MPMI Vol. 8, No. 2, 1995, pp. 331–334.

For the Record

Structure-Based Multiple Alignment of Extracellular Pectate Lyase Sequences

Susan Heffron,[1] Bernard Henrissat,[2] Marilyn D. Yoder,[1] Susan Lietzke,[1] and Frances Jurnak[1]

[1]Department of Biochemistry, University of California, Riverside, California 92521 U.S.A.; [2]Centre de Recherches sur les Macromolecules Vegetales, Centre National de la Recherche Scientifique, BP 53X - 38041 Grenoble, France

Received 26 August 1994. Accepted 8 December 1994.

Pectate lyases are secreted virulence factors which degrade the pectate component of plant cell walls. The evolutionary-based multiple alignment of extracellular pectate lyases has been corrected using three-dimensional structural information derived from *Erwinia chyrsanthemi* pectate lyases C and E. The new multiple alignment reveals invariant amino acids likely to be involved in two different enzymatic functions.

Additional keywords: pectate lyase sequence alignment.

Pectate lyases (EC 4.2.2.2) are depolymerizing enzymes which degrade the pectate component of the plant cell wall, leading to tissue maceration and death. The most-studied enzymes are secreted from the phytopathogenic *Erwinia* species and are the primary virulence factors in 'soft-rot' diseases (Collmer and Keen 1986; Kotoujansky 1987). Sequence studies have identified four distinct families: the extracellular *pelADE* family; the extracellular *pelBC* family; the periplasmic pectate lyases from *Yersinia pseudotuberculosis* and *Erwinia carotovora* strains; and some fungal pectate lyases with a small molecular weight of 24 kDa (Gonzalez-Candelas and Kolattukudy 1992). Within each family, the genes share more than 50% sequence similarity, but only the *pelBC* and *pelADE* subfamilies appear to be evolutionarily related. In the latter subfamilies, three conserved sequence patterns have been identified which have been used to characterize a new DNA sequence as belonging to the pectate lyase family. The patterns include AxDIKGxxxxVTxS, VxxRxPxxRxGxxHxxxN, and vWiDH (Tamaki et al. 1988; Hinton et al. 1989; Hugouvieux-Cotte-Patat et al. 1992; Barras et al. 1994). The extracellular pectate lyases also share the patterns with pectin lyases (Gysler et al. 1990; Kuster-van Someren et al. 1992), plant pollen proteins (Wing et al. 1989; Rafnar et al. 1991), and plant style proteins (Budelier et al. 1990), but none of the latter proteins has yet been found to possess in vitro pectate lyase activity.

The three-dimensional structures of three pectate lyases have recently been determined: *E. chrysanthemi* pectate lyase C (PelC) (Yoder et al. 1993); *E. chrysanthemi* pectate lyase E (PelE) (Lietzke et al. 1994); and *Bacillus subtilis* pectate lyase (Pickersgill et al. 1994). All reveal that the enzymes have an unusual structure, that of parallel β strands folded into a large helix. Because the protein fold is novel, various studies have been undertaken to understand the unique features. The present study involves a superposition of the atomic coordinates for PelC and PelE using a least-square fit between αCs in the parallel β helix domain. The PILEUP multiple sequence alignment program of the Genetics Computer Group package has been used to align the sequences of other extracellular pectate lyases to the structurally aligned sequences of PelC and PelE. The multiple alignment of 14 extracellular pectate lyases is shown in Table 1. Of the 353 amino acids in PelC, 49 are invariant, using only evolutionary relationships (Hinton et al. 1989), but a different set of only 27 is invariant, using the structural information. An analysis of the discrepancies indicates that, in addition to new sequence information, those alignments which relied, in part (Barras et al. 1994) or in total (Hinton et al. 1989), on evolutionary relationships contain misplaced gaps in sequences in loop regions of PelC and PelE, but correctly align most sequences within the parallel β helix. As a consequence, the evolutionary-based multiple sequence alignments of pectate lyases and their homologues are, at best, 70% correct.

