# The crystal structure of a major allergen from plants

Kurt S Thorn[1], Hans EM Christensen[2], Ron Shigeta Jr[1], Don Huddler Jr[1], Lamaat Shalaby[3], Uno Lindberg[4], Nam-Hai Chua[5] and Clarence E Schutt[1]*

**Background**: Profilins are small eukaryotic proteins involved in modulating the assembly of actin microfilaments in the cytoplasm. They are able to bind both phosphatidylinositol-4,5-bisphosphate and poly-L-proline (PLP) and thus play a critical role in signaling pathways. Plant profilins are of interest because immunological cross-reactivity between pollen and human profilin may be the cause of hay fever and broad allergies to pollens.

**Results**: The determination of the *Arabidopsis thaliana* profilin isoform I structure, using multiwavelength anomalous diffraction (MAD) to obtain structure-factor phases, is reported here. The structure of *Arabidopsis* profilin is similar to that of previously determined profilin structures. Conserved amino acid residues in profilins from plants, mammals, and lower eukaryotes are critically important in dictating the geometry of the PLP-binding site and the overall polypeptide fold. The main feature distinguishing plant profilins from other profilins is a solvent-filled pocket located in the most variable region of the fold.

**Conclusions**: Comparison of the structures of SH3 domains with those of profilins from three distinct sources suggests that the mode of PLP binding may be similar. A comparison of three profilin structures from different families reveals only partial conservation of the actin-binding surface. The proximity of the semi-conserved actin-binding site and the binding pocket characteristic of plant profilins suggests that epitopes encompassing both features are responsible for the cross-reactivity of antibodies between human and plant profilins thought to be responsible for type I allergies.

Addresses: [1]Henry H Hoyt Laboratory, Department of Chemistry, Princeton University, Princeton NJ 08544, USA, [2]Laboratory of Plant Cell Biology, Institute of Molecular Agrobiology, National University of Singapore, Science Park, Singapore 0511, [3]DuPont Agricultural Products, Experimental Station, Wilmington, DE 19880-0402, USA, [4]Department of Zoological Cell Biology, WGI, Arrhenius Laboratories for Natural Sciences, Stockholm University, S-10691 Stockholm, Sweden and [5]Laboratory of Plant Molecular Biology, The Rockefeller University, New York, NY 10021, USA.

*Corresponding author.
E-mail: schutt@chemvax.princeton.edu

Key words: actin binding, allergies, *Arabidopsis thaliana*, MAD, plant profilin, poly-L-proline binding

Received: 31 July 1996
Revisions requested: 10 September 1996
Revisions received: 9 October 1996
Accepted: 15 October 1996

Electronic identifier: 0969-2126-005-00019

Structure 15 January 1997, 5:19–32



## Introduction

Actin is the principal structural element of the microfilament system in eukaryotic cells. This system determines cell shape and motility and is involved in many processes where the conversion of chemical energy to mechanical work is required [1]. Actin filaments, as well as the actin monomer sequestering protein profilin, are found in close apposition to the inner surface of the plasma membrane and are found in particularly high concentrations in regions actively engaged in motile activity [2,3]. In plants, actin plays an essential role in many processes, including cytoplasmic streaming, positioning of organelles, growth tip extension, and pollen tubule growth [4,5].

Profilins, ubiquitous cytosolic actin-binding proteins with molecular weights in the 12–15 kDa range [6,7], appear to control the preferential addition of actin monomers at the barbed ends of microfilaments [8]. Profilin binds actin and poly-L-proline (PLP) simultaneously, an important but not fully understood reaction [9–11]. Phosphatidylinositol-4,5-bisphosphate ($PIP_2$), a component of the phosphatidylinositol cycle employed in cell signalling events [12], also binds to profilin but causes the profilin–actin complex to dissociate [7,13,14]. The localization and binding properties of profilin thus suggest that it acts at a critical control point in signaling pathways initiated by events at the plasma membrane and plays an important role in regulating the activity in the microfilament system and intracellular calcium levels. Growth factors binding to their cognate receptors not only turn on the transcription of new genes, but also cause an immediate activation of the microfilament system [15,16]. Small GTPases [17] and proteins possessing PLP-binding SH3 domains [18,19] have recently been shown to be central elements in the control of the microfilament system.

Profilins can be divided into four classes, depending on their source: plants, mammals, other eukaryotes, and viruses. Between the different classes of profilins sequence homology is low (~30%) but within a class homology is much higher (50 to 80%). *Arabidopsis thaliana* (mouse ear cress) has at least four profilin genes encoding four isoforms: isoforms I and II are expressed in all organs of the plant while isoforms III and IV are found in floral tissues only. In plants, profilin appears to perform many of the same roles as in animals. Birch (*Betula verrucosa*) pollen profilin exhibits similar effects on the polymerization of actins from animal sources *in vitro* as do animal profilins

[20]. In addition, the microinjection of mammalian profilin into the stamen hair cells of birch results in an abrupt cessation of cytoplasmic streaming and collapse of the cytoskeleton [4]. However, neither *Arabidopsis* profilins I nor II increase the rate of ATP exchange on α-actin *in vitro* [21], as mammalian and *Acanthamoeba* profilins do [22–24].

Type I hypersensitivities, characterized by rhinitis, conjunctivitis, and bronchial asthma affect 15% of the population of industrialized countries [25–27]. Type I allergic symptoms can be induced in sensitive individuals by plant profilin preparations. The further importance of plant profilins in generating allergies is demonstrated by the observation that IgEs from allergic patients cross-react with multiple plant profilins as well as human profilins [26,27]. The binding of IgE to human profilin can be blocked by incubation with birch profilin, indicating that the antibodies recognize common epitopes between human and birch profilin [25]. Furthermore, individuals hypersensitive to plant profilins show allergic reactions to human profilin, suggesting that continued sensitization, even in the absence of further exposure to plant profilins, is maintained by an autoimmune reaction to human profilin. Plant profilins are involved in 20–30% of type I allergies, contributing to a major health problem.

The structures of bovine profilin, both alone [28] and in complex with bovine actin [11], as well as the structure of *Acanthamoeba* profilins I and II [29–31] have been previously published. Together with the structure of a plant profilin reported here, a representative structure of each of the profilin classes (except viral profilins) is now known, allowing a comparison of profilins across the spectrum of organisms in which they are found. We report a structure-based alignment of mammalian, plant, and other profilins which shows that the overall profilin fold is well conserved. However, there are some differences between the profilin classes, most of which are located within three loop regions. Two of these loops form a solvent-filled pocket adjacent to the actin-binding site; this binding pocket is unique to plant profilins. It is proposed that immune epitopes which contain the plant-specific pocket and the semi-conserved actin-binding surface are responsible for immune cross-reactivity between human and plant profilins.

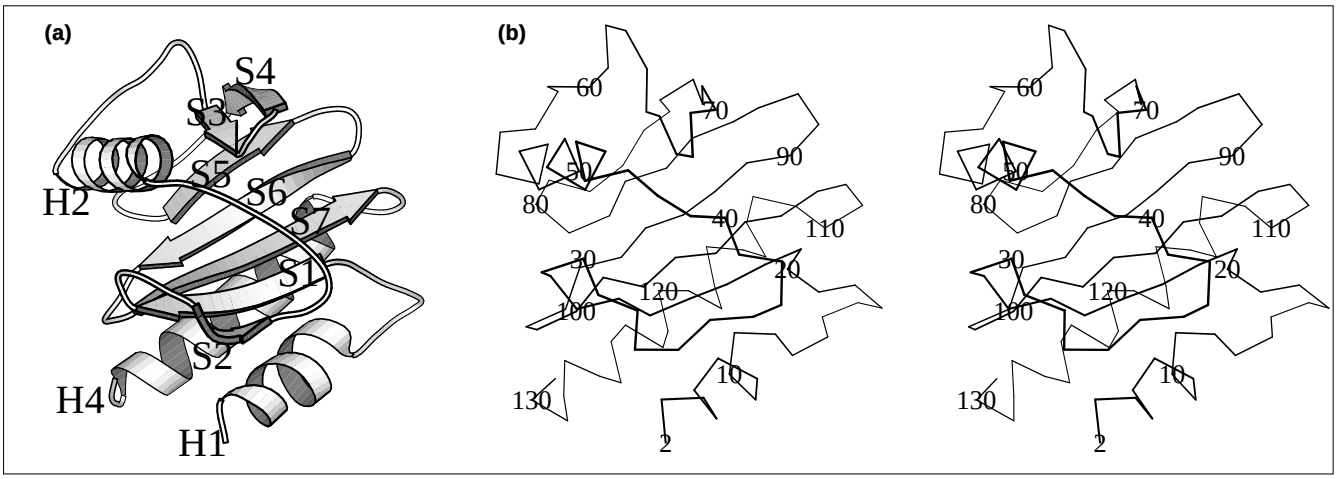
## Results and discussion

### Plant profilins are conserved with other classes of profilins

The profilin fold consists of a central seven-stranded β sheet sandwiched between the N- and C-terminal helices on one side, and two short consecutive helices on the other (Fig. 1) [11,28–31]. The amino acid sequence of *Arabidopsis* profilin is 39% identical to that of *Acanthamoeba* profilins I and II (SwissProt entries P07763 and P19984) and 30% identical to bovine profilin (SwissProt entry P02584) (Fig. 2). The overall tertiary structure of the profilins is conserved; the *Arabidopsis* profilin crystal structure shares the same fold as the *Acanthamoeba* and bovine structures and aligns to these structures with a mean deviation in Cα coordinates of 1.64 and 1.61 Å, respectively (Fig. 3). For the systematic comparison of the three profilin families, except where noted, the residue numbers will be adopted from the bovine profilin numbering system used in the alignment shown in Figure 2.

