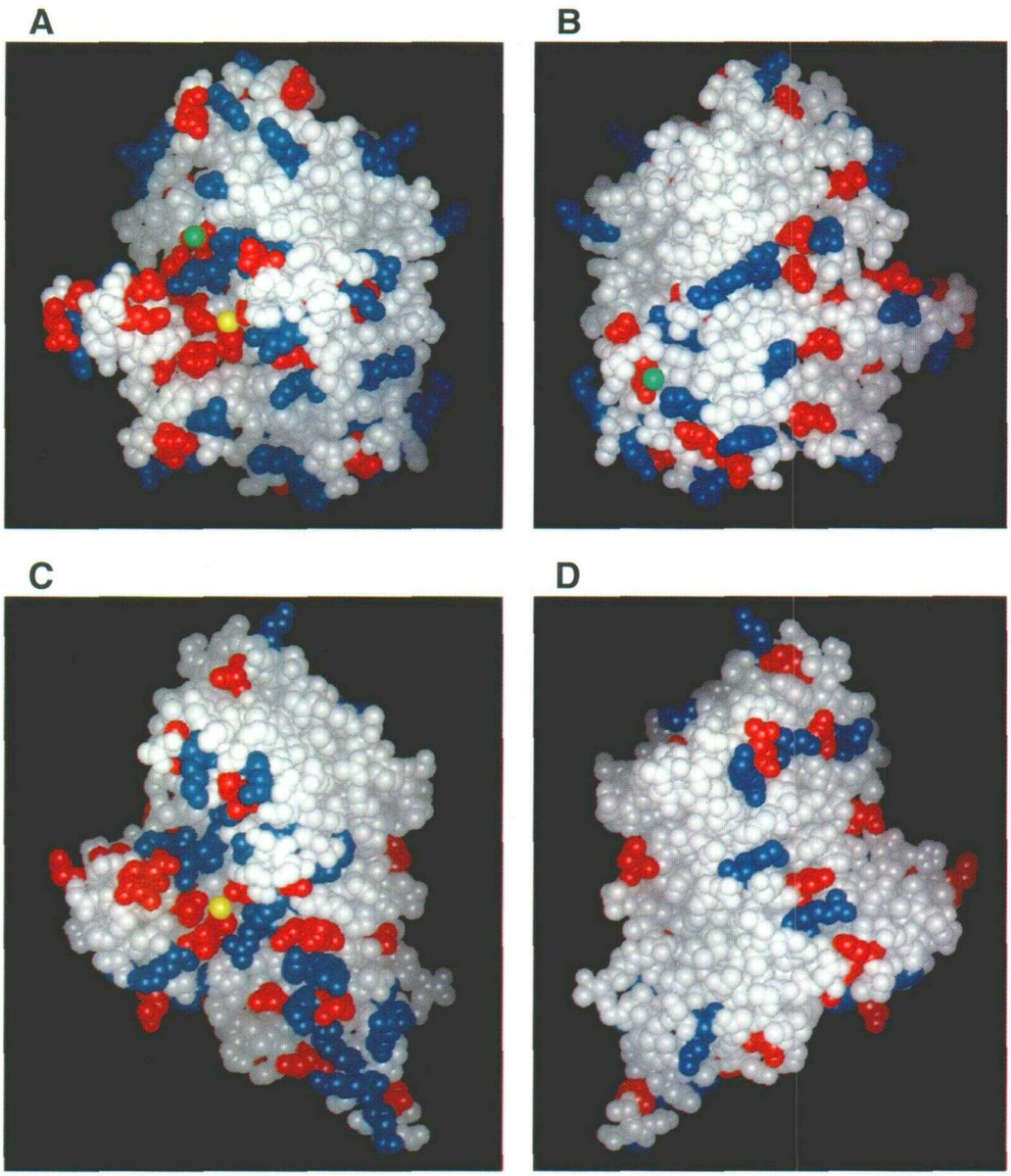

Figure 9. Space-filling models of PelE and PelC. Neutral residues are shown in light gray and charged residues are shown in color. The basic residues are illustrated in blue and acidic residues are illustrated in red. The putative Ca^{2+} ions are indicated in yellow and potential cationic ions are indicated in green. A, Standard view of PelE, as shown in Figure 5A, illustrating the groove with the putative Ca^{2+}-binding site. B, PelE rotated 180° in the vertical axis. C, Standard view of PelC, analogous to view of PelE in A. D, PelC rotated 180° in the vertical axis.