The corrected multiple sequence alignment reveals several new features in extracellular pectate lyases. Of the 27 invariant amino acids, 14 possess chemical properties that are compatible with catalysis (Zvelebil and Sternberg 1988). In PelC, these include T10, T92, D131, D144, H145, D170, T179, S181, K190, T206, R218, R223, H28, and K342. On the three-dimensional structures of PelC and PelE, the 'potentially catalytic' amino acids cluster into two widely separated regions, suggesting that there are two active sites. The first group, D131, D170, K190, R218, and R223, are found near the Ca^{2+} binding site on the protein, with the two invariant aspartic acids coordinating directly to Ca^{2+} (Pickersgill et al. 1994). Because Ca^{2+} is essential for in vitro catalytic activity, the region around the Ca^{2+} site is believed to be the pectinolytic active site (Yoder et al. 1993; Pickersgill et al. 1994). The remaining 'potentially catalytic' invariant amino acids are widely separated in the primary sequence but cluster spatially in a region surrounding the vWiDH sequence. Interestingly, the cluster includes, not only the

Corrresponding author: F. Jurnak, Department of Biochemistry, University of California, Riverside, California 92521 U.S.A.

Table 1. Multiple alignment of pectate lyase sequence[a]

Seq. Name		Res #
	[b]N-Terminal Branch ß Helix	
Pelc_Erwch	ATDT..... ....GGYAA.. ..TAGGNVTG AV.....SKT ATSMQDIVNI	32
Pelb_Erwch	.ADT..... ....GGYTK.. ..TDGGDVSG AV.....KKT ASSMQDIVNI	31
Pel2_Erwca	.ANT..... ....GGYAT.. ..TDGGEVSG AV.....KKT ARSMKEIVDI	31
Pel3_Erwca	.ANT..... ....GGYAT.. ..TDGGDVAG AV.....KKT ARSMQDIIDI	31
Pela_Erwca	.ANT..... ....GGYAT.. ..TDGGDVSG AV.....KKT ARSLQEIINL	31
Pelb_Erwca	.ANT..... ....GGYAT.. ..TDGGDVSG AV.....KKT ARSLQEIVDI	31
Pelc_Erwca	.ANT..... ....GGYAT.. ..TDGGDVAG AV.....KKT ARSMQDIIDI	31
Pela_Erwch	....AELVSD KALESAPT.. ....VGWASQN GFTTGGAA.. ..ATSDNIYI VTNISEFTSA	47
Peld_Erwch	.....TLQTT KATEAA.S.. ....TGWATQ. GGTTGGAK.. ..AASAKIYA VKNISEFKAA	44
Pele_Erwch	.........A VETDAA.T.. ....TGWATQN GGTTGGAK.. ....AAKAVE VKNISDFKKA	39
Pelf_Erwch	....ASLQTT KATEAA.S.. ....TGWATQS GGTTGGAK.. ..ASSSKIYA VKSISEFKAA	46
Pel_Bacsub	....AD LGHQTLGSN. ....DGWGAYS TGTTGGSK.. ..ASSSNVYT VSNRNQLVSA	44
Pel_Psemar	....A DIWLDVAT.. ....TGWATQN GGTKGGSR.. ..AAANDIYT VKNAAELKKA	42
Pela_Emnid	....AR HELTRRQASE SCPIGYCTQN GGTTGGAA.. .....GDTVT VTNLADLTEA	45
	ß ß	
Pelc_Erwch	IDAARLDANG KKVKGGA... ..YPLVITYT GNEDSLINAA AANICGQWSK DPRG.....	81
Pelb_Erwch	IEAAKVDANG KKVKGGA... ..YPLVITYT GNEDSLINAA AANICGQWSK DARG.....	80
Pel2_Erwca	IEAAQVDSKG KKVKGGA... ..YPLIITYS GNEDSLIKAA EKNICGQWSK DARG.....	80
Pel3_Erwca	IEAAKLDSNG KKVKGGA... ..YPLVITYN GNEDALIKAA EANICGQWSK DARG.....	80
Pela_Erwca	IEEAQLDSKG KKLKGGA... ..YPLVITYN GNEDALIKAA EANICGQWSK DPRG.....	80
Pelb_Erwca	IEAAKVDSKG KKVKGGA... ..YPLIITYN GNEDSLIKAA EKNICGQWSK DARG.....	80
