

Leonard M. Thomas,<sup>†</sup> Chuong N. Doan,<sup>‡</sup> Randall L. Oliver<sup>§</sup> and Marilyn D. Yoder\*

School of Biological Sciences, University of Missouri-Kansas City, 5007 Rockhill Road, Kansas City, MO 64110-2499, USA

<sup>†</sup> Current address: Howard Hughes Medical Institute, Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.

<sup>‡</sup> Current address: The Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110, USA.

<sup>§</sup> Current address: Zymogenetics, 1201 Eastlake Avenue East, Seattle, WA 98102, USA.

Correspondence e-mail: yoderm@umkc.edu

## Structure of pectate lyase A: comparison to other isoforms

Pectate lyase A is a virulence factor secreted by the plant-pathogenic bacteria *Erwinia chrysanthemi*. The enzyme cleaves the glycosidic bond of pectate polymers by a calcium-dependent  $\beta$ -elimination mechanism. The crystal structure of pectate lyase A from *E. chrysanthemi* EC16 has been determined in two crystal forms, monoclinic C2 to 1.8 Å and rhombohedral R3 to 2.1 Å. The protein structure is compared with two other pectate lyase isoforms from *E. chrysanthemi* EC16, pectate lyase C and pectate lyase E. Pectate lyase A is unique as it is the only acidic pectate lyase and has end products that are significantly more varied in length in comparison to those of the other four major pectate lyase isozymes. Differences and similarities in polypeptide trace, size and volume of the active-site groove and surface electrostatics are discussed.

Received 5 December 2001

Accepted 2 April 2002

**PDB References:** C2 PelA, 1jta, r1jtasf; R3 PelA, 1jrg, r1jrgsf.

### 1. Introduction

Pectate lyases (Pels) are virulence factors of necrotrophic pathogens that cause bacterial soft rots in plants (Collmer & Keen, 1986; Alfano & Collmer, 1996). *Erwinia* species are one of the most extensively characterized of the soft-rot pathogens and secrete several pectate-degrading enzymes, including pectate lyases, polygalacturonases and pectin methylesterases. Pels cleave the  $\alpha$ -1,4-galacturonosyl linkages in pectate polymers, which are components of both the plant cell wall and the middle lamella. The enzymes utilize a calcium-dependent  $\beta$ -elimination mechanism. *E. chrysanthemi* produces five major isozymes of pectate lyase, named alphabetically in order of increasing pI: PelA, PelB, PelC, PelD and PelE.

The major *pel* genes in *Erwinia* are organized as two multigene families on separate chromosomal clusters. The *pelA*, *pelE* and partially deleted *pelD* genes of *E. chrysanthemi* EC16 form one cluster, presumed to be created by a duplication event. The second gene cluster consists of the *pelB* and *pelC* genes. PelB and PelC share extensive genetic homology, with 84% amino-acid identity. The *pelA/pelE* genes have diverged more than the *pelB/pelC* genes, exhibiting 61% amino-acid identity (Tamaki *et al.*, 1988; Barras *et al.*, 1987).

The first structure of a pectate lyase, PelC, was determined from *E. chrysanthemi* EC16 (Yoder, Keen *et al.*, 1993). The polypeptide chains folds in a large right-handed helical coil, forming three parallel  $\beta$ -sheets. This novel fold is referred to as a parallel  $\beta$ -helix. Other Pel structures reported to date include *E. chrysanthemi* PelE (Lietzke *et al.*, 1994), *E. chrysanthemi* PelC bound to a pentagalacturonate substrate (Scavetta *et al.*, 1999) and the pectate lyase from *Bacillus subtilis* (Pickersgill *et al.*, 1994). The structure of PelC bound to a pectate fragment has been particularly enlightening in

**Table 1**  
Crystal parameters for PelA C2 and R3 crystal forms.

Space group	C2	R3
Unit-cell parameters (Å, °)	$a = 87.549$ $b = 53.802$ $c = 71.039$ $\beta = 109.448$	$a = 98.055$ $b = 98.055$ $c = 217.162$ $\gamma = 120$
Unit-cell volume (Å <sup>3</sup> )	$3.1 \times 10^5$	$1.8 \times 10^6$
$V_M$ (Å <sup>3</sup> Da <sup>-1</sup> )	2.03	2.54
Molecules per asymmetric unit	1	2
Estimated solvent content (%)	29.9	44.0
Diffraction limit (Å)	1.6	2.1

postulating a catalytic role for an Arg in the enzymatic mechanism (Herron *et al.*, 2000).

The oligosaccharide end products of PelA catalysis on a polygalacturonate substrate vary in length from dimers to octamers and nonamers. In contrast, the end products of PelC and PelB are predominantly trimeric saccharides, while the PelE end products are primarily dimeric (Preston *et al.*, 1992). It is not clear from comparing the known structures of PelC and PelE why this end-product difference occurs.

From biochemical studies, it is clear that PelA has several unique features relative to the other isozymes that warrant a structural characterization. PelA is the only acidic isozyme, with a pI of 4.6. While the specific activity of PelA is lower but similar to that of PelE (Barras *et al.*, 1987), PelA and the other Pel isozymes exhibit significant differences in tissue-maceration ability. PelB, PelC and PelE cleave polygalacturonate to small oligosaccharide end products. In contrast, PelA end products are considerably longer and more varied (Preston *et al.*, 1992). The structure determination of PelA was initiated to elucidate possible structural rationalizations in order to address these biochemical observations.

## 2. Materials and methods

### 2.1. Protein purification and crystallization

The protein was purified and crystallized as described previously (Doan *et al.*, 2000). *Escherichia coli* cells, which contain pPEL812 containing a high-expression plasmid construct with the *pelA* gene from *E. chrysanthemi* EC16 (Tamaki *et al.*, 1988), were grown in LB media and the protein was purified from the periplasmic fraction. The protein was sufficiently pure for crystallization after DEAE ion-exchange chromatography followed by hydrophobic interaction chromatography on phenyl Sepharose. During the course of the study, two additional crystal forms of PelA were found. The C2 form crystallizes under the same conditions as a previously described P2<sub>1</sub> form (Doan *et al.*, 2000). The R3 crystal form was grown at room temperature using sitting-drop vapor-diffusion techniques. A 0.006 ml solution containing 6.0 A<sub>280</sub> of PelA in 10 mM HEPES pH 7.0 with 0.5 M Li<sub>2</sub>SO<sub>4</sub> and 1% polyethylene glycol (PEG) 8K was equilibrated over a 0.8 ml reservoir of 1 M Li<sub>2</sub>SO<sub>4</sub> and 2% PEG 8K. Crystal parameters for the two new crystal forms are given in Table 1.