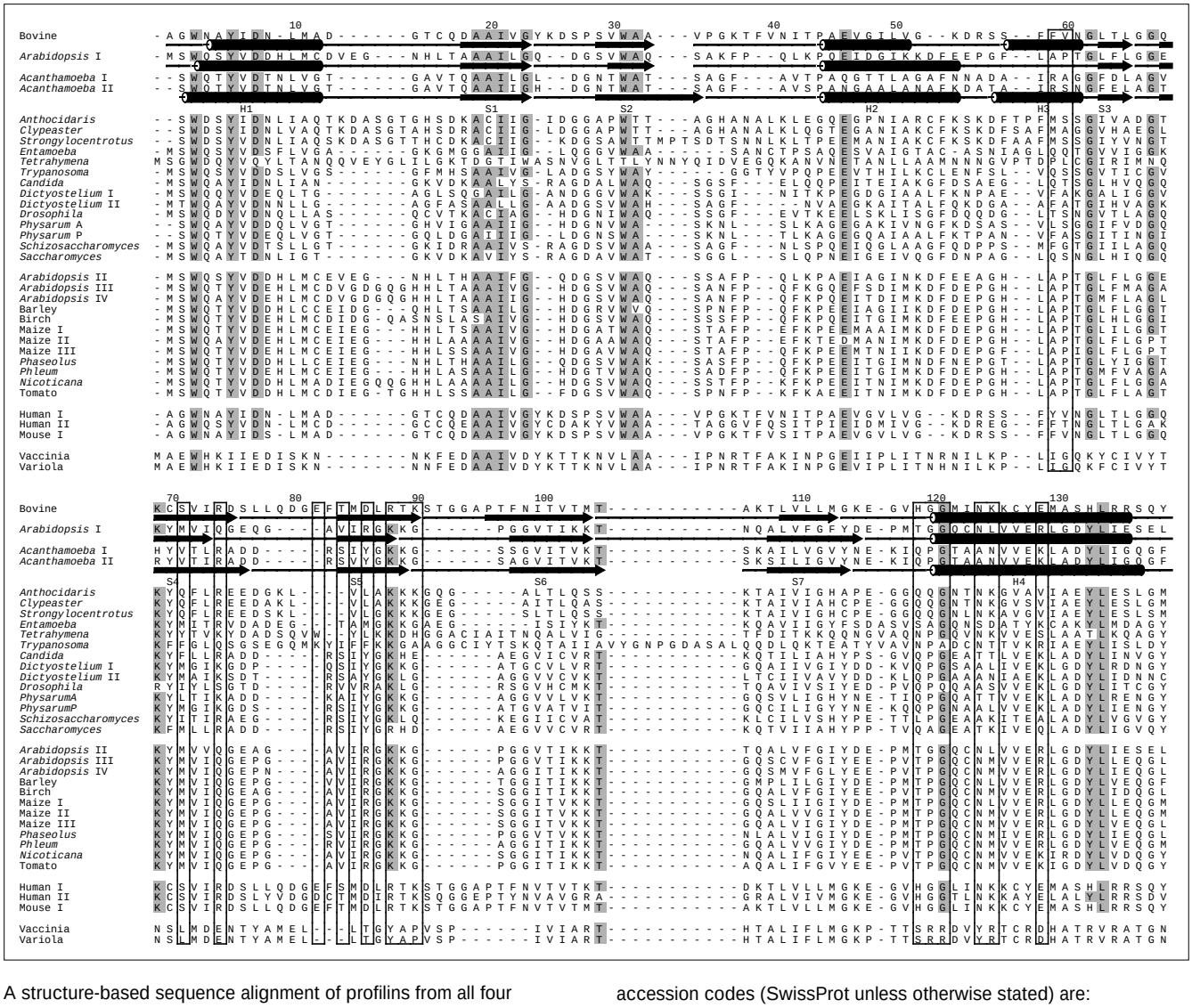
In the amino acid sequence and three-dimensional structure, three characteristic loops distinguish plant profilins from other profilins (Fig. 3). The first loop, between the N-terminal helix 1 and β strand 1, contains an insertion of

**Figure 1**




Overall structure of *Arabidopsis* profilin. (a) Ribbon diagram of the fold of *Arabidopsis* profilin I. Secondary structure elements are labeled: helices (H1, H2 and H4), β strands (S1–S7). (b) Stereo view Cα trace of *Arabidopsis* profilin I; every tenth residue is numbered. (Figure generated using the program MOLSCRIPT [69].)