Pelc_Erwca	IEAAKLDSNG KKVKGGA... ..YPLVITYN GNEDALIKAA ENDICGQWKK DARG.....	80
Pela_Erwch	LS..... .....A GAEAKIIQIK GT..... ....IDISGG TPYTDFADQK ARSQ.....	82
Peld_Erwch	LN..... .....GT DTDPKIIQVT GA..... ....IDISGG KAYTSFDDQK ARSQ.....	80
Pele_Erwch	LN..... .....GT DSSAKIIKVT GP..... ....IDISGG KAYTSFDDQK ARSQ.....	75
Pelf_Erwch	LN..... .....GT DSSPKIIQVT GA..... ....IDISGG KAYTSFDDQK ARSQ.....	82
Pel_Bacsub	LG..... .....KET NTTPKIIYIK GT..... ....IDMNVD DNLKPLGLND	77
Pel_Psemar	LS..... .....ASA GSNGRIIKIT GI..... ....IDVSEG KVYTKTADMK VRGR.....	79
Pela_Emnid	AE..... ..... SDGPLTIIVS GS.....	59
	ß ß T3 ß T1	
Pelc_Erwch	.... .VEIKEFTK. .GITIIGANG S.SAN.FGIW I...KKSSDV	113
Pelb_Erwch	.... .VEIKDFTK. .GLTIIGANG S.SAN.FGIW I...VNSSDI	112
Pel2_Erwca	.... .VQIKEFTK. .GITIQGTNG S.SAN.FGVW I...VNSSNV	112
Pel3_Erwca	.... .VEIKEFTK. .GITIIGTNG S.SAN.FGIW L...TKSSDI	112
Pela_Erwca	.... .VEIKEFTK. .GLTILGTNG S.SAN.FGIW L...VNSSNV	112
Pelb_Erwca	.... .VQIKEFTK. .GITIQGTNG S.SAN.FGVW I...VNSSNV	112
Pelc_Erwca	.... .VEIKEFTK. .GITIIGTNG S.SAN.FGIW L...TKSSDI	112
Pela_Erwch	.... .INI.....P ANTTVIGLGT DAKFINGSLI IDGTDGTNNV	116
Peld_Erwch	.... .ISV.....P SNTTIIGIGS NGKFTNGSLV I...KGVSNV	111
Pele_Erwch	.... .ISI.....P SNTTIIGVGS NGKFTNGSLV I...KGVKNV	106
Pelf_Erwch	.... .ISI.....P SNTTIIGIGN KGKFTNGSLV V...KGVSNV	113
Pel_Bacsub	YKDPEYDLDK YLKAYDPSTW GKKEPSGTQE EARARSQKNQ KARVM..... .VDI.....P ANTTIVGSGT NAKVVGGNFQ I....KSDNV	152
Pel_Psemar	.... .LDI.....P GKTTIVGIGS NAEIREGFFY A....KENDV	109
Pela_Emnid	....ISGSA KIRV.....A SDKTIFGESG S.SITGIGFY I...RRVSNV	95
	ß ß T3 ß ß ß T3	
Pelc_Erwch	VVQNMRIGY. LPGGAK.... ....DGDMIRV DDSP.NVWVD HNELFAANHE CDGTPDNDT.	163
Pelb_Erwch	VVRNMRIGY. LPGGAQ.... ....DGDMFRI DNSP.NVWLD HNELFAANHE CDGTKDGDT.	162
Pel2_Erwca	VVRNMRFGY. MPGGAQ.... ....DGDAIRI DNSP.NVWID HNEIFAKNFE CKGTPDNDT.	162
Pel3_Erwca	VIRNMRFGY. MPGGAQ.... ....DGDAIRI DNTP.NVWID HNEIFAKNFE CAGTKDGDT.	162
Pela_Erwca	VVRNMRFGY. MPGGAK.... ....DGDAVRI DNSP.NVWID HNEIFAKNFE CAGTPDNDT.	162
Pelb_Erwca	VVRNMRFGY. MPGGAQ.... ....DGDAIRV DNSP.NVWID HNKIFAQSFE CKGTPDNDT.	162
Pelc_Erwca	VIRNMRFGY. MPGGAQ.... ....DGDAIRI DNTP.NVWID HNEIFAKNFE CAGTKDGDT.	162
Pela_Erwch	IIRNVYIQ.T PID...VEPH YEKGDG.... WNAEWDAMNI TNGAHHVWID HVTISD.... .....G NFTDDMYTTK DGETYV....	181
Peld_Erwch	ILRNLYIE.T PVD...VAPH YEEGDG.... WNAEWDAAVI DNST.RVWVD HVTISD.... .....G SFTDDKYTTK NGEKYV....	175
Pele_Erwch	ILRNLYIE.T PVD...VAPH YESGDG.... WNAEWDAAVI DNST.NVWVD HVTISD.... .....G SFTDDKYTTK DGEKYV....	170
Pelf_Erwch	ILRNLYIE.T PVD...VAPH YEEGDG.... WNAEWDAVVI DSTD.HVWVD HVTISD.... .....G SFTDDKYTTK NGEKYV....	177
Pel_Bacsub	IIRNIEFQ.D AYD...YFPQ WDPTDGSSGN WNSQYDNITI NGGT.HIWID HCTFND.... .....G SRPDSTSPKY YGRKY.....	219
Pel_Psemar	IIRNITVE.N PWD...PEPI FDKDDGADGN WNSEYDGLTV EGAN.NVWVD HVTFTD.... .....G RRTDD..... .....QNGT	170
Pela_Emnid	IMRNLKI..S KVD.AD.... ....NGDAIGI DASS.NVWVD HCDLSG....	130