**Table 2**  
Statistics for X-ray data collection for the C2 and R3 crystal forms.

	C2	R3
Wavelength (Å)	0.9795	1.543
Temperature (K)	110	153
Resolution (Å)	1.8	2.1
Reflections collected	88736	111710
Unique reflections	36568	44138
Completeness (%)	88.3 (64.9)	98.3 (99.8)
$\langle I/\sigma(I) \rangle$	21.2 (6.28)	12.6 (3.5)
$I > 3\sigma(I)$ † (%)	90.5 (78.3)	81.1 (50.3)
$R_{\text{sym}}^\ddagger$ (%)	4.1 (16.0)	7.7 (31.0)
Mosaicity (°)	0.35	0.64

† Percentage of data with  $I > 3\sigma(I)$ . ‡  $R_{\text{sym}} = \sum(I_{hkl} - \langle I_{hkl} \rangle) / \sum I_{hkl}$  (single measurements excluded).

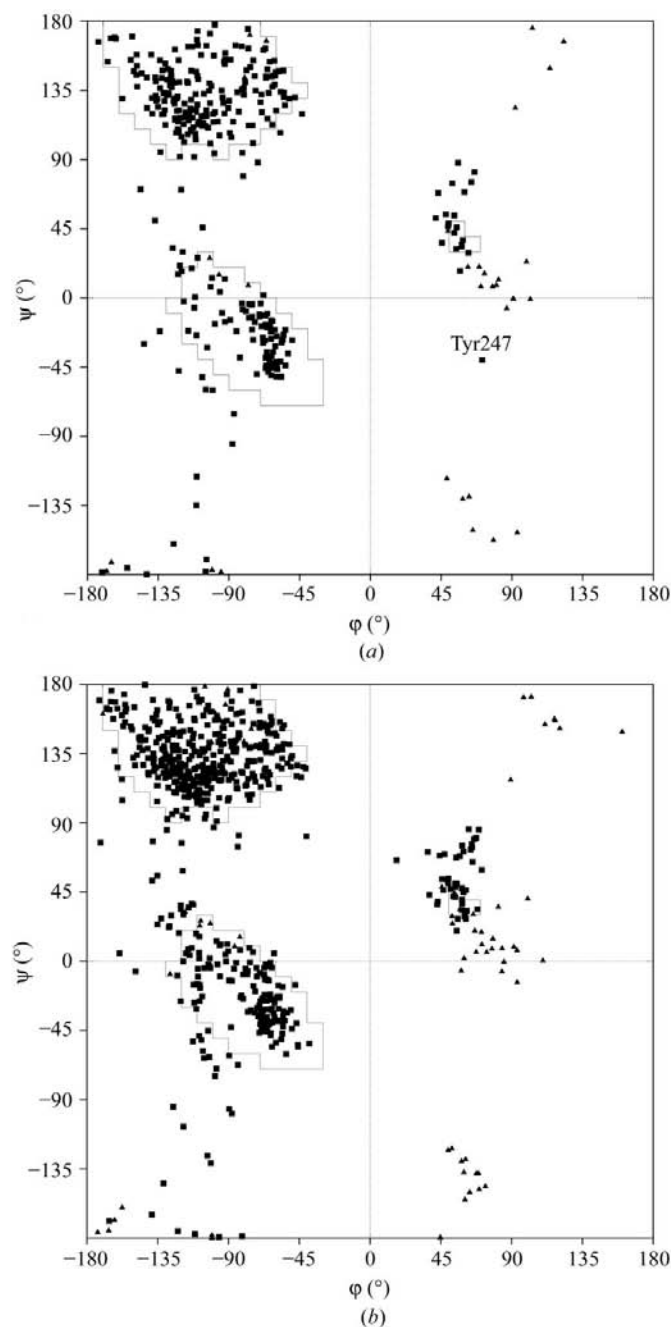
### 2.2. Data collection and refinement

Data-collection statistics are given in Table 2. The R3 crystal data were collected on a MAR300 phosphoimaging-plate detector mounted on a rotating-anode generator operating at 50 kV and 100 mA with a Supper graphite monochromator. The loop-mounted crystal was flash-cooled after a serial transfer to a cryosolution of 2 M Li<sub>2</sub>SO<sub>4</sub>, 40% PEG 8K containing increasing percentages of ethylene glycol to a final concentration of 25%. X-ray data were collected at 153 K with an exposure time of 6 min, an oscillation range of 1° and a crystal-to-detector distance of 150 mm.

The C2 crystal data were collected with synchrotron radiation at NSLS beamline X8C. Single-wavelength X-ray data were collected at 0.9795 Å on an ADSC Quantum4 detector with a crystal-to-detector distance of 150 mm. The oscillation range was 1° per frame, with an exposure time of 60 s. Data were collected at 110 K from a flash-cooled crystal after a serial transfer to a cryosolution of 10% PEG 8K, 4% PEG 1K containing increasing percentages of ethylene glycol to a final concentration of 20%.

X-ray diffraction images were indexed and scaled with *HKL* (Otwinowski & Minor, 1997). The initial structure was determined by molecular replacement as implemented in *AMoRe* (Navaza, 1994). The molecular-replacement search model was based on PelE residues 30–61, 70–100, 105–251, 258–280, 293–314 and 323–332. Positional parameters were optimized by rigid-body refinement against the C2 data. Missing portions of the PelA model were manually built into electron-density maps using the computer-graphics program *O* (Jones *et al.*, 1991). The models were checked using simulated-annealing OMIT maps and careful monitoring of  $R_{\text{free}}$  statistics (Brünger, 1993) as calculated by *CNS* (Brünger *et al.*, 1998). The refined model from the C2 data consists of all residues in the protein (1–361), 244 water molecules and one sulfate ion. The side chains of residues Glu2, Lys135, Asp137 and Lys334 were modeled as Ala owing to a lack of detectable electron density beyond the C<sup>β</sup> atom. One residue, Tyr247, has an unfavorable combination of  $\psi$ ,  $\phi$  backbone torsion angles as determined by *PROCHECK* (Laskowski *et al.*, 1992) (Fig. 1a).