and a trimer for PelC (Preston et al., 1992). The same study suggests that the minimal substrate length for normal cleavage by either PelE or PelC is a hexamer of galacturonic acid. The pH of plant surfaces and tissues is 6 or lower (Yang et al., 1992); however, the pH optimum for both Pels is much higher: 9.0 for PelE and 9.5 for PelC (Kotoujansky, 1987; M. Garrett, F. Jurnak, N.T. Keen, unpublished observations). *E. chrysanthemi* PelE and PelC appear to have the same host specificity, but the maceration rate of recombinant PelE on potato tubers is 10-fold that of recombinant PelC (Tamaki et al., 1988). Such subtle differences in functional properties are likely to be correlated with small but detectable structural differences.

A comparison of the PelE and PelC structures was undertaken at atomic resolution to elucidate any distinguishing features relevant to their functional properties. Not surprisingly, the overall three-dimensional structures of PelE and PelC are very similar. The major differences between the two enzymes appear to be in the size and conformation of the loops that protrude from the core of the parallel β helix. The

surface loops dramatically alter the surface charge distribution and overall shape of the two enzymes. The cluster of charged amino acids in PelC in a narrow groove led Yoder et al. (1993a) to speculate that this region is a binding site for a negatively charged polygalacturonate substrate in PelC. A similar inference is not possible with only the PelE structure because the distribution of the surface charges is more random. The lack of an extensive charge cluster on the PelE surface suggests that the tentative identification of the pectinolytic active site in PelC must be re-evaluated and other possibilities considered.

The most plausible pectinolytic active site still remains the region around the putative Ca^{2+}-binding site. Ca^{2+} is essential for in vitro Pel activity, but its role is undetermined. A putative Ca^{2+} site toward one end of the charged groove in PelC was identified by the substitution results of three Ca^{2+}-like heavy atom derivatives. Uranyl, Pb^{2+}, and Lu^{3+} substituted at the same atomic location in PelC, each with a reasonable coordination to invariant or highly conserved amino acids in all Pels. In PelE, La^{3+} and one of three uranyl sites substitute at an analogous location to that in PelC with a similar coordination sphere. A second uranyl site, observed only in PelE, substitutes strongly 12.8 Å from the putative Ca^{2+} site. The coordination to a single, nonconserved amino acid as well as the lack of substitution by the La^{3+} derivative suggests that the second uranyl site is a cationic binding site, but not specific for Ca^{2+}. The significance of a second cationic site in PelE, but not in PelC, lies in a possible difference in the mode of saccharide binding between the two enzymes. PGA is a saccharide that contains repetitive, negatively charged uronate groups. Potentially, a charged substrate could bind to the protein via one or more ionic interactions. In the Pels, ionic interactions with the substrate are likely to occur via Lys's or Arg's. In the putative saccharide-binding regions, PelE and PelC share four positively charged amino acids: Lys^{101}, Lys^{197}, Arg^{230}, and Arg^{235} in PelE nomenclature. Another potential source of an ionic interaction is the essential Ca^{2+} found in all Pels. In addition, the putative cationic site in PelE, spaced 12.8 Å from a Ca^{2+} site, is suggestive of another strong ionic interaction between PelE and the substrate, thereby reducing the need for many of the positively charged amino acids found in the groove of PelC. Alternatively, PelE may have a greater affinity for a different saccharide substrate in vivo than PelC. Given the relatively low number of positively charged amino acids in this region of PelE compared to that in PelC, it is quite plausible that the in vivo PelE substrate may not consist entirely of galacturonic acid units. A difference in saccharide specificity in vivo may be responsible for some of the differences in maceration rates between the two enzymes. Unfortunately, no studies have been carried out that define the optimal saccharide substrate for any of the Pels.