(continued on next page)

[a] Sequences shown and residue numbers listed correspond to mature proteins. The difference between residue number for precursor and residue number for mature protein for each of the sequences shown is: Pelc_Erwch, 22; Pelb_Erwch, 22; Pel2_Erwca, 22; Pel3_Erwca, 22; Pela_Erwca, 22; Pelb_Erwca, 22; Pelc_Erwca, 22; Pela_Erwch, 32; Peld_Erwch, 31; Pele_Erwch, 30; Pelf_Erwch, 41; Pel_Bacsub, 21; Pel_Psemar, 29; Pela_Emnid, 22. Invariant residues in the alignment are shown in bold.

[b] The nomenclature describes the structural elements that are similar between PelC and PelE. The solid lines represent the residue range of a particular structural element found in PelC that is also structurally conserved in PelE. T1 refers to the polypeptide segment, either a turn or a loop, connecting parallel β strands 3 and 1. The portions of the T1 and T3 turns which are not structurally conserved in both PelC and PelE are represented by a dotted line. The * designates the position of the asparagine ladder found as the second residue of a 2-residue polypeptide connection between parallel β strands 2 and 3.

Table 1. *(Continued)*

Seq. Name		Res #
Pelc_Erwch	.....TFESA VDIKGASNTV TVSYNYIHGV KKVGLDGSSS SDTG.....RNITYHH NYYNDVNARL PLQRGGLVHA	229
Pelb_Erwch	.....TFESA IDIKKGATYV TISYNYIHGV KKVGLSGFSS SDTAE.....RNITYHH NIYSDVNARL PLQRGGNVHA	229
Pel2_Erwca	.....TFESA VDIKKGSTNV TVSYNYIHGI KKVGLSGASN TDTG.....RNLTYHH NIYRDVNSRL PLQRGGLVHA	228
Pel3_Erwca	.....TFESA IDIKKASTNV TISYNYIHGI KKVGLSGFSS SDTG.....RDLTYHH NIYDDVNARL PLQRGGQVHA	228
Pela_Erwca	.....TFESA VDIKKGSTNV TVSYNYIHGV KKVGLSGSSN TDTG.....RNLTHHH NIHSDVNSRL PLQRGGQVHA	228
Pelb_Erwca	.....TFESA VDIKKGSTNV TVSYNYIHGI KKVGLSGASN TDTG.....RNLTYHH NIYRDVNSRL PLQRGGLVHA	228
Pelc_Erwca	.....TFESA IDIKKASTNV TVSYNYIHGI KKVGLSGFSS SDTG.....RDLTYHH NIYDDVNARL PLQRGGQVHA	228