The R3 X-ray data were twinned; the twin law was determined by the Merohedral Crystal Twinning Server for partial twinning (Yeates, 1997). The twinning operator was identified to be a twofold along *a* and *b*, with *h*, *k*, *l* related to *h*,  $-h - k$ ,  $-l$  and a twinning fraction of 0.231. Detwinned data were generated by *CNS* (Brünger *et al.*, 1998) and used for molecular replacement against the C2 model of PelA. The model was refined using *CNS* (Brünger *et al.*, 1998) against detwinned data using strict NCS restraints. After all residues were built into the electron-density maps, except two loops at



**Figure 1**  
Ramachandran plots for the final PelA models. Triangles represent glycine residues; squares represent all other amino acids. The core regions are outlined. (a) C2 crystal form, (b) R3 crystal form.

**Table 3**  
Final refinement statistics as a function of resolution.

C2			R3		
Resolution range (Å)	$R_{\text{cryst}}$ (%)	$R_{\text{free}}$ (%)	Resolution range (Å)	$R_{\text{cryst}}$ (%)	$R_{\text{free}}$ (%)
3.60–50.0	18.51	21.84	4.20–50.0	13.01	16.52
2.86–3.60	18.54	21.89	3.33–4.20	12.71	18.79
2.50–2.86	18.80	23.06	2.91–3.33	15.50	21.33
2.27–2.50	18.00	23.86	2.65–2.91	17.35	22.73
2.11–2.27	17.88	21.91	2.46–2.65	18.85	25.79
1.98–2.11	17.53	23.79	2.31–2.46	19.40	24.37
1.88–1.98	16.75	22.23	2.20–2.31	20.29	23.42
1.80–1.88	17.65	23.19	2.10–2.20	22.07	26.10
Overall (1.80–50.0)	18.15	22.52	Overall (2.20–50.0)	15.96	21.09

**Table 4**  
Refinement and model statistics for C2 and R3 crystal forms of PelA.

	C2	R3
Resolution range (Å)	50.0–1.8	50.0–2.1
$R_{\text{cryst}}$ † ( $R_{\text{free}}$ ) (%)	18.15 (22.53)	15.96 (21.09)
No. of reflections (in test set)	26888 (2667)	37994 (3990)
No. of atoms		
Protein	2732	5308
Water molecules	244	325
Sulfate molecules	1	4
R.m.s.d. <i>B</i> factors (Å <sup>2</sup> )		
Main-chain bonds	1.075	1.057
Main-chain angles	1.539	1.753
Side-chain bonds	1.892	1.526
Side-chain angles	2.615	2.175
Average <i>B</i> values (Å <sup>2</sup> )		
Protein main chain	11.904	21.117
Protein side chain	13.900	21.841
Water molecules	23.978	22.900
Sulfate molecules	31.244	51.900
R.m.s.d. from ideality		
Bond length (Å)	0.005	0.007
Bond angle (°)	1.401	1.418
Dihedrals (°)	26.482	27.001
Ramachandran plot‡		
Favored (%)	85.1	83.4
Allowed (%)	15.2	15.9
Generously allowed (%)	0.3	0.7
Disallowed (%)	0.3 (Tyr247)	0.0

†  $R_{\text{cryst}} = \sum [F(\text{obs})_{hkl} - F(\text{calc})_{hkl}] / \sum F(\text{obs})_{hkl}$ . ‡ Glycines and prolines excluded from calculations.

134–138 and 214–222, NCS restraints were relaxed using  $R_{\text{free}}$  to guide NCS weighting. Ultimately, no NCS restraints were used. The current refined model consists of two molecules of PelA, 325 water molecules and four sulfate ions. Molecule *A* consists of residues 1–214 and 222–361, while molecule *B* consists of residues 1–213 and 225–361. Side chains for residues LysA135, LysA223, LysA225, LysA299 and LysB225 were modeled as Ala owing to a lack of detectable electron density beyond the  $C^{\beta}$  atom. There are no residues with unfavorable combinations of  $\psi$ ,  $\phi$  backbone torsion angles as determined by *PROCHECK* (Lawkowski *et al.*, 1992) (Fig. 1*b*). Final refinement statistics as a function of resolution for both crystal forms are presented in Table 3.

Solvent molecules were added after all observed amino-acid residues were modeled. Water peaks were assigned using the

standard protocol as implemented with *CNS*. Briefly, peaks greater than  $3\sigma$  were identified from a  $F_{\text{obs}} - F_{\text{calc}}$  electron-density map. Potential solvent molecules had to meet a minimum and maximum distance to any atom of 2.6 and 4.0 Å, respectively, and a minimum and maximum hydrogen-bonding distance between appropriate atoms of 2.0 and 3.2 Å, respectively. All potential solvent molecules were visually checked in *O*. Sulfate ions were identified by visual inspection of  $2F_{\text{obs}} - F_{\text{calc}}$  electron-density maps calculated with water molecules included. Sulfate ions were initially identified by non-spherical electron density and were the top solvent peaks in both the *C2* and *R3* crystal forms. These solvent peaks had very similar electron-density levels to those observed for the S atoms in Cys and Met residues when the maps were drawn at high contour levels. Model statistics for both crystal forms are shown in Table 4.

### 2.3. Volume of saccharide-binding groove

The solvent-accessible surface and volume of the saccharide-binding pocket were calculated with *SURFNET* (Laskowski, 1995) using a minimum and maximum sphere radius of 1.4 and 4.0 Å, respectively, and a grid spacing of 0.8 Å.