The results of the PelE and PelC structural comparison is surprising, not only for the observed differences in the loop structures around the putative Ca^{2+} site but also for the structural conservation of some loops with unusual conformations and for the discovery of an apparent aromatic specificity pocket. The most striking similarity between PelE and PelC is the conformation of a segment of the N-terminal loop and the large omega loop, neither of which shares extensive sequence homology among the Pels. Both loops fold over and cover one exterior surface of the parallel β helix. The two loops meet in a mid-region that has highly conserved sequences, not only among pectate and pectin lyases but among other plant homologs. Many of the invariant amino acids in this region (Kuster-Van Someren, 1991) cluster around the Asn ladder in PelE and PelC, with several invariant residues lying on the exterior of the parallel β helix. With the exception of the invariant Trp, Asp, and His in a highly conserved vWiDH sequence, most of the invariant amino acids are not exposed to the bulk solvent region. Curiously, a highly conserved aromatic residue is located at the first turn of the N-terminal branch. The side chain of Trp^{11} in PelE and Tyr^{7} in PelC is oriented in a cavity surrounded on one side by five invariant ringed amino acids located on the exterior surface of the parallel β helix or on the interior surface of the large omega loop (Fig. 8). The presence of an apparent aromatic specificity pocket suggests an essential role, perhaps an alignment function, for the conserved aromatic residue in the N-terminal branch. The entire region is also a potential candidate for the pectinolytic active site. The catalytic amino acids would most likely include the invariant Asp^{147} and His^{148}, whose side chains form a hydrogen bond and have a structural orientation reminiscent of the catalytic amino acids in Ser proteases. The concentration of invariant and highly conserved ringed residues in the region also is suggestive of a common mode of saccharide-protein interactions involving aromatic groups. However, if this region were the pectinolytic active site, it would be difficult to postulate an essential role for the Ca^{2+}, whose putative binding site lies approximately 65 to 95 Å from Asp^{147} around the circumference of the parallel β helix. Given the large separation between the two regions, it is rather unlikely that both regions are part of the pectinolytic active site. Rather it is more probable that Pels have another, as-yet unidentified enzymatic function that may be related to secretion or pathogenesis.

Several investigators have noted that the Pels share sequence homologies with pollen and style plant proteins (Wing et al., 1989; Kuster-Van Someren, 1991; Rafnar et al., 1991). By comparison with the PelE and PelC structures, the homologies include all amino acids except Lys^{178} around the putative Ca^{2+}-binding site, two invariant Arg's in the putative saccharide binding region, the three-residue Asn ladder observed in PelE, and the highly conserved vWiDH region. The homologies suggest that not only are the plant proteins likely to have the parallel β helix topology, but they are also likely to share similar functions, such as some type of saccharidase activity. Such functional similarities have important ramifications for pollination as well as for the allergenic response in which the pollen homologs play a major role.

In the last 30 years, most macromolecular structural determinations have been preceded by a plethora of biochemical studies to facilitate the structural interpretation of functional domains. Because Pels are generally secreted as isozymes, which pose a purification problem, the enzymes were not suitable for crystallographic analyses until the genes for single isozymes had been cloned and overexpressed in *E. coli*. By coupling recombinant DNA technology with advancements in crystallographic methodology, the Pel system now dem-

onstrates that it can be more expedient to obtain detailed information at atomic resolution than to characterize an enzyme by other methods. Unfortunately, the lack of supporting biochemical data for Pels has complicated the task of correlating the structures with functional properties. However, comparison of the PelE and PelC structures has revealed unforeseen structural differences in loops and surface properties that may ultimately be correlated with differences in enzymatic and maceration properties of the virulence factors. Even more unexpected is the structural conservation of the N- and C-terminal regions and the discovery of an apparent specificity pocket for an aromatic residue. Just as the initial structure of a Pel was a surprise, the first comparative analysis of two Pels suggests that more surprises are likely to follow when the enzymatic, secretory, and pathogenic functions of these virulence factors are probed.