Pela_Erwch	.....QHDGA LDIKRGSDYV TISNSLIDQH DKTMLIGHND T....NSAQD KGKLHVTLFN NVFNRVTERA PRVRYGSIHS	252
Peld_Erwch	.....QHDGA LDIKKGSDYV TISSSRFELH DKTILIGHSD S....NGSQD SGKLRVTFHN NVFDRVTERT PRVRFGSIHA	246
Pele_Erwch	.....QHDGA LDIKKGSDYV TISYSRFELH DKTILIGHSD S....NGSQD SGKLRVTFHN NVFDRVTERA PRVRFGSIHA	241
Pelf_Erwch	.....QHDGS LDIKRGSDYV TVSNSRFELH DKTILIGHSD N....NGSQD AGKLRVTFHN NLFDRVGERT PRVRFGSVHA	248
Pel_Bacsub	....QHHDGQ TDASNGANYI TMSYNYYHDH DKSSIFGSSD S....KTSDD .GKLKITLHH NRYKNIVQRA PRVRFGQVHV	290
Pel_Psemar	EHERPK.... ....QHHDGA LDVKNGANFV TISYSVFKSH EKNNLIGSSD S....RTTDD .GKLKVTIHN TLFENISARA PRVRYGQVHL	247
Pela_Emnid	......DLSG GKDD..LDGL VDISHGAEWI TVSNTYFHDH WKGSLIGHSD N....NEDED LGHLHVTYAN NYWYNVYSRT PLIRFATVHI	208

Pelc_Erwch	YNNLYTNI.. ......TGSG LNVRQNGQAL IENNWFEKAI ....NP VTSRYDGKNF G.....	274
Pelb_Erwch	YNNLYTGI.. ......TSSG LNVRQNGKAL IENNWFENAV ....SP VTSRYDGSNF G.....	274
Pel2_Erwca	YNNLYDGI.. ......TGSG FNVRQKGIAL IESNWFENAL ....NP VTARNDSSNF G.....	273
Pel3_Erwca	YNNLYTGI.. ......TSSG LNVRQKGIAL IERNWFENAK ....NP VTSRYDGSNF G.....	273
Pela_Erwca	YTNLYDGI.. ......KSSG FNVRQKGIAL IESNWFENAL ....NP VTARNDDSNF G.....	273
Pelb_Erwca	YTNLYDGI.. ......TGSG FNVRQKGIAL IESNWFEIAL ....NP VTARNDSSNF G.....	273
Pelc_Erwca	YNNLYTGI.. ......TSSG LNVRQKGIAL IERNWFENAK ....NP VTSRYDGSNF G.....	273
Pela_Erwch	FNNVFKG.DA KDPVYRYQYS FGIGTSGSVL SEGNSFTI.. ANL..... ..SASKACKV VKKF..... .NGS.....	307
Peld_Erwch	YNNVYLG.DV KNSVYPYLYS FGLGTSGTIL SESNSFTL.. SNLKSIDGKN PE.....CSI VKQF..... .NSK.....	305
Pele_Erwch	YNNVYLG.DV KHSVYPYLYS FGLGTSGSIL SESNSFTL.. SNLKSIDGKN PE.....CSI VKQF..... .NSK.....	300
Pelf_Erwch	YNNVYVG.DV NHKAYRYQYS FGIGTSGSLL SESNAFTI.. DNMKKISGRD KE.....CSV VKAF..... .NGK.....	307
Pel_Bacsub	YNNYYEG.ST SSSSYPFSYA WGIGKSSKIY AQNNVIDV.. ....PGLSAA KTISVFSGGT	343
Pel_Psemar	YNNYHVG.ST SHKVYPFSYA HGVGKNSKIF SERNAFEI.. ....AGISGC DKIAGDYGGS	300
Pela_Emnid	INNYWDSL.. ......IDTG VNCRMDAQVL IQSSAFHN..	238