## 3. Results

### 3.1. Structure of PelA

The crystal structure of PelA was determined by molecular replacement using the PelE structure as the phasing model. PelA consists of a single domain which folds as a classic parallel  $\beta$ -helix (Yoder, Lietzke *et al.*, 1993) (Fig. 2). The core of the protein contains three parallel  $\beta$ -sheets, which form eight coils along an axis of constant diameter. The parallel  $\beta$ -sheets are referred to as PB1, PB2 and PB3 and the loop

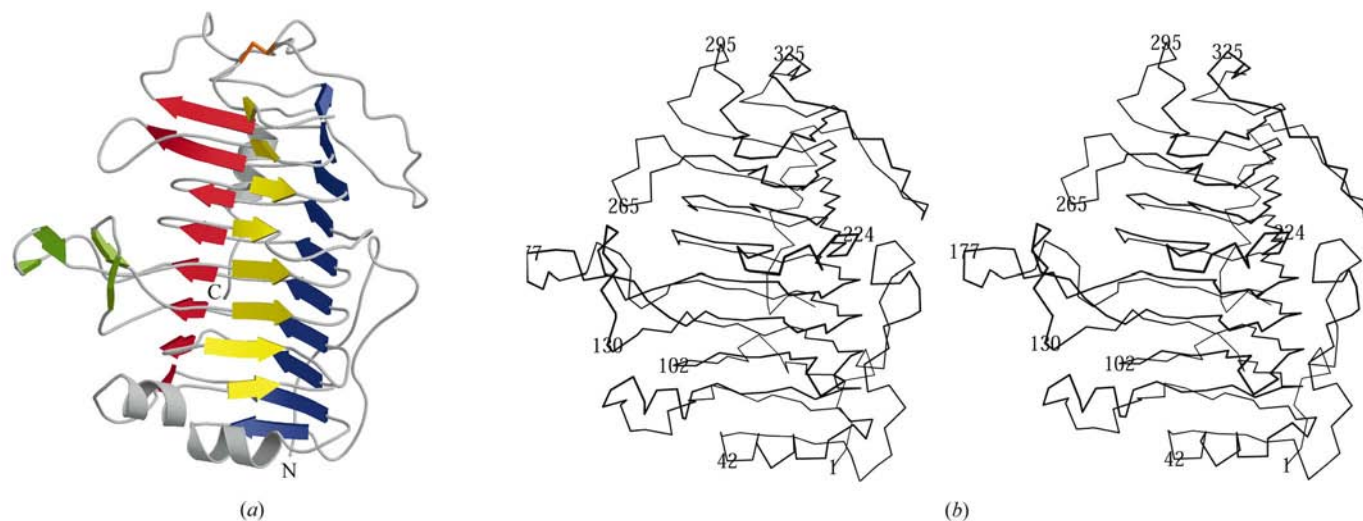
regions are T1 (connecting PB1 with PB2), T2 (between PB2 and PB3) and T3 (between PB3 and PB1). An individual loop is referred to as its general type followed by its location on the parallel  $\beta$ -helix (*i.e.* T1.5 refers to the T1 loop on the fifth coil of the parallel  $\beta$ -helix, counting from the N-terminus).

The two Cys residues, 298 and 326, form a disulfide bond across the C-terminal end of the parallel  $\beta$ -helix. There are three  $\alpha$ -helices consisting of more than one turn. One  $\alpha$ -helix caps the N-terminal end of the parallel  $\beta$ -helix, one is part of the T3.1 loop and one runs coincident to the axis of the parallel  $\beta$ -helix at the C-terminal end of the protein. There are two  $\beta$ -ribbons not participating in the core parallel  $\beta$ -helix, on the T3.3 and T3.4 loops. The saccharide-binding site is assumed to be located in the same place as observed in the PelC–pentagalacturonate crystal structure (Scavetta *et al.*, 1999): in a long groove with one side defined by several T3 loops and the other side by parts of PB1.

The PelA structures contain one *cis*-proline residue, Pro243. This is a conserved residue, both in amino-acid type and *cis* conformation.

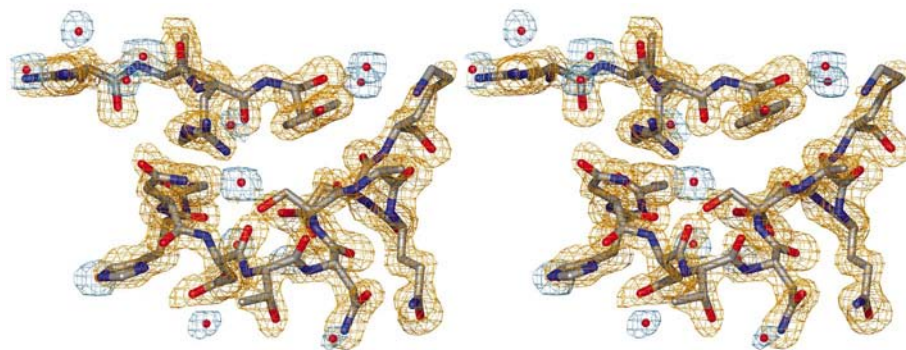
### 3.2. Differences between the *R3* and *C2* PelA structures

The structures of two crystallographic forms of PelA were determined. The root-mean-square differences (r.m.s.d.s) between all shared residues in the *C2* PelA form and molecules *A* and *B* of the *R3* form are 0.515 and 0.368 Å, respectively. The differences in structure between the *R3* and *C2* forms of PelA are located in the loop regions. The T3.4 loop between residues 173 and 180 exhibits the largest difference in backbone  $C^\alpha$  atoms. The other difference is in the T1.5 loop. Both molecules in the *R3* form have an untraceable region in the T1.5 loop between residues 214 and 224. The electron density in the T1.5 loop region in the *C2* form is observed (Fig. 3). The r.m.s.d. between  $C^\alpha$  atoms of all shared