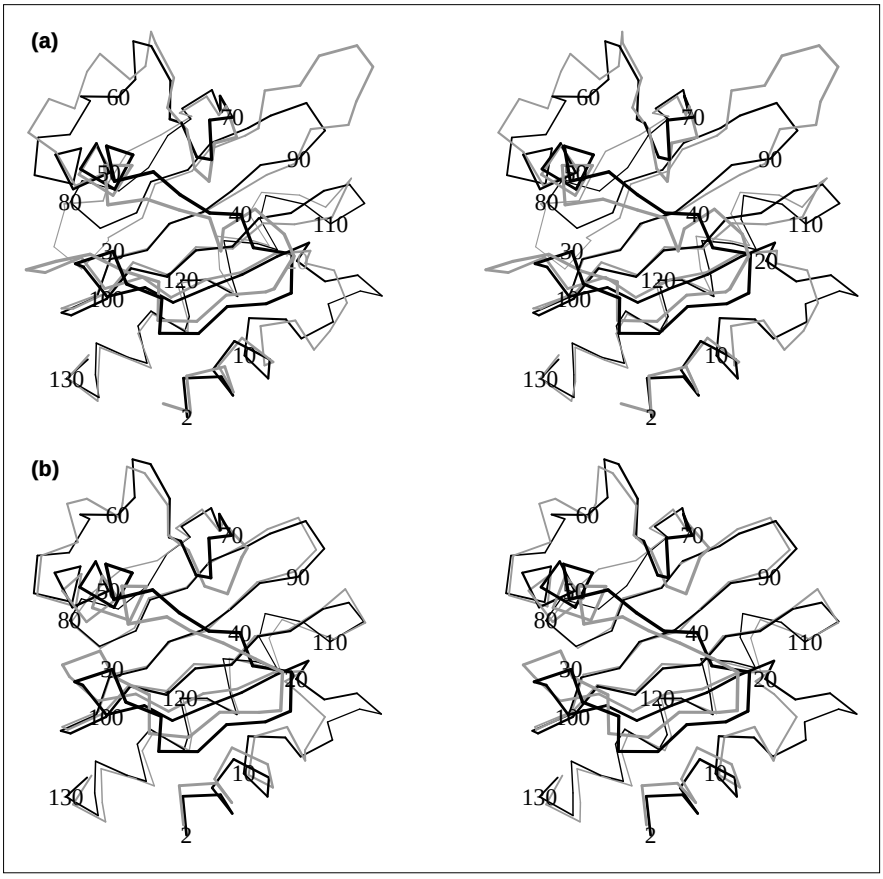
**Figure 2**



A structure-based sequence alignment of profilins from all four classes. The crystal structures of bovine profilin, *Acanthamoeba* profilins I and II, and *Arabidopsis* profilin I were superimposed by a least-squares fit of their Cα coordinates (see Materials and methods section) and the structurally homologous residues were then used as the basis for a sequence alignment. The residue numbering of the bovine structure is used for this figure. Secondary structures for the first four profilins are indicated below the primary sequence: cylinders for α helices and arrows for β sheets. The alignment was created with the program AMPS [70] and the table formatted with ALSCRIPT [71]. Amino acids which are conserved in more than 80% of profilin sequences are shaded in gray. Amino acids which have been identified as structurally homologous to the actin-binding site of bovine profilin are shown boxed. The remaining profilin sequences are divided into groups on the basis of sequence homology and taxonomy: non-plant and non-mammalian profilins; plant profilins; mammalian profilins; and viral profilins. The organisms and their accession codes (SwissProt unless otherwise stated) are: *Anthocidaris crassipina* (P18320); *Clypeaster japonicus* (P18321); *Strongylocentrotus purpuratus* (P32006); *Entamoeba histolytica* (P49230); *Tetrahymena pyriformis* (P23412); *Trypanosoma brucei* (NCBI 1311627); *Candida albicans* (Genbank L37834); *Dictyostelium discoideum* isoforms I and II (P26199, P26200); *Physarum polycephalum* isoforms A and P (P22271, P18322); *Schizosaccharomyces pombe* (P39825); *Arabidopsis thaliana* isoforms II, III, and IV [56]; *Hordeum vulgare* (barley) (Genbank U49505); *Betula verrucosa* (white birch) (P25816); *Zea mays* (maize) isoforms I, II, and III (P35081, P35082, P35083); *Phaseolus vulgaris* (kidney bean) (P49231); *Phleum pratense* (P35079); *Nicotiana tabacum* (tobacco) (P41372); *Solanum lycopersicum* (tomato) (Genbank U50195); Human isoforms I and II (P07737, P35080); Mouse (P10924); Vaccinia virus (P20844); and Variola virus (P33828). (Figure generated using the program MOLSCRIPT [69].)

three to six residues in plant profilins. In *Arabidopsis*, a water-mediated hydrogen bond links strands 1 and 7 of the β sheet (Fig. 4) at the edge of the actin-binding site. This water molecule is bonded to the amide of Asp115 and the carbonyl of His15 and is further stabilized by a hydrogen bond to the sidechain of residue Thr17, which is in turn

**Figure 3**




Stereo view comparison of *Arabidopsis* profilin (dark lines) with (a) bovine and (b) *Acanthamoeba* profilins (both in gray lines). Every tenth residue of the *Arabidopsis* profilin is numbered. The differences between the plant profilins and mammalian profilins are particularly well illustrated: the characteristic mammalian loop insert between residues 89 and 87; the plant inserts between β strands 4 and 5 and between β strands 5 and 6; and the *i→i+5* helical turn at residue 10. All three profilins show the most variation in the region between residues 20 and 65 running along the top of the figures. (Figure generated using the program MOLSCRIPT [69].)

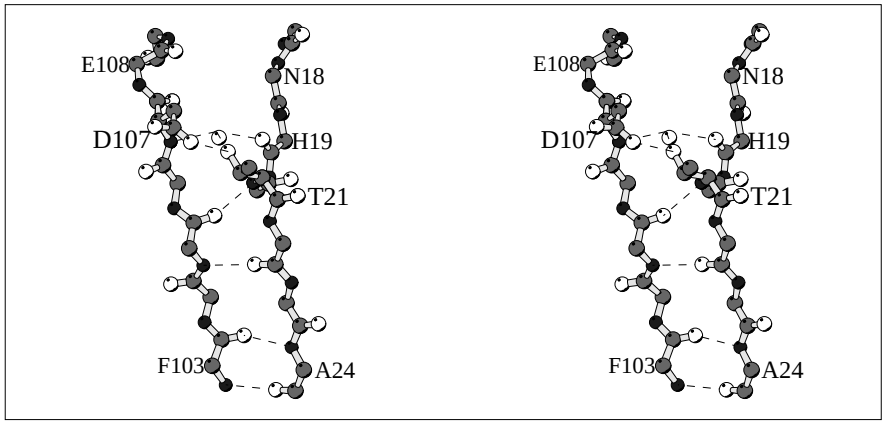
polarized by the carboxylate of Asp115. The water molecule stabilizes the β sheet where the hydrogen-bond donor and acceptor would otherwise be too far apart. This water molecule is not found in either the bovine or the *Acanthamoeba* structures, although both contain a similar bend in the β sheet. The water is well ordered, appearing in low temperature 2.8Å single isomorphous replacement multiwavelength anomalous diffraction (SIRMAD) electron-density maps, as well as in the room temperature structure. The remaining two loops, between β strands 4 and 5 and between strands 5 and 6, are characteristically shorter in plant profilins. These loops may modulate the interaction of profilin with actin, as discussed at length below.

In addition to the loops discussed above, the peptide chain between β strands 2 and 3 has little secondary structure and differs significantly between *Arabidopsis* profilin and other reported profilin structures. These residues, which correspond to helix 3 in bovine and *Acanthamoeba* profilin, are non-helical in *Arabidopsis* profilin and form a loop.

As was reported for *Acanthamoeba* profilin [29], an *i→i+5* hydrogen bond is found in the N-terminal helix of the *Arabidopsis* profilin structure. Structural alignments show that the underwound helix is due to a single amino acid insertion between positions 10 and 11 relative to bovine profilin, which has an α helix at the N terminus. Such α helix unwindings have been observed in mutants of T4 lysozyme [32] in which alanine residues have been inserted into the ends of helices. All non-mammalian profilins sequenced to date have this insertion (Fig. 2).

Despite the low homology between profilins in general, 18 residues are identical in 80% of sequenced profilins (Fig. 2). Of these, Ala19, Ala20, Glu46, Gly62, Gly67, and Thr105 appear to be intrinsic to the profilin fold. Ala19 and Ala20 are buried in the core of the protein and make close contacts with hydrophobic sidechains. Glu46 is also buried and hydrogen bonds to the amides of Gly66, Gly67, and Lys43. Only residues capable of making these hydrogen bonds can substitute at this location. Gly62 is in a type I turn and is the only residue with unusual ϕ/ψ values (116° and 156°, respectively). Gly67 is involved in a turn, as is Thr105; Thr105 stabilizes the turn via a hydrogen bond between its Oγ1 and the amide of Ala108. The remaining conserved residues play important roles in maintaining the PLP and actin-binding surfaces of profilin, which are described below and summarized in Table 1.

**Figure 4**




Stereo view showing the water mediated contact between β strands 1 and 7 of *Arabidopsis* profilin. The water molecule is hydrogen bonded to the amide of Asp107 and the carbonyl of His19 and is further stabilized by a hydrogen bond to the sidechain of residue Thr21; hydrogen bonds are shown as dashed lines. All residue numbers are in *Arabidopsis* sequence numbering. (Figure generated using the program MOLSCRIPT [69].)

Lys69 is conserved in 30 out of 35 profilins including *Arabidopsis*. This conservation has previously been recognized [29] and Lys69 was proposed to form part of the $PIP_2$-binding site. Of the other proposed residues in the *Acanthamoeba* profilin $PIP_2$-binding site, Lys88 (also a lysine residue in *Arabidopsis*) is conserved but Arg74 (glutamine in *Arabidopsis*) is not, although Arg86 is found in a similar location to Arg74 within the structure. In all non-viral profilins in which residue 74 is not basic, residue 86 is found to be basic instead. The residues analogous to 74 and 88 have been implicated in $PIP_2$-binding by mutagenesis studies on yeast and human profilin. However, while mutagenesis studies of yeast implicate residue 74 [33], mutagenesis of the corresponding residue in human profilin has no effect on $PIP_2$-binding [34]. Similarly, residue 88 is implicated in $PIP_2$-binding in mutagenesis studies of human profilin [34], but mutation of the analogous residue in yeast has no effect [33]. This comparison suggests that the localization of the $PIP_2$-binding site is not yet conclusive.