Pelc_Erwch	.... TWVLKGNNIT KPADFSTYSI TWTADTKPYV NADSWTS...	311
Pelb_Erwch	.... TWVLKGNNIT KPADFATYNI TWTPDTKEYR NADTWTS...	311
Pel2_Erwca	.... TWELRNNNIT SPSDFAKYKI TWGKPSSPHI NADNWKS...	310
Pel3_Erwca	.... TWELRNNNVM SPADFAKYNI TWDKDSKPYV NAEDWKS...	310
Pela_Erwca	.... TWELRNNNIT SPSDFAKYKI TWGKPSTPHI NADDWKS...	310
Pelb_Erwca	.... TWELRNNNIT KPADFSKYNI TWGRPSTPHV NADDWKS...	310
Pelc_Erwca	.... TWELRNNNVM SPADFAKYNI TWDKDTKPYV NAEDWKS...	310
Pela_Erwch	.... IFSDNGSVL. ....NGS A.VDLSGCGF	328
Peld_Erwch	.... VFSDNGSLV. ....NGS STTKLDTCAV	327
Pele_Erwch	.... VFSDKGSLV. ....NGS TTTKLDTCGL	322
Pelf_Erwch	.... IFSDKGSII. ....NGA S.YNLNGCGF GF.....	330
Pel_Bacsub	.... ALYDSGTLL. ....NGT Q..... ..INASAANG ....LSSS..	368
Pel_Psemar	.... VYRDTGSTL. ....NGS A..... ..... LSCSWSSS..	321
Pela_Emnid	CPDRAIFFAD SDYTGYAVVD DVDLGGSSN. ....SV	369

Pelc_Erwch	...TGTF.PT VAYNY..... ...SPVSAQC V...KDKLPG YAGVGKNLAT LTSTACK	353
Pelb_Erwch	...TGTY.PT VPYSY..... ...SPVSAQC V...KDKLAN YAGVGKNLAT LASSACK	353
Pel2_Erwca	...TGKF.PS ISYKY..... ...TPVSAQC V...KDKLAN YAGVGKNLAV LTAANCK	352
Pel3_Erwca	...TGTF.AS VPYSY..... ...SPVSAQC V...KDKLAN YAGVNKNLAV LTAANCN	352
Pela_Erwca	...TGKF.PA VPYSY..... ...SPVSAQC V...KDKLAN YAGVGKNQAV LTAANCK	352
Pelb_Erwca	...TGKF.PS ISYKY..... ...TPVSAQC V...KDKLAN YAGVGKNLAV LTAANCK	352
Pelc_Erwca	...TGTF.AS VPYSY..... ...SPVSAQC V...KDKLAN YAGVNKNLAV LTAANCN	352
Pela_Erwch	...SAY.TSK IPYIY..... ...DVQPMT. .TELAQSITD NAGSGKL...	361
Peld_Erwch	...TAY.KPT LPYKY..... ...SAQTMT. .SSLASSINS NAGYGKL...	360
Pele_Erwch	...TAY.KPT LPYKY..... ...SAQTMT. .SSLATSINN NAGYGKL...	355
Pelf_Erwch	...SAY.SAK IPYKY..... ...SAQTIT. .TSLAGSISS NAGYGKL...	363
Pel_Bacsub	...VGW.TPS L..... ....KSNVIN QAGAGKLN..	399
Pel_Psemar	...IGW.TP. .PYSY..... ...TPLAADK V...AADVKA KAGAGKL...	351
Pela_Emnid	PEGTLT.PSS LPY..... AAITALGSGQ V...ASVIPG TAGQ.KL...	304