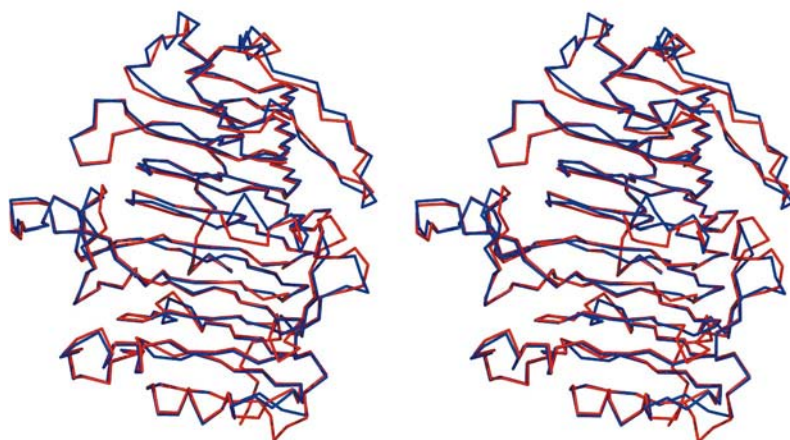
**Figure 2**

Structure of PelA from *E. chrysanthemi*. (a) A ribbons-style representation with the three parallel  $\beta$ -sheets of the parallel  $\beta$ -helix shown in different colors: yellow for PB1, blue for PB2 and red for PB3. The two  $\beta$ -ribbons not participating in the parallel  $\beta$ -helix are shown in green. The disulfide bond is drawn as an orange thunderbolt. The N- and C-termini are labeled 'N' and 'C', respectively. (b) Stereo diagram of the  $C^\alpha$  chain trace of PelA in the *C2* crystal form. Residue numbers throughout the molecule are indicated.



**Figure 3**

Stereoview of the electron-density map. The  $2F_o - F_c$  electron density, calculated with the final model of the *C2* crystal form, is shown in the vicinity of the T1.5 loop. In this orientation, the polypeptide chains have their N-terminus on the left. The top protein segment is of residues 244–247 and the bottom fragment shows residues 213–223, both with gold electron density. Solvent molecules are represented as small red spheres with blue electron density. The protein model is shown in sticks, with standard CPK element colors.



**Figure 4**

Stereoview of the superposition of the backbone polypeptide trace of PelA and PelE from *E. chrysanthemi*. The backbone trace is drawn through all  $C^\alpha$  atoms. PelA is drawn in red, PelE in blue.

residues of molecule *A* and molecule *B* in the *R3* crystal of the current model refined without NCS restraints is 0.250 Å.

The solvent structure of both the *C2* and *R3* crystal forms contain sulfate ions. The *R3* crystals were grown in the presence of  $\text{Li}_2\text{SO}_4$ , but there is no apparent source of sulfate in the *C2* crystallization conditions. It is presumed that sulfate was introduced during the hydrophobic interaction chromatography owing to elution by ammonium sulfate. All sulfate ions in both crystal forms are observed near the active-site groove area and none are superimposable with the sulfate ions reported in the PelC structure (PDB code 1air). In the *C2* crystal form one sulfate ion is observed, whereas in the *R3* crystal form two sulfate ions in each molecule in the asymmetric unit are apparent. One of these sulfate ions has similar hydrogen bonding to protein and water atoms as the sulfate in the *C2* crystal form.

### 3.3. Structural comparison to other Pel isoforms

The polypeptide folds of PelA and PelE are strikingly similar (Fig. 4): superposition of the  $C^\alpha$  coordinates of PelA

and PelE gives an r.m.s.d. of 0.728 Å between residues 36–361 in PelA, discounting small regions of insertion/deletions and the 215–220 loop of PelA. Secondary-structure elements are highly conserved, both in number of  $\beta$ -sheets, number of  $\beta$ -strands and length of each  $\beta$ -strand. The disulfide bond in PelA is in a similar location and orientation as observed in the PelE structure.

Several structural features characteristic of other pectate lyases (Yoder, Lietzke *et al.*, 1993) are observed in PelA. PelA contains an Asn ladder in the interior of the T2 turn, which is identical to that observed in PelE. The PelA Asn ladder is composed of three Asn residues (233, 255 and 285) on PB3.6, PB3.7 and PB3.8, flanked by two Ser residues (314 and 201) at each end of the ladder. The aromatic stack on PB3 adjacent to the Asn ladder is different in PelA, PelC and PelE. Both PelA and PelC have three aromatic residues in this stack, while PelE has four residues in the stack. Unlike PelC and PelE, all residues in the aromatic stack of PelA (235, 257 and 287) are Phe. The position of the fourth aromatic residue in PelE is an Ile in both PelA and PelC. All *Erwinia* Pels exhibit four aliphatic stacks in the interior of the molecule between the  $\beta$ -sandwich formed by PB1 and PB2. The three

major  $\alpha$ -helices observed peripheral to the core parallel  $\beta$ -helix in PelA are superimposable on helices found in PelE.

PelA exhibits a unique second Asn ladder on the exterior side of the parallel  $\beta$ -helix. This second ladder consists of three Asn residues, 200, 232 and 254, and is located on T2.5, T2.6 and T2.7 on the immediate N-terminal residue from the previously described Asn ladder. The hydrogen-bonding pattern of the second Asn ladder utilizes two water molecules. The hydrogen-bond donor and acceptor atoms of the stacked Asn residues are too far apart to form hydrogen bonds; the water molecules appear to form hydrogen bond bridges to compensate. At the C-terminal end of the ladder, Asn254  $N^{\delta 1}$  forms hydrogen bonds with carbonyl O atoms from residues 284 and 342, while at the N-terminal end of the ladder, Asn200  $N^{\delta 1}$  forms a hydrogen bond with Asp15  $O^{\delta 1}$ .