### A plant-specific binding pocket

A major structural difference between plant and other profilins is the presence of a solvent-filled pocket located near the actin-binding surface. One of the glycerol molecules from the cryoprotectant solvent was found to be specifically bound in the pocket with an average temperature factor of 13 $Å^2$. Its carbon backbone makes contact

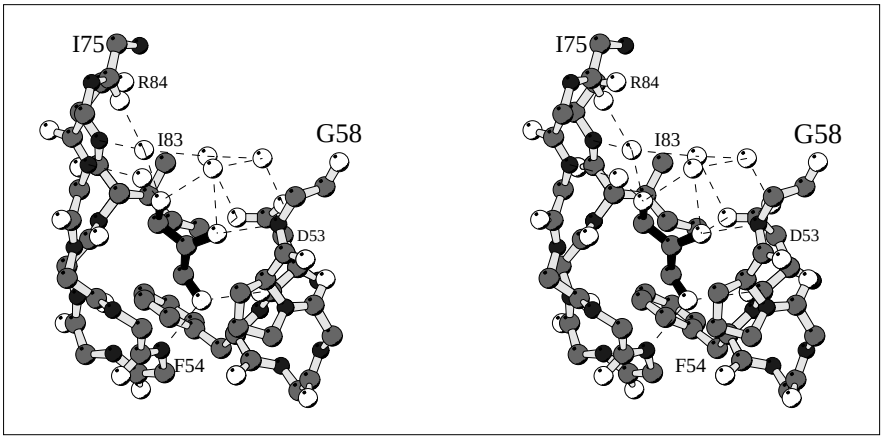
**Table 1**

**Biological function and structural role of the 18 residues shown to be identical in 80% of sequenced profilins.**

| Residue | Frequency* | Biological function and structural role† |
|---|---|---|
| W3 | 35 | PLP binding; direct contact to PLP |
| Y6 | 33 | PLP binding; direct contact to PLP |
| D8 | 32 | Unknown |
| A19 | 31 | Fold conservation; packed in core, contacts V7, L10, L16, M11 |
| A20 | 29 | Fold conservation; packed in core, contacts L23, F37, Q33 |
| I21 | 33 | PLP binding; packs behind W3 and W31 |
| G23 | 30 | PLP binding; packs against W31 |
| W31 | 32 | PLP binding; direct contact to PLP |
| A32 | 30 | PLP binding; packs against helix supporting W3 |
| E46 | 32 | Fold conservation; hydrogen bonds to N66, N67, N44 |
| G62 | 33 | Fold conservation; $\phi/\psi = 115/156°$; involved in turn |
| G67 | 29 | Fold conservation; $\phi/\psi = 83/0°$; involved in turn |
| K69 | 30 | Conserved positive patch |
| K88 | 28 | Actin-binding site; conserved basic residue |
| T105 | 29 | Fold conservation; stabilized turn, hydrogen bonds to N100 |
| G121 | 31 | Actin binding; near actin C terminus in complex |
| Y133 | 29 | PLP binding; direct contact to PLP |
| L134 | 33 | PLP binding; direct contact to PLP |

*The frequency of occurrence of a particular residue in the 35 sequences examined. †Eight of these residues are involved in PLP binding; six appear to be important for fold conservation.

**Figure 5**



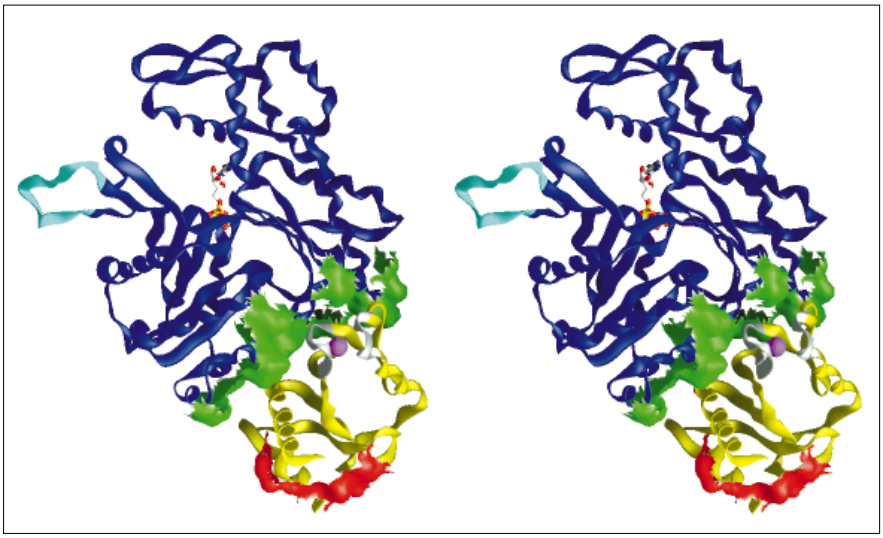

Stereo view of the plant-specific binding pocket. The backbone of *Arabidopsis* profilin is shown for residues Ile75 to Arg84 and Asp53 to Gly58. Sidechains are omitted for all residues except Asp53, Phe54, Ile83, and Pro57. Phe54 and Ile85 form a hydrophobic floor that the glycerol carbon backbone rests upon, while Pro57 and Gln79 close the top face of the pocket. Dashed lines represent hydrogen bonds. (Figure generated using the program MOLSCRIPT [69].)

with hydrophobic residues (Phe54 and Ile83) and the carboxylate of Asp52b (where b relates to an insertion relative to the bovine sequence), which is buried at the base of the pocket (Fig. 5). The alcohol groups are involved in a complex hydrogen-bonding network where two of the hydroxyls make hydrogen bonds to the protein and the third makes hydrogen bonds to three water molecules, enclosing a volume of 590 Å$^3$.

This pocket is not seen in either the *Acanthamoeba* or bovine profilin structures and represents a unique feature of the plant profilins. It is notable that it is positioned just adjacent to, indeed forms part of, the actin-binding surface (Fig. 6). In non-plant profilins, the polypeptide chain occupies the space filled by the glycerol in the *Arabidopsis* structure. Due to the possibility that the glycerol-binding site may have been an artifact of freezing, the structure was solved at room temperature to 2.2 Å (Table 2). The frozen and room temperature structures differed very little with a root mean square (rms) deviation of 0.295 Å over all Cα atoms, confirming that the pocket is an intrinsic feature of the plant profilin fold. At room temperature, four ordered water molecules are observed in the pocket (B factors = 6.35, 16.12, 16.03 and 9.01 Å$^2$ for waters 4, 16, 17 and 18, respectively). Two of the water molecules occupy the same space as the glycerol while two others are coincident with waters observed in the low temperature structure. The residues which make up the pocket are well conserved among plant profilins (Figs 2,5) indicating that the pocket is a structural signature of the plant profilins.

A number of other small molecules are also identifiable in the low temperature electron-density map. A second, less well ordered, glycerol (average B factor of 25 Å$^2$) and

**Figure 6**



Stereo view model of the *Arabidopsis* profilin–actin complex. The model was prepared by alignment with the bovine profilin–actin structure and shows the actin (dark blue) backbone with the DNAse-binding loop (light blue) and the ATP (in standard colors) of the actin. The profilin (yellow) backbone demonstrates the PLP-binding surface (in red) and the plant-specific binding pocket (white ribbon) as well as the position of the glycerol (purple sphere) relative to the profilin–actin interface (green). The plant-specific binding pocket is located immediately adjacent to the actin–profilin interface. (Figure generated using the program GRASP [72].)

**Table 2**

**Overall data collection and phasing statistics.**

| Derivative | λ (Å) | Resolution (Å) | Completeness (%) | $R_{sym}$* | Total reflections | Unique reflections | $R_{iso}$† | < FH >/E | Figure of merit‡ |
|---|---|---|---|---|---|---|---|---|---|
| MAD phases | | | | | | | | | |
| SeMet rising | 0.9791 | 30.0–1.60 | 96.0 | 0.055 | 122 584 | 27 092 | | | |
| SeMet peak | 0.9786 | 30.0–1.60 | 95.7 | 0.056 | 123 354 | 27 114 | 0.048 | 1.40 | |
| SeMet remote | 0.9500 | 30.0–1.50 | 95.8 | 0.062 | 146 788 | 32 355 | 0.080 | 1.77 | 0.5933 |
| SIRAS+MAD phases | | | | | | | | | |
| SeMet peak | 0.9786 | | | | | | | 1.62 | |
| SeMet remote | 0.9500 | | | | | | | 1.99 | |
| $CH_3HgCl$ | 1.0030 | 30.0–1.75 | 94.3 | 0.034 | 98 327 | 11 483 | 0.399 | 1.58 | 0.7195 |
| Room temperature | | | | | | | | | |
| data | 1.54 | 30.0–2.2 | 90.4 | 0.066 | 48 373 | 5545 | | | |