'potentially catalytic' invariant amino acids, but many other invariant residues found in the amino (N)-terminal branch (GyatxxxxTxGG), the midregion of the parallel β helix (vWiDH; TxS; HaynN) and the carboxy (C)-terminal branch of the enzymes (AGxxK), all folded together in three dimensions. The vWiDH cluster is separated from the Ca^{2+} site by approximately 95 Å, too far to be involved in the pectinolytic function. Thus, the identification of a second unique cluster of 'potentially catalytic' residues suggests that the pectate lyases have a second, as yet unidentified, enzymatic function. Although site-directed mutational studies will aid the search for a second function, the relative orientation of the invariant DH side chains in the PelC and PelE structures is reminiscent of active sites found in many protease structures (Branden and Tooze 1991).

An analysis of the invariant amino acids lacking in catalytic properties is also revealing. Two, I54 and N232 in PelC, are involved in the internal side-chain stacking interactions which appear to confer stabilizing properties on the pectate lyases in solution as well as in a hostile extracellular environment. The invariant proline, P220 in PelC, has an unusual *cis* conformation in the three-dimensional structures, apparently to properly align the invariant arginine, R218 in PelC, for an important catalytic function. At 26%, the percentage of invariant glycines is unusually high, indicating crucial roles in controlling conformational changes or in avoiding steric clashes. Most notable is the GyatxxxxTxGG sequence in the N-terminal branch. A database search reveals that a similar pattern is found in many precursor sequences (personal observations). Andersson and von Heijne (1991) have predicted that exported proteins contain sequence signals in the first 30 amino acids following the N-terminal signal peptide which is removed during or after export of the protein. Indeed, they found that export was blocked by the insertion of a string of lysines within the first 30 amino acids. From an enzymatic and structural viewpoint, it is also logical to expect that some type of signal may exist. For example, during enzymatic cleavage of an internal bond, in this case, the cleavage of the peptide bond between the signal sequence and the mature protein, there are generally groups on both sides of the scissile bond which contribute to the recognition of the substrate by the enzyme. The structural comparison of PelC and PelE suggest that the GyatxxxxTxGG sequence may form part of the signal on the mature protein side of the scissle bond. Not only is the unusual conformation of the N-terminal branch structurally similar, but the conserved aromatic group, Y7 in PelC, is tightly bound in an apparent specificity pocket lined with conserved aromatic groups. Moreover, a small rotation about the ϕ bond of the invariant glycine preceding the conserved aromatic group brings the N terminus to the invariant histidine in the vWiDH cluster. The invariant residues in the TxGG sequence are linked to the invariant tryptophan in the vWiDH sequence through a hydrogen bonding network. Thus, the invariant amino acids in the N-terminal branch appear to have some type of alignment function for a catalytic activity occurring at the vWiDH sequence.

In summary, a structure-based multiple alignment of extracellular pectate lyase sequences corrects errors in the evolutionary-based alignments and reveals new sequence patterns which suggest a second catalytic function for the pectate lyases.

ACKNOWLEDGMENTS

The research is supported by the United States Department of Agriculture (award #94-37303-0730) and the San Diego Supercomputer Center.