### 3.4. Active-site differences

The active site is located in a long groove defined by PB1 and parts of the T3 loops (Scavetta *et al.*, 1999). In the superposition of PelE and PelA, there are no insertions or

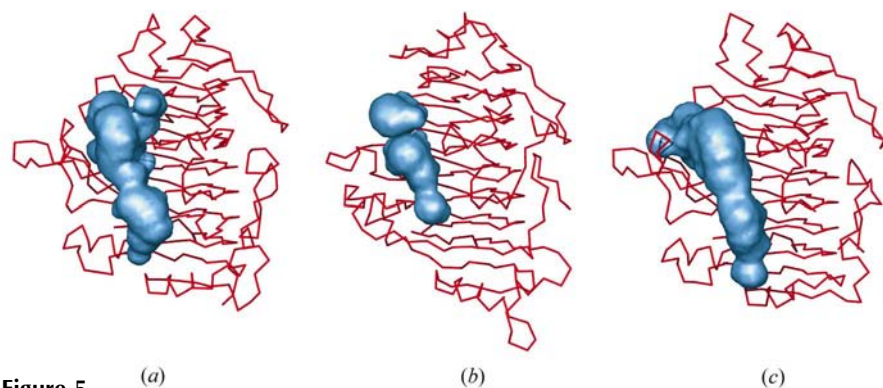
deletions of residues within this area. The amino-acid residues within the groove are highly conserved between PelA and PelE: 83% of the residues are absolutely conserved.

The volume of the saccharide-binding groove in PelA is approximately 2530 Å<sup>3</sup>. This is similar to the volume in PelE (2930 Å<sup>3</sup>), but about two times larger than the same groove in PelC (1230 Å<sup>3</sup>) (Fig. 5). The coordinates of the tetra-galacturonate fragment determined in the PelC–pentagalacturonate structure fit into the groove defined for PelA.

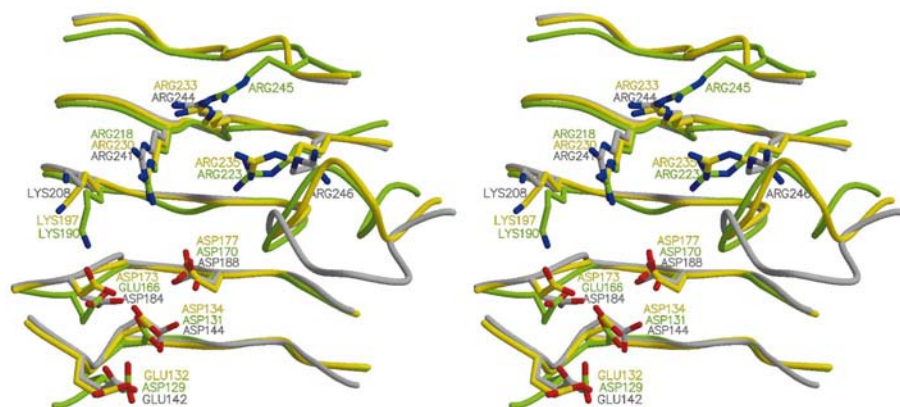
PelA, PelC and PelE were superimposed and amino acids demonstrated to interact with a bound pectate fragment in PelC were analyzed (Fig. 6). Two Arg residues that interact with the pectate fragment by forming hydrogen bonds with galacturonates on both sides of the putative scissile bond (Scavetta *et al.*, 1999) exhibit structural differences. Using the PelC–pentagalacturonate complex as a reference, the residues are (i) Arg223 on PB1.6 and (ii) Arg245 on PB1.7. In the first case, the orientation of the side chain of PelC Arg223 differs in PelA (Arg246) relative to both PelC and PelE (Arg235). This is likely to be attributable to steric constraints at the T1.5 loop

composed of residues 214–224 in PelA (203–213 in PelE) (Fig. 6). In PelA, the Arg246 side chain is oriented away from the saccharide-binding site. In both PelE and PelC, the T1.5 loop constrains this Arg residue to orient towards the substrate-binding site. The T1.5 loop conformation in PelA may be stabilized by a salt bridge between Asp216 O<sup>δ1</sup> and Arg191 NH1. Although the Asp residue is conserved in PelE (there is no structurally superimposable residue in the PelC structure), Arg191 is not conserved in either of the other two isozymes (Lys180 in PelE and Gly173 in PelC). The T1.5 loop from 214–224 cannot be modeled in the PelA R3 crystal form, yet the orientation of Arg246 is the same as that observed in the C2 crystal form.

The second structurally different saccharide-binding residue involves Arg244 of PelA. The structurally similar residue in PelE is Arg233. In both PelE and PelA, the hydrogen-bond donor groups of the side-chain atoms superimpose with the side-chain atoms of PelC Arg245. The main-chain atoms of PelC Arg245 reside on the β-strand immediately adjacent on the C-terminal side of PB1 (PB1.6) relative to that of PelA244 and PelE234. Although the arginine residues in the three isozymes are not necessarily contributed from the same backbone positions, the side-chain atoms are superimposable and are likely to be able to interact with the substrate in a similar manner.



**Figure 5** Active-site groove of PelA, PelC and PelE. The masked active-site groove is displayed in cyan, with the C<sup>α</sup> chain trace of the proteins in red. (a) PelA, (b) PelC, (c) PelE.



**Figure 6** Stereoview of the superposition of amino acids implicated in saccharide and calcium binding. Selected residues demonstrated to interact with a bound pectate fragment and with calcium ions in the structure of PelC–pentagalacturonate (Scavetta *et al.*, 1999) are shown with side-chain atoms. Homologous residues based on the structural superposition of PelA and PelE are also shown. PelA is drawn with a grey backbone trace, gray text numbering and gray carbon sticks; PelC is drawn with a green backbone trace, green text numbering and green carbon sticks; PelE is drawn with a yellow backbone trace, yellow text numbering and yellow carbon sticks.

### 3.5. Electrostatic differences

Although there is highly conserved backbone superposition of PelE and PelA, the electrostatic charge distribution around the active-site groove is significantly different between the isozymes. PelA is the only acidic Pel in the PelABCDE family. An electrostatic surface-contour plot was used to compare charge distributions in PelA, PelC and PelE (Fig. 7). PelA has the most acidic active-site groove, which is consistent with known differences in amino-acid composition and measured pI values. In *E. chrysanthemi*, the Pel isozymes have pI values between 4.6 (PelA) and 9.8 (PelE) (Tamaki *et al.*, 1988). The increased acidic nature of PelA arises from a relative decrease in the number of basic residues, particularly in the N-terminal end of the active-site groove. The charge-distribution difference is most noticeable in the active-site groove defined by the first two β-sheets of PB1 (PB1.1 and PB1.2) and the preceding T3 turns.