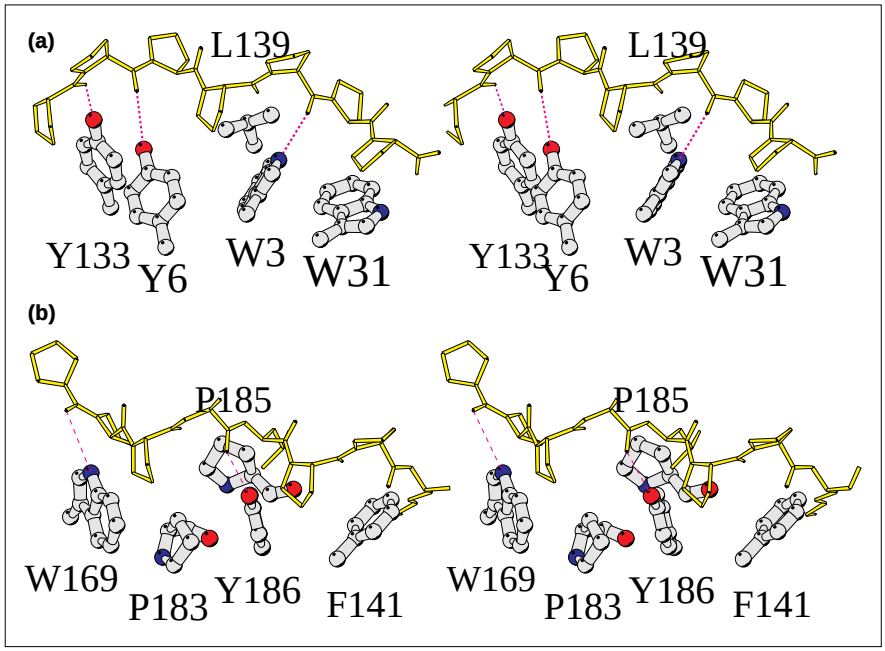
*$R_{sym} = (\Sigma_{hkl} | I_{hkl} - < I > |) / \Sigma_{hkl} I_{hkl}$, and †$R_{iso} = (\Sigma_{hkl} | F_{PH}^{hkl} - F_P^{hkl} |) / \Sigma_{hkl} F_P$, where the average $< I >$ is taken over all equivalent measurements, $I_{hkl}$, and $F_P$ and $F_{PH}$ refer to the native protein structure factor and the protein heavy-atom structure factor, respectively. ‡The figure of merit as defined by Blow and Crick [74] for both the selenomethionyl (SeMet) MAD data and methylmercury chloride ($CH_3HgCl$) SIR + SeMet MAD phasing as a function of resolution.

a sulfate ion are located between two symmetry-related molecules. The sulfate ion probably explains the failure to obtain crystals when other ammonium salts were used as precipitants. Of the 99 water molecules associated with each molecule in the crystal structure, all but ten are in the first solvation layer, making at least one hydrogen bond to the protein. In the room temperature structure, the sulfate ion is not identifiable and only 35 water molecules were placed with all of them directly contacting the protein.

### The PLP-binding site of profilins is the most highly conserved feature

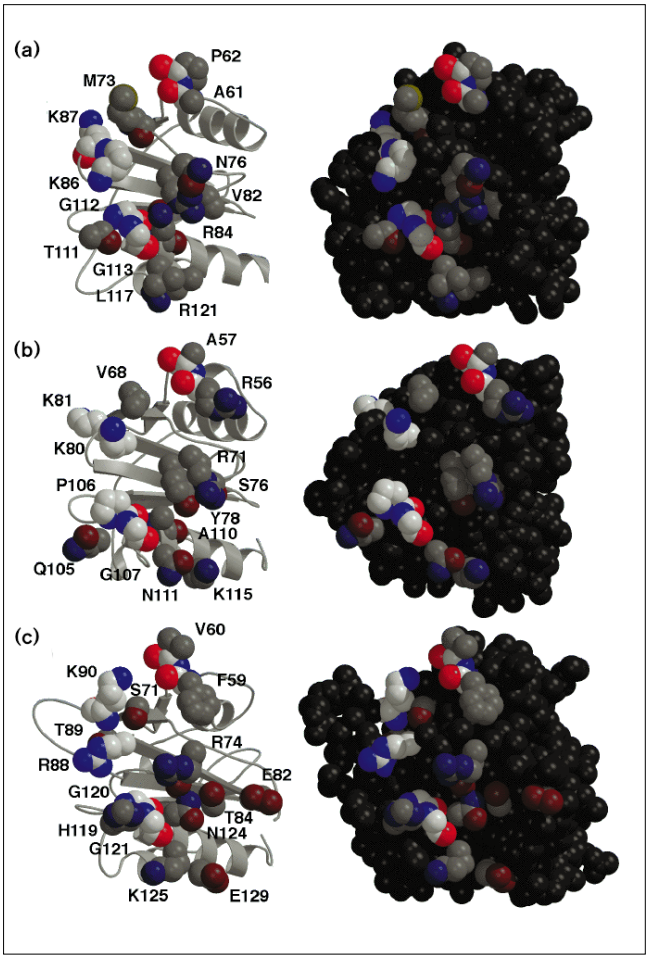
Of the 18 residues conserved amongst 80% of profilins, nearly half (Trp3, Tyr6, Ile21, Gly23, Trp31, Ala32, Tyr133 [His133 in bovine profilin] and Leu134) are found in a hydrophobic patch on the surface of profilin distinct from the actin-binding site [11]. This hydrophobic patch has been determined to be responsible for PLP-binding by mutagenesis [35,36], NMR spectroscopy [37,38], and fluorescence quenching [38,39]. Except for the putative

**Figure 7**




Stereo view showing the refined model of PLP binding to profilin. Polar contacts analogous to those found in SH3–PRP complexes are found in profilin–PLP models. **(a)** A proline helix (yellow stick model) makes hydrogen bonds (dotted lines) to the heteroatoms of the sidechains of Tyr6, Trp3, and Tyr133 of the 'aromatic ladder' on the profilin surface. **(b)** The c-Crk SH3 domain associates with PRP (yellow stick model) only through polar contacts between the proline carbonyls and sidechains which form a similar stack of sidechains for SH3 domains and profilin [43]. (Figure generated using the program MOLSCRIPT [69].)

**Figure 8**




The actin-binding face of profilins. The figure shows the residues corresponding to the actin-binding surface of (a) *Arabidopsis* profilin I, (b) *Acanthamoeba* profilin I and (c) bovine profilin. Residues contacting actin in the model are displayed on a ribbon representation (left) and on a dark gray CPK van der Waals surface (right). Six residues appear to be conserved actin-binding determinants: four residues in the backbone (residues Ala61, Pro62, Gly112, and Gly113 of *Arabidopsis*), and two conserved basic residues (Lys86 and Lys87 of *Arabidopsis*) which are shown in brightly colored CPK representation. Comparison of the dark CPK images (right) demonstrates how the loop insertions associated with the different profilin classes affect the shape of the actin-binding face. (Figure generated using the programs MOLSCRIPT [69] and RASTER 3D [73].)

**Table 3**

**Residues in *Arabidopsis* and *Acanthamoeba* profilins corresponding to residues in bovine profilin which are known to contact actin in the profilin–β-actin complex.**

| *Arabidopsis* | *Acanthamoeba* I | Bovine | Actin contact |
|---|---|---|---|
| A61 | R56 | F59 | H173, D286 |
| P62 | A57 | V60 | D286, V287, R290 |
| M73 | V68 | S71 | Y166, D286 |
| N76 | R71 | R74 | H371 |
| - | - | E82* | K113 |
| V82 | S76 | T84 | R372 |
| R84 | Y78 | D86 | F375, C terminus |
| K86 | K80 | R88 | E167, Y169 |
| K87 | K81 | T89/K90* | Y166/D286, D288 |
| T111 | Q105 | H119 | Y169, Y133, K373 |
| G112 | P106 | G120 | Y169, F375 |
| G113 | G107 | G121 | F375 |
| N116 | A110 | N124 | R372 |
| L117 | N111 | K125 | E364, E361 |
| R121 | K115 | E129 | E364 |

Of the 16 residues found in van der Waals contact in the bovine profilin–β actin complex, only four are found to be strongly structurally conserved (shown in bold). *Denotes a residue which appears on the loops which are characteristically longer only in mammalian profilins.