ACKNOWLEDGMENTS

F.J. gratefully acknowledges the support of the Academic Computing Graphics and Visual Imaging Lab, University of California, Riverside, and the San Diego Supercomputer Center.

Received March 29, 1994; accepted July 1, 1994.
Copyright Clearance Center: 0032-0889/94/106/0849/14.

LITERATURE CITED

Baker EN, Hubbard RE (1984) Hydrogen bonding in globular proteins. Prog Biophys Mol Biol **44:** 97–179

Baumann U, Wu S, Flaherty KM, McKay DB (1993) Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel β roll motif. EMBO J **12:** 3357–3364

Bhat TN, Cohen GH (1984) OMITMAP: an electron density map suitable for the examination of errors in a macromolecular model. J Appl Crystallogr **17:** 244–248

Blundell TL, Johnson LN (1976) Protein Crystallography. Academic Press, New York, pp 373–375

Brünger AT (1993a) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. Acta Crystallogr **D49:** 24–36

Brünger AT (1993b) X-PLOR Manual, Version 3.1. Yale University, New Haven, CT

Brünger AT, Kuriyan J, Karplus M (1987) Crystallographic R factor refinement by molecular dynamics. Science **235:** 458–460

Burley SK, Petsko GA (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. Science **229:** 23–28

Cohen FE (1993) The parallel β helix of pectate lyase c: something to sneeze at. Science **260:** 1444–1445

Collmer A, Keen NT (1986) The role of pectic enzymes in plant pathogenesis. Annu Rev Phytopathol **24:** 383–409

Condemine G, Dorel C, Hugouvieux-Cotte-Pattat N, Robert-Baudouy J (1992) Some of the *out* genes involved in the secretion of pectate lyases in *Erwinia chrysanthemi* are regulated by kdgR. Mol Microbiol **6:** 3199–3211

Crawford M, Kolattukudy PE (1987) Pectate lyase from *Fusarium solani* f. sp. *pisi*: purification, characterization, in vitro translation of the mRNA, and involvement in pathogenicity. Arch Biochem Biophys **258:** 196–205

Fitzgerald P, Madsen N (1986) Improvement of limit diffraction and useful x-ray lifetime of crystals of glycogen debranching enzyme. J Crystal Growth **76:** 600–606

He SY, Schoedel C, Chatterjee AK, Collmer A (1991a) Cloned *Erwinia chrysanthemi out* genes enable *Escherichia coli* to selectively secrete a diverse family of heterologous proteins to its milieu. Proc Natl Acad Sci USA **88:** 1079–1083

He SY, Schoedel C, Chatterjee AK, Collmer A (1991b) Extracellular secretion of pectate lyase by the *Erwinia chrysanthemi* out pathway is dependent upon sec-mediated export across the inner membrane. J Bacteriol **173:** 4310–4317

Hinton JCD, Sidebotham JM, Gill DR, Salmond GPC (1989) Extracellular and periplasmic isoenzymes of pectate lyase from *Erwinia carotovora* subspecies *carotovora* belong to different gene families. Mol Microbiol **3:** 1785–1795

Hodel A, Kim S-H, Brünger AT (1992) Model bias in macromolecular crystal structures. Acta Crystallogr **A48:** 851–858

Howard AJ, Nielson C, Xuong N-H (1985) Software for a diffractometer with multiwire area detector. Methods Enzymol **114:** 452–472

Jones TA (1985) Interactive computer graphics: FRODO. Methods Enzymol **115:** 157–171

Jones TA, Cowan S, Zou J-Y, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr **A47:** 110–119

Jones TA, Kjeldgaard M (1993) O: The Manual, Version 5.9.1. Uppsala University, Uppsala, Sweden

Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen bonded and geometrical features. Biopolymers **22:** 2577–2637

Keen NT, Tamaki S (1986) Structure of two pectate lyase genes from *Erwinia chrysanthemi* EC16 and their high-level expression in *Escherichia coli*. J Bacteriol **168:** 595–606

Kim C, Mosser V, Keen NT, Jurnak F (1989) Preliminary crystallographic analysis of a plant pathogenic factor: pectate lyase. J Mol Biol **208:** 365–367