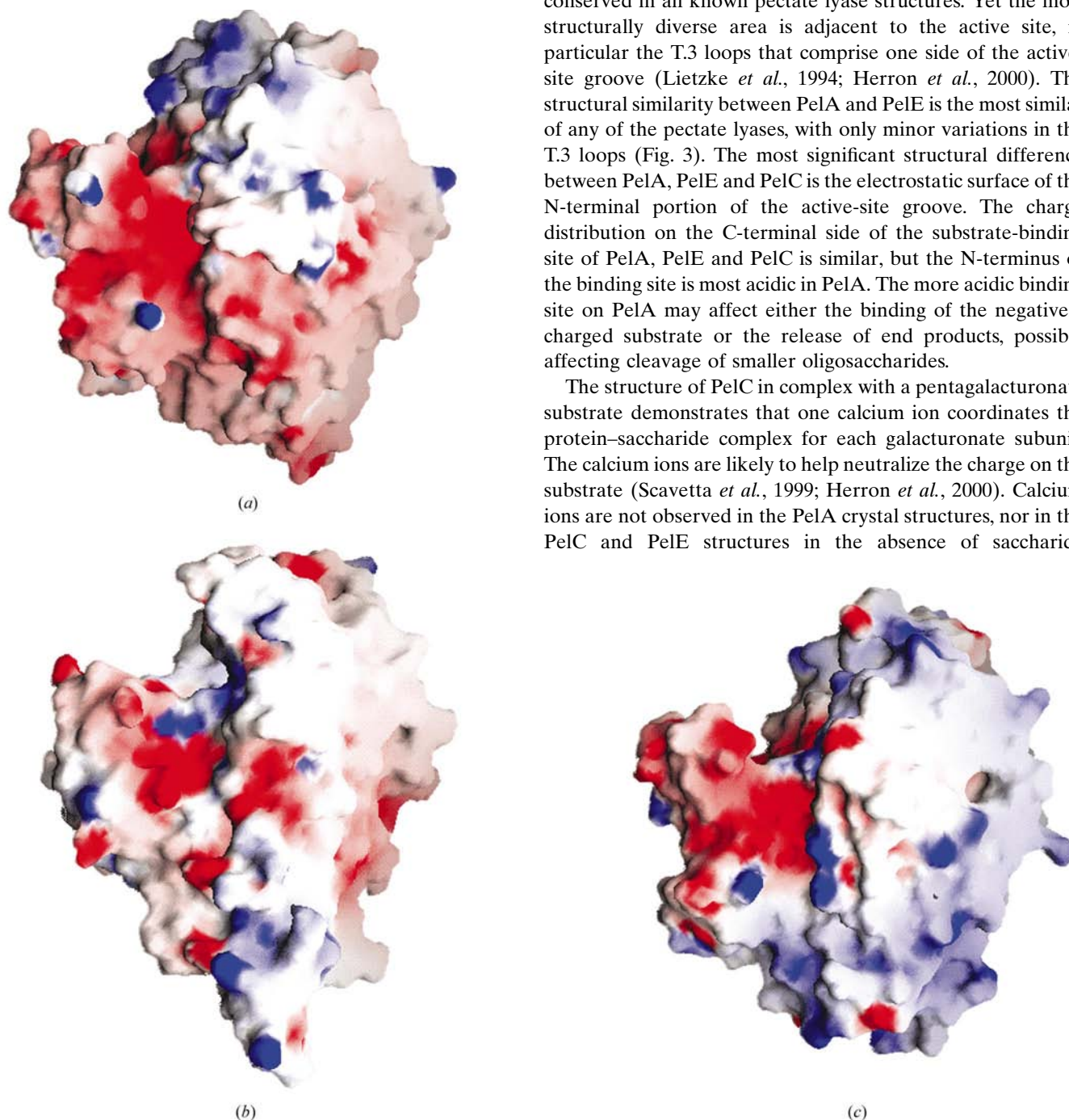
#### 4. Conclusions

The structure determination of PelA was undertaken to provide structural explanations of the unique physical and enzymatic properties of PelA. The results presented provide three apparent structural features in the structure of PelA in comparison to the other *E. chrysanthemi* isozymes that may provide insight. The features of interest include (i) charge distribution in the substrate-binding groove, (ii) size and volume of the substrate-binding groove and (iii) orientation of

a pivotal Arg involved in substrate binding adjacent to the scissile bond of the substrate.

The superposition of the polypeptide chains of the two isozymes PelA and PelE illustrates a high degree of similarity. The genes for the two proteins reside within the same gene cluster and share 222 identical amino acids (61.5%). Despite a sequence identity of 83% in the saccharide-binding groove, there is a significant difference in the electrostatic surface around the substrate-binding groove. It has been observed that the immediate region around the catalytic site is highly conserved in all known pectate lyase structures. Yet the most structurally diverse area is adjacent to the active site, in particular the T.3 loops that comprise one side of the active-site groove (Lietzke *et al.*, 1994; Herron *et al.*, 2000). The structural similarity between PelA and PelE is the most similar of any of the pectate lyases, with only minor variations in the T.3 loops (Fig. 3). The most significant structural difference between PelA, PelE and PelC is the electrostatic surface of the N-terminal portion of the active-site groove. The charge distribution on the C-terminal side of the substrate-binding site of PelA, PelE and PelC is similar, but the N-terminus of the binding site is most acidic in PelA. The more acidic binding site on PelA may affect either the binding of the negatively charged substrate or the release of end products, possibly affecting cleavage of smaller oligosaccharides.

The structure of PelC in complex with a pentagalacturonate substrate demonstrates that one calcium ion coordinates the protein–saccharide complex for each galacturonate subunit. The calcium ions are likely to help neutralize the charge on the substrate (Scavetta *et al.*, 1999; Herron *et al.*, 2000). Calcium ions are not observed in the PelA crystal structures, nor in the PelC and PelE structures in the absence of saccharide



**Figure 7**

Electrostatic charge distribution of PelA, PelC and PelE (GRASP; Nicholls *et al.*, 1991). Regions of negative and positive potential are colored over an approximate range of  $-15$  (red) and  $+15$  (blue)  $k_B T$ , where  $k_B$  is the Boltzman constant and  $T$  is temperature. (a) PelA, (b) PelC, (c) PelE.