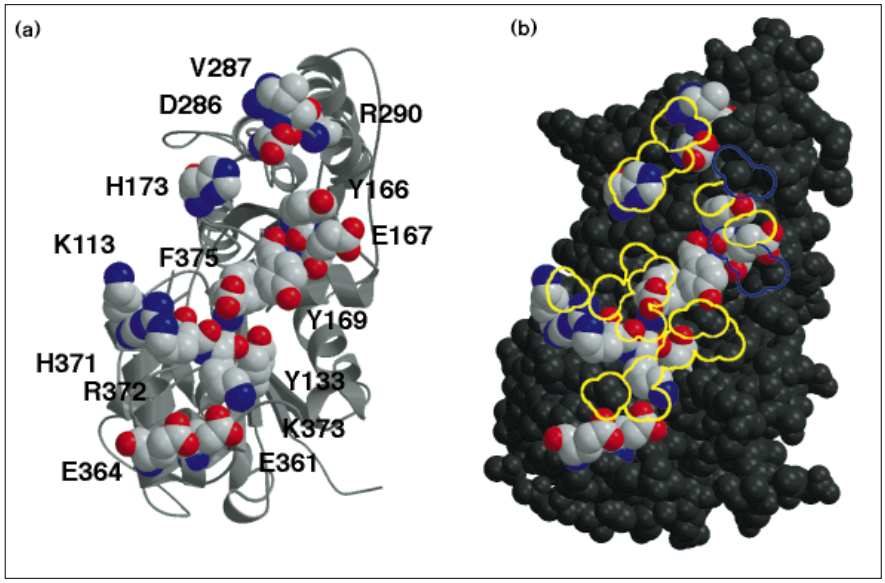
profilins from viruses [40], all profilins bind PLP. This function of profilin was initially discovered in surveys for PLP-binding proteins using PLP affinity chromatography [9,10,41] and has been extensively exploited in the purification of profilin and the profilin–β-actin complex [42].

The conservation of this hydrophobic patch reflects the conserved PLP-binding function of profilins. In all four profilin crystal structures, the sidechains of Trp3, Tyr6, and Trp31 are maintained in an 'aromatic ladder' conformation by interactions with conserved hydrophobic residues Tyr133 (histidine in bovine profilin) and Ile21, as well as the conserved hydrophobic residue 139 (leucine in *Arabidopsis*). The conservation of Ala32 and Gly23 appears to be required for the tight packing of their sidechains against Trp3 and Trp31 in the conserved ladder (Fig. 7).

The arrangement of solvent-exposed aromatic residues in the PLP-binding site on profilin is reminiscent of the binding site for proline-rich peptides (PRPs) on SH3 domains (Fig. 7b). The SH3 PRP-binding site contains a conserved hydrophobic ladder, consisting of tryptophan, proline, and tyrosine residues. SH3 domains are thought to localize proteins to the eukaryotic cytoskeleton via the binding of PRPs in the proline helix II configuration [43]. Sidechains in the conserved aromatic ladder of the PLP-binding site of profilin are spaced between 4 and 5 Å apart, an arrangement similar to the conserved hydrophobic ladder of SH3 domains. The sidechains of the three conserved residues in the PRP-binding site support two turns of a proline II helix (Fig. 7a).

Structures of SH3 domains with and without bound PRP [43–46] show that the PRP-binding site is pre-organized (i.e. the configuration of the sidechains does not change in bound structures). Although the PLP- and PRP-binding sites are composed of residues presented by different secondary structure elements in the profilin and SH3 domains, respectively, the near identity of the PLP-binding sites on the three profilin crystal structures [28,29] and their similarity to the PRP-binding sites of SH3 domains suggests that

**Figure 9**




The profilin-binding site of actin. (a) The ribbon depiction of actin displays the sidechains which contact profilin in the bovine profilin–actin complex. On a CPK representation of the actin molecule (b) these same contact residues are shown in standard colors with the outline of the residues of the bovine profilin-binding site shown in yellow (compare with Fig. 8c). (Figure generated using the program MOLSCRIPT [69].)

PLP binding may also occur without major changes in sidechain positions.

To investigate whether the mode of binding might be similar between SH3 domains and profilins, a PLP heptamer was modeled into the profilin PLP-binding site. Modeling was performed using the known structure of PRP co-crystallized with the N-terminal SH3 domain from c-Crk (PDB entry 1cka [43]). The model was initially fit by manual rigid-body rotation and translation onto the profilin-binding site and then refined by Newton-Raphson minimization [47]. In the resulting model, the ridge of the PLP helix ridge rests between the Trp3 and Tyr6 sidechains and proline carbonyls make hydrogen bonds with donors from Trp3, Tyr6 and Tyr133, a bonding pattern analogous to that of SH3 domains (Fig. 7a). Again in a manner similar to that found in all SH3 complexes (e.g. Phe141 of c-Crk; Fig. 6b), the indole ring of Trp31 lies flat along the protein surface and a turn of the PLP helix lies on top of it. The PLP helix is supported along its length by a conserved hydrophobic residue (Leu134) in a manner very similar to that found in SH3–PRP complexes (Fig. 7b).

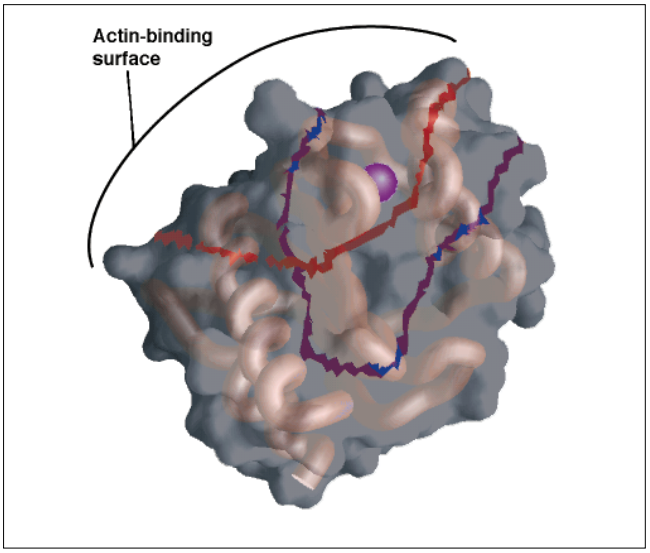
Because of equatorial pseudo-twofold symmetry in the PLP helix, equally favorable hydrogen-bonding interactions may be found in either orientation along the aromatic ladder [46]. The specific PRP sequence which a given SH3 domain binds is dictated by the arrangement of polar sidechains surrounding the PRP-binding site [43] which orients the PRP in the binding site [46]. The *in vivo* target of the profilin PLP-binding site has not been elucidated, although profilin affinity chromatography of platelet lysates yielded [48] vasodilator-stimulated phosphoprotein (VASP). VASP is involved in the regulation of both cytoskeletal dynamics and signal transduction. VASP contains a $(GlyPro_5)_3$ repeat [33] and $(GlyPro_5)_3$ peptides competitively dissociate VASP from profilin affinity columns [48].

### The actin-binding surfaces of profilins are not conserved

The highly conserved nature of actin [49] would seem to dictate that actin-binding proteins are equally well conserved at sites of contact with actin. This is not the case for the profilin family, as can be seen by comparing models of the *Arabidopsis* and *Acanthamoeba* profilin–actin interfaces (Fig. 8) to the crystal structure of the bovine profilin–β-actin complex [11] obtained by a least squares superposition based on the bovine profilin–β actin crystal structure.

The major actin-binding site of profilin consists of the face of β strands 4, 5, and 6 and the C-terminal helix [11]. The residues of the actin-binding site are poorly conserved (see boxed residues; Fig. 2); only Lys88 and Gly121 are highly conserved among all profilins (Table 3). Nevertheless, there are consistent determinants of the actin-binding surface of profilin: a basic patch comprising residues Lys88 and Lys89, a knob composed of residues Ala59 and Pro60 which contacts the actin backbone, and a further backbone contact with actin at Gly120 and the highly conserved Gly121. The loop between β strands 4 and 5 contains Glu82 which forms a contact with Lys113 of actin in the bovine profilin–actin complex. This residue is not found in the shorter loop located between these two strands in *Arabidopsis* profilin.

**Figure 10**




A structural model of the allergenic immune response with cross-reactivity centering on the binding pocket. The IgE response engendered by plant profilins with epitopes on the binding-pocket region will often overlap with the adjacent semi-conserved actin-binding site (light gray), resulting in cross-reactivity with human profilins. Individuals whose IgE epitopes only contact non-conserved regions (dark gray region) will not have an immune response maintained by their own native profilin. The red and blue boundaries enclose two hypothetical epitopes, where the red represents an auto-reactive IgE-binding region and the blue area a non-autoreactive site. (Figure generated using the program GRASP [72].)