LITERATURE CITED

Andersson, H. and von Heijne, G. 1991. A 30-residue-long "export initiation domain" adjacent to the signal sequence is critical for protein translocation across the inner membrane of *Escherichia coli*. Proc. Natl. Acad. Sci. USA 88:9751-9754.

Barras, F., van Gijsegem, F., and Chatterjee, A. K. 1994. Extracellular enzymes and pathogenesis of soft-rot *Erwinia*. Annu. Rev. Phytopathol. 32:201-34.

Branden, C. and Tooze, J. 1991. Introduction to Protein Structure. Garland Publishing Co., New York. pp. 231-246.

Budelier, K. A., Smith, A. G., and Gasser, C. S. 1990. Regulation of a stylar transmitting tissue-specific gene in wild-type and transgenic tomato and tobacco. Mol. Gen. Genet. 224:183-192.

Collmer, A., and Keen, N. T. 1986. The role of pectic enzymes in plant pathogenesis. Annu. Rev. Phytopathol. 24:383-409.

Gonzales-Candelas, L., and Kolattukudy, P. E. 1992. Isolation and analysis of a novel inducible pectate lyase gene from the phytopathogenic fungus *Fusarium solani* f. sp. *pisi* (*Nectria haematococca*, mating population VI). J. Bacteriol. 174:6343-6349.

Gysler, C., Harmesen, J. A., Kester, H. C., Visser, J., and Heim, J. 1990. Isolation and structure of the pectin-lyase D-encoding gene from *Aspergillus niger*. Gene 89:101-108.

Hinton, J. C. D., Sidebotham, J. M., Gill, D. R., and Salmond, G. P. D. 1989. Extracellular and periplasmic isoenzymes of pectate lyase from *Erwinia carotovora* subspecies *carotovora* belong to different gene families. Mol. Microbiol. 3:1785-1795.

Hugouvieux-Cotte-Pattat, N., and Robert-Baudouy, J. 1992. Analysis of the regulation of the *pelBC* genes in *Erwinia chrysanthemi* 3937. Mol. Microbiol. 6:2363-2376.

Kotoujansky, A. 1987. Molecular genetics of pathogenesis by soft-rot *Erwinias*. Annu. Rev. Phytopathol. 25:405-430.

Kuster-van Someren, M., Flipphi, M., de Graff, L., van den Broeck, H., Kester, H., Hinnen, A. and Visser, J. 1992. Characterization of the *Aspergillus niger pelB* gene: Structure and regulation of expression. Mol. Gen. Genet. 234:113-120.

Lietzke, S. E., Yoder, M. D., Keen, N. T., and Jurnak, F. 1994. The three-dimensional structure of pectate lyase E, a plant virulence factor from *Erwinia chrysanthemi*. Plant Physiol. 106:849-862.

Pickersgill, R., Jenkins, J., Harris, G., Nasser, W., and Robert-Baudouy, J. 1994. The structure of *Bacillus subtilis* pectate lyase in complex with calcium. Nature Struct. Biol. 1:717-723.

Rafnar, T., Griffith, I. J., Kuo, M. C., Bond, J. F., Rogers, B. L., and Klapper, D. G. 1991. Cloning of *Amb aI* (antigen E), the major family of short ragweed pollen. J. Biol. Chem. 266:1229-1236.

Tamaki, S. J., Gold, S., Robeson, M., Manulis, S., and Keen, N. T. 1988. Structure and organization of the *pel* genes from *Erwinia chrysanthemi* EC16. J. Bacteriol. 170:3468-3478.

Wing, R. A., Yamaguchi, J., Larabell, S. K., Ursin, V. M. and 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. 14:17-28.

Yoder, M. D., Keen, N. T., and Jurnak, F. 1993. New domain motif: The structure of pectate lyase C, a secreted plant virulence factor. Science 260:1503-1507.

Zvelebil, M. J. J. M. and Sternberg, M. J. E. 1988. Analysis and prediction of the location of catalytic residues in enzymes. Prot. Eng. 2:127-138.