Kotoujansky A (1987) Molecular genetics of pathogenesis by soft-rot erwinias. Annu Rev Phytopathol **25:** 405–430

Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr **24:** 946–950

Kuster-Van Someren M (1991) Characterization of an *Aspergillus niger* pectin lyase gene family. PhD dissertation. University of Utrecht, The Netherlands

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr **26:** 283–291

Leszczynski JF, Rose GD (1986) Loops in globular proteins: a novel category of secondary structure. Science **234:** 849–855

Lindeberg M, Collmer A (1992) Analysis of eight *out* genes in a cluster required for pectic enzyme secretion by *Erwinia chrysanthemi*: sequence comparison with secretion genes form other gram-negative bacteria. J Bacteriol **174:** 7385–7397

Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature **356:** 83–85

Matthews B (1968) Solvent content of protein crystals. J Mol Biol **33:** 491–497

Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) Stereochemical quality of protein structure coordinates. Proteins Struct Funct Genet **12:** 345–364

Pickersgill R, Jenkins J, Harris G, Nassar W, Robert-Baudouy J (1994) The structure of *Bacillus subtilus* pectate lyase in complex with calcium. Nature Struct Biol (in press)

Preston JF, Rice JD, Ingram LO, Keen NT (1992) Differential depolymerization mechanisms of pectate lyases secreted by *Erwinia chrysanthemi* EC16. J Bacteriol **174:** 2039–2042

Rafnar T, Griffith IJ, Kuo MC, Bond JF, Rogers BL, Klapper DG (1991) Cloning of *Amb a I* (antigen E), the major family of short ragweed pollen. J Biol Chem **266:** 1229–1236

Read R (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr **A42:** 140–149

Ring CS, Kneller DG, Langridge R, Cohen FE (1992) Taxonomy and conformational analysis of loops in proteins. J Mol Biol **224:** 685–699

Singh J, Thornton JM (1985) The interaction between phenylalanine rings in proteins. FEBS Lett **191:** 1–6

Sprang SR (1993) On a β-roll. Trends Biochem Sci **18:** 313–314

Tamaki SJ, Gold S, Robeson M, Manulis S, Keen NT (1988) Structure and organization of the *pel* genes from *Erwinia chrysanthemi* EC16. J Bacteriol **170:** 3468–3478

Ten Eyck L, Weaver LA, Matthews BW (1976) A method of obtaining a stereochemically acceptable protein model which fits a set of atomic coordinates. Acta Crystallogr **A32:** 349–355

Terwilliger TC, Eisenberg DE (1983) Unbiased three dimensional refinement of heavy atom parameters by correlation of origin removed Patterson functions. Acta Crystallogr **A39:** 813–817

Walkinshaw MD, Arnott S (1981) Conformations and interactions of pectins. I. X-ray diffraction analyses of sodium pectate in neutral and acidified forms. J Mol Biol **153:** 1055–1073

Wang B-C (1985) Resolution of phase ambiguity in macromolecular crystallography. Methods Enzymol **115:** 90–112

Wilmot CM, Thornton JM (1990) β-Turns and their distortions: a proposed new nomenclature. Protein Eng **3:** 479–493

Wilson AJC (1942) Determination of absolute from relative x-ray intensity data. Nature **150:** 90–112

Wing RA, Yamaguchi J, Larabell SK, Ursin VM, McCormick S (1989) Molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen *Erwinia*. Plant Mol Biol **14:** 17–23

Yang Z, Cramer CL, Lacy GH (1992) *Erwinia carotovora* subsp. *carotovora* pectic enzymes: *in planta* gene activation and roles in soft-rot pathogenesis. Mol Plant Microbe Interact **5:** 104–112

Yoder MD, Keen NT, Jurnak F (1993a) New domain motif: the structure of pectate lyase c, a secreted plant virulence factor. Science **260:** 1503–1507

Yoder MD, Lietzke SE, Jurnak F (1993b) Unusual structural features in the parallel β helix in pectate lyases. Structure **1:** 241–251