The lack of strict conservation of the actin-binding surface in profilin is perhaps surprising, but not unprecedented. Several structures of distinct proteins binding to a common protein epitope indicate that a given protein region can be recognized by different protein surfaces. The crystal structures of antibodies NC10 and NC41 in complex with neuraminidase [50] show that 80% of the antibody-binding site on neuraminidase is common between the binding sites even though only one residue is shared by both $F_{ab}$ structures. Similarly, the structures of the antilysozyme idiotope–anti-idiotope complex [51] and the human growth hormone bound to its receptor [52], show that a given protein region can be recognized by different protein surfaces. Examination of the profilin-binding site on actin reveals that a basic patch on profilin (Lys90 and Arg88) aligns with an acidic patch (Glu167 and Phe375 at the C terminus) on the surface of actin (Fig. 9).

In the model of the *Arabidopsis* profilin–actin complex (Fig. 6), it is notable that one wall of the plant-specific binding pocket forms part of the actin-binding site. Presumably any effector which binds to the pocket would influence the energetics of actin–profilin complex formation. This model predicts that a specialized regulator of actin polymerization exists in plants. The search for such an effector is underway.

The rate of ATP exchange by rabbit muscle actin is decreased by up to 250-fold by *Arabidopsis* profilin association [23,24], although it remains to be seen if *Arabidopsis* profilin–actin exhibits a similar effect. The structure of an open-state of the bovine profilin–actin complex [53] shows that concerted subdomain motions in the actin monomer link changes at the C terminus of actin and the DNAse-binding loop to movements of the nucleotide in the cleft between the two major domains of actin (Fig. 6). Residue His119 of bovine profilin (threonine in *Arabidopsis*) makes a focal contact near the C terminus of actin so that substitutions at this site may account for the decreased rate of nucleotide exchange.

In summary, the protein fold is conserved within the profilin families; however, there are four unique features of plant profilins: three interstrand loops and a putative binding pocket. Two of these loops (between β strands 4 and 5 and between β strands 5 and 6) are part of the actin-binding interface. The binding pocket, located adjacent to the actin-binding interface, is composed of residues that are conserved in the known plant profilins. All of these features may affect interaction with actin and the kinetics of actin ATP hydrolysis.

Among the 35 known profilins, 18 residues are greater than 80% conserved; eight of these residues form the PLP-binding site. Comparison of SH3 domains with profilins suggests a similar mode of PRP helix association in the two distinct protein families. This appears to be an example of convergent molecular evolution, as there is no apparent relationship between the SH3 and profilin families except for their involvement in intracellular signaling. Another PRP-binding protein family, the WW domain, characterized by two stacked tryptophan residues, has been reported [54]. The structure [54] of the complex of the WW domain of the proto-oncogene YES associated protein with its bound peptide [55] is another example of the kind of interaction predicted here for profilins. The consensus PRP sequence for WW domains is Pro-X-X-Pro-Tyr where the phenol of the tyrosine stacks to complete a ladder of tyrosine, tryptophan, and proline sidechains which occupy the grooves of the WW PRP-binding site.

In contrast, the actin-binding site contains only two of the residues conserved among the profilins. In spite of this variation of primary sequence, the general fold and distribution of charges within the actin-binding site is similar to other profilins.

### A structural model for plant allergies

When pollen is hydrated, for example when it contacts a human mucus membrane following inhalation, pollen tube formation is initiated and large amounts of profilin are released. Valenta *et al.* [27] have demonstrated that IgEs isolated from allergic patients react with profilins from a wide

variety of plants as well as weakly with human profilin. This weak binding to human profilin is thought to sustain allergenic hypersensitivity to a broad spectrum of plant pollens even after the cessation of the hay fever season [25,27]. The identification of a plant-specific binding pocket suggests that it is the major immunogenic region of plant profilins. It is unlikely that the PRP-binding site is the source of auto-reactivity as it is so highly conserved. The IgE epitope probably encompasses the adjacent actin-binding site, generating a cross-reacting response to human as well as plant profilins (Fig. 10), as the actin site is only semi-conserved. Individuals whose IgE epitopes include the pocket but not the actin-binding site would not exhibit allergenic responses, while those with cross-reacting IgEs maintain appreciable levels of self and plant profilin specific IgEs.

## Biological implications

Profilin binds to actin and regulates its assembly into microfilaments. It is also involved in signal transduction pathways mediated by binding poly-L-proline (PLP) sequences and phosphatidylinositol-4,5-bisphosphate. Plant profilins represent a conserved family of proteins distinct from those of mammals and other eukaryotes, but until now a structure of a plant profilin has not been reported.

The structure of *Arabidopsis* profilin I allows a structural comparison between the major families of profilins. Such a comparison reveals that nearly half of the residues which are known to be conserved across all profilins, are involved in the interaction of profilin with PLP. Many of the remaining residues appear to be important for maintaining the tertiary structure of profilin; they are involved in turns or in close packing of core residues. Surprisingly, only two conserved residues are involved in actin binding. Modeling of the actin-binding face of different classes of profilin reveals binding patterns in which residues make similar, but not identical, contacts with their respective actin molecules. The differences, such as His119→Thr119, may have significance for understanding how profilin can have different effects on cytoskeletal function in various genetic isoforms and different organisms.

Plant profilins have been the subject of great interest after their discovery as major human allergens [25]. The plant-specific binding site, as the only novel plant-specific structural feature, represents the most likely immunogenic site of plant profilins. The close proximity of the plant-specific and actin-binding sites suggests that IgEs whose epitopes include both the plant-specific site and portions of the actin-binding site are responsible for pollen allergies.

## Materials and methods

### *Overexpression and crystallization*

*Escherichia coli* transformed with a DE3 plasmid containing the *Arabidopsis* profilin I gene were used to produce profilin essentially as previously described [56] except that 35% dimethyl sulfoxide (DMSO) was used to elute the profilin from the PLP affinity column, as described in [42]. DMSO was removed by overnight dialysis against 20 mM Tris/HCl, 30 mM sodium pyrophosphate, 1 mM EDTA, pH 7.5 (TPPE).

The protein was then concentrated by ultrafiltration to 10 mg $ml^{-1}$ and further purified by anion exchange chromatography on a Mono-Q FPLC column (Pharmacia). The column was equilibrated with 20 mM Tris/HCl, pH 7.5, and the profilin was applied in TPPE buffer. Profilin was eluted using a linear gradient of 0 to 0.5 M NaCl, 20 mM Tris/HCl, pH 7.5; profilin eluted at 200 mM NaCl. The profilin was desalted into 20 mM Tris/HCl pH 7.5 on a PD-10 column (Pharmacia), concentrated by ultrafiltration to 7 mg $ml^{-1}$, centrifuged and used for crystallization. Protein purity was monitored by gel electrophoresis. Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry was used to further verify the quality of the protein and determined that Met1 was cleaved from the recombinant protein [57].

Selenomethionyl substituted profilin for MAD experiments was prepared by producing protein in a culture via the methionine biosynthesis pathway inhibition procedure [58]. The resulting protein was purified by the same methods as the native profilin except that all buffers contained 10 mM dithiothreitol (DTT) and 0.2 mM EDTA to prevent oxidation of the selenomethionine. DTT was not included during the DMSO elution step to prevent formation of dimethyl sulfide.

Crystallization conditions were evaluated with sparse matrix screens from Hampton Research [59] using the hanging drop method. The optimum crystallization conditions for the selenomethionyl protein were found to be 2.0 M ammonium sulfate, 0.1 M citrate pH 5.0, 10 mM DTT, 0.2 mM EDTA. The crystallization of selenomethionyl protein was carried out in the absence of oxygen. These crystals are isomorphous with the native crystals which crystallize under identical conditions and belong to space group $P2_12_12_1$, with a = 42.04, b = 43.07, and c = 59.45 Å.

### *Data collection*

Multiwavelength anomalous diffraction (MAD) data were collected at beamline X12-C of the National Synchrotron Light Source at Brookhaven National Laboratory using a CAD4 goniometer and a 300 mm MAR image plate controlled by the MARMAD software package [60]. Data sets were collected on a single crystal frozen in synthetic mother liquor with 25% glycerol added [61]. The crystal was transferred to freezing buffer by increasing the glycerol concentration by 5% every 15 min, and the crystal was then flash-frozen in a 100 K nitrogen gas stream. The crystal was aligned so that the c axis of the crystal was along the rotation axis of the goniometer, allowing near simultaneous collection of Freidel pairs. Native data were collected on an RU-200 rotating-anode X-ray generator with an R-AXIS II detector.