**Table 5**

Homologous carbohydrate and calcium-binding residues in superimposed structures of PelC–pentagalacturonate and PelA.

	PelA	PelC
Carbohydrate ligands	NR†	Asp162
	Lys208	Lys190
	His214	Ser196‡
	Arg241	Arg218
	Arg246	Arg223
	Arg244	Arg245
Calcium ligands	NR	Tyr268
	Glu142	Asp129
	Asp144	Asp131
	Asn140§	Asp160
	NR	Asp162
	Asp184	Glu166
	Asp188	Asp170
	Arg241	Arg218

† No residue superimposes at site. ‡ Main-chain O involved in carbohydrate binding. Carbonyl O atoms of these structurally homologous residues are superimposable.

§ Side chains of the two superimposed structures intersect, but main-chain atoms are not superimposable.

substrates. The PelC–pentagalacturonate structure identified seven charged residues involved in coordinating four calcium ions. Upon superposition with the PelA structure, homologous residues involved in saccharide and calcium binding are identified (Table 5). The one calcium protein ligand not present in PelA is Asp162 (PelC), which resides on the T3.4 loop in PelC. In PelA, the T3.4 loop has a different conformation such that there is no related residue that superimposes to PelC Asp162. The saccharide-binding residues are also well conserved, with the exception of PelC Tyr268, which forms a stacking interaction with one of the galacturonate-ring structures. In PelC, Tyr268 is on the T1.8 loop which is longer and reaches closer towards the active-site groove than the comparable T1.8 loop in PelA or PelE. PelA has two aromatic residues in this region, with the closest being Phe304, although the distance between the C<sup>γ</sup> atoms of the superimposed PelA and PelC aromatic residues is approximately 3.9 Å.

The orientation of homologous residues that interact with bound substrate as determined in the PelC–pentasaccharide complex is highly conserved in PelA, except for Arg246. In the PelE structure as well as the PelC structure in both the bound pentapectate and unbound form, this Arg residue is oriented so that the side-chain atoms are directed towards the active site. In PelC, the Arg side-chain atoms interact with the galacturonate adjacent to the scissile bond. The side-chain atoms of the homologous Arg residue in PelA, Arg246, are oriented away from the active site. If the Arg residue in PelA is still involved in substrate coordination in a similar manner as observed with PelC, then the side chain must reorient. The Arg orientation towards the active site is sterically permissible, but the T1.5 loop in PelA must move. It seems apparent that there is flexibility in this loop, particularly in consideration of the PelA R3 crystal form where electron density is not visible for this loop, presumably owing to conformational flexibility.

Interestingly, the end products of PelA enzymatic action on a polypectate substrate have the most unique oligomeric patterns in comparison to the other Pel isozymes. PelB, PelC and PelE cleave the substrate predominantly to either small di- or trisaccharides. PelA end products are more varied in length, from dimers to nonamers. Yet the active-site groove is most dissimilar between PelC and either PelA or PelE. The overall shape of the substrate-binding groove is quite similar between PelE and PelA. It does not appear that the length or the shape of the active-site groove strongly dictates the end products. It remains to be determined which residues or regions in the active-site groove dictate end-product cleavage profiles.

The high-expression construct of PelA was graciously provided by Noel T. Keen. We are indebted to Chris Snook and Lesa Beamer for collecting X-ray data of the C2 crystal form at the National Synchrotron Light Source at Brookhaven National Laboratory, beamline X8C. Coordinates of PelE and the PelC–pentagalacturonate structures were graciously provided by Frances Jurnak. This work is supported by a grant from the National Science Foundation (MDY).

## References

- Alfano, J. R. & Collmer, A. (1996). *Plant Cell*, **8**, 1683–1698.
- Barras, F., Thurn, K. K. & Chatterjee, A. K. (1987). *Mol. Gen. Genet.* **209**, 319–325.
- Brünger, A. T. (1993). *Acta Cryst.* **D49**, 24–36.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Collmer, A. & Keen, N. T. (1986). *Annu. Rev. Phytopathol.* **24**, 383–409.
- Doan, C. N., Caughron, M. K., Myers, J. C., Breakfield, N. W., Oliver, R. L. & Yoder, M. D. (2000). *Acta Cryst.* **D56**, 351–353.
- Herron, S. R., Benen, J. A. E., Scavetta, R. D., Visser, J. & Jurnak, F. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 8762–8769.
- Jones, T. A., Cowan, S., Zou, J.-Y. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- Laskowski, R. A. (1995). *J. Mol. Graph.* **13**, 323–330.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1992). *J. Appl. Cryst.* **26**, 283–291.
- Lietzke, S. E., Yoder, M. D., Keen, N. T. & Jurnak, F. (1994). *Plant Physiol.* **106**, 849–862.
- Navaza, J. (1994). *Acta Cryst.* **A50**, 157–163.
- Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pickersgill, R., Jenkins, J., Harris, G., Nasser, W. & Robert-Baudouy, J. (1994). *Nature Struct. Biol.* **1**, 717–723.
- Preston, J. F. III, Rice, J. D., Ingram, L. O. & Keen, N. T. (1992). *J. Bacteriol.* **174**, 2039–2042.
- Scavetta, R. D., Herron, S. R., Hotchkiss, A. T., Kita, N., Keen, N. T., Benen, J. A. E., Kester, H. C. M., Visser, J. & Jurnak, F. (1999). *Plant Cell*, **11**, 1081–1092.
- Tamaki, S. J., Gold, S., Robeson, M., Manulis, S. & Keen, N. T. (1988). *J. Bacteriol.* **170**, 3468–3478.
- Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.
- Yoder, M. D., Keen, N. T. & Jurnak, F. (1993). *Science*, **260**, 1503–1507.
- Yoder, M. D., Lietzke, S. & Jurnak, F. (1993). *Structure*, **1**, 241–251.