MAD data sets were collected at the inflection point of the rising edge of the absorption spectrum, at the peak of the absorption spectrum, and at a high energy remote wavelength (0.95 Å) with recalibrations performed at every wavelength change. 90° data sweeps were collected in blocks of 16 degrees followed by a 20° cusp sweep. Each block was collected sequentially at all three wavelengths. Integration and scaling were carried out with DENZO and SCALEPACK [62]; scaled intensities were converted to structure factors using TRUNCATE [63].

A derivative data set was collected using the same protocols on a crystal soaked for 16 h in methylmercury chloride. Collection was done at the peak absorption wavelength (1.003 Å) so as to maximize the anomalous difference. The statistics for these data sets are given in Table 4.

The room temperature data set was collected on a Rigaku RU-200 Cu rotating anode with graphite mirrors running at 50 kV and 100 mA. The detector was a RAXIS II-C image plate system, scanning at high resolution. Statistics for this data set are in Table 2.

**Table 4**

**Refinement statistics.**

| | Low temperature model | Room temperature model |
|---|---|---|
| Reflections used | 14 596 | 5530 |
| Atoms refined | 1157 | 1280 |
| Resolution (Å) | 1.6 | 2.2 |
| R (%) | 18.08 | 17.20 |
| $R_{free}$ (%)* | 20.49 | 23.80 |
| Completeness (%) | 92.94 | 84.33 |
| Rms deviation | | |
| bonds (Å) | 0.009 | 0.014 |
| angles (°) | 1.563 | 1.907 |
| Average B factor (Å$^2$) | | |
| mainchain atoms | 8.6 | 8.6 |
| sidechain atoms | 11.4 | 11.4 |
| solvent | 20.3 | 20.5 |

All data from 30–1.6 Å were used in the refinement. The selenomethionyl remote data set was used for refinement of the model without anomalous differences. *$R_{free}$ is the same as the standard crystallographic R factor except computed for 10% of data left out of the refinement.

### *Molecular replacement*

Initial phasing of the native data was attempted using molecular replacement as implemented in AMoRe [64]. Both the *Acanthamoeba* I (PDB entry 1acf) [29] and bovine (PDB entry 1pne) [28] profilin structures were used as search models. The structure of *Arabidopsis* profilin is expected to be similar to that of other profilins on the basis of sequence conservation (39% identity between *Arabidopsis* and *Acanthamoeba* profilins), the conservation of function between profilins, and the conservation of the fold between those profilin structures already determined. However, no clear molecular replacement solution could be found.

The failure of molecular replacement was investigated subsequent to the determination of the *Arabidopsis* structure reported here, using the refined model in molecular replacement searches. The refined model without water gave a very clear molecular replacement solution, even when sidechains were removed and B factors were uniformly set to 15. This implies that the inability of molecular replacement to yield a clear solution must be due to the observed differences in structure.

The low temperature MAD coordinates were used as a starting model for fitting the low temperature data. The amino acid residues surrounding the plant-specific binding pocket were excised (residues 52 to 60 and 74 to 84) and the data was initially fit by rigid-body rotation of the model and 150 cycles of positional refinement; the R factor dropped to 31.8%.

### *Phasing*

A conventional search for heavy-atom derivatives yielded only one isomorphous derivative: methylmercury chloride. The major mercury site was located by anomalous and isomorphous difference Patterson methods. However, because of low occupancy of the mercury sites and substantial non-isomorphism with the native data, SIRAS phases were not sufficient to determine the structure.

To obtain phases, selenomethionyl profilin was prepared and MAD data were collected as described above. The three expected selenium positions were found via cross-difference Fourier maps in both hands calculated with the anomalous differences from the rising edge MAD data and SIRAS phases from the major mercury site. The two sets of potential sites from both enantiomers were refined in MLPHARE [63] using the rising edge data as native and the peak and remote wavelengths as derivatives with anomalous scattering [65] with values of f′ and f″ taken from [65]. The phasing power and the overall anomalous statistics of one hand refined very well (Table 3) and yielded an interpretable density map.

A second set of phases for the selenomethionine crystals were calculated by including the mercury derivative data (with anomalous scattering) as a third derivative to the selenomethionyl rising edge data. Two mercury sites (located by Patterson methods) were included in the model and the selenium sites were refined with negative occupancy. This combination of SIRAS and MAD phasing (SIRMAD) improved the figure of merit (Table 4) and yielded somewhat better maps despite the severe non-isomorphism ($R_{iso}$ = 39.9%) between this derivative and the selenomethionyl data.

### *Model building*

The 2.2 Å resolution MAD and SIRMAD electron-density maps were quite clear, permitting the unambiguous determination of the correct sidechain in most places. The *Acanthamoeba* profilin I backbone was placed piece-wise in the electron density and the correct sidechains based on the sequence were built using the program O [66].

Conjugate gradient positional refinement of this initial model (residues 2–74 and 81–131) in X-PLOR [67] using data from 6.0 to 2.2 Å gave an R factor of 31.01% ($R_{free}$ = 38.38%; with 6% of reflections in the test set [68]). After a few rounds of refinement, water molecules were placed at peaks higher than 4$\sigma$ in the $F_o$–$F_c$ map if there was corresponding density in the SIRMAD or MAD maps and the water made a hydrogen bond to at least one protein atom. Phase extension was performed to 1.6 Å resulting in an R factor of 19.23% ($R_{free}$ = 22.02%) for the complete model. A bulk solvent mask using all data from 30–1.6 Å with parameters K = 0.45, B = 50, radius = 0.29 Å improved this to 17.95% ($R_{free}$ = 20.95%).

Alternate conformers for several residues could be seen in both the $F_o$–$F_c$ and 2$F_o$–$F_c$ maps, and were built for Asp9, Cys13, Asp14, Ser31, Gln35, Ile47, Glu78, Pro89, Gly90, Lys96, and Met110. The current model contains residues 2–131 of profilin (the N-terminal methionine residue was removed by the *E. coli* host), 99 water molecules, two glycerols, and one sulfate with R = 0.1806 ($R_{free}$ = 0.2052) and good stereochemistry.

The room temperature structure including the plant-specific binding pocket was rebuilt manually and refined using X-PLOR Powell refinement mode. 26 water molecules were then identified using $F_o$–$F_c$ maps and placed manually using the program O and the structure was completed with positional and annealing refinement.

### *Construction of profilin–actin interfaces*

The LSQ facilities in the program O [66] were used to align the alpha carbons of *Arabidopsis* profilin I and *Acanthamoeba* profilin I by least squares algorithm to residues in the actin-binding face of bovine profilin from the profilin–actin complex (PDB entry pdb2btf.ent). A list of the residues which are in contact in the bovine profilin–actin complex were generated with the program CONTACT from the CCP4 suite [63]. The alignment was refined using profilin residues structurally homologous to those which contact actin residues found in CONTACT (bovine profilin residues Phe59, Ser60, Ser71, Arg74, Glu82, Thr84, Asp86, Arg88, Thr89, His119, Gly120, Gly121, Asn124, Lys125, and Glu129).

### *Refinement of models of PLP binding to profilin*

A model type II PLP helix complex was prepared by aligning bound PRPs of several SH3 domain–PRP complexes (PDB entries 1abl, 1abo, 1cka, 1ckb, 1lck, 1prm, 1sem, 1shf, 1shg) using the LSQ facility in O [66]. The aromatic ladder of *Arabidopsis* profilin I was then aligned to the consensus binding site and the peptide from the poly-proline-binding structure (lck) was taken as an initial model. This initial model was mutated into a PLP sequence using the MUTATE facilities in O and manually rotated to maximize the number of hydrogen-bonding contacts between the proline carbonyls and profilin sidechain heteroatom donors. This PLP model was

then refined *in vacuo* with 1200 iterations of Newton-Raphson minimization [47] in the DISCOVER molecular modeling module in the InsightII suite, keeping the coordinates of the profilin fixed.

### *Accession numbers*

The refined coordinates for both models have been deposited in the Brookhaven Protein Data Bank (*Arabidopsis* PDB code 3NUL).

## Acknowledgements

We acknowledge the contributions of Jennifer Jones, Dale Jefferson, Brian Johnson, and Rebecca Schutt who obtained the original crystals of plant profilin, and Michael Rozycki for advice.

We also thank Dr Robert Sweet and the staff at X12C where diffraction data for this study were collected. Beamline X12C at the National Synchrotron Light Source is supported by the United States Department of Energy, Office of Health and Environmental Research, and by the National Science Foundation.

This work was supported by NIH grant GM-44038, by the National Science and Technology Board, Singapore, and by a grant-in-aid from DuPont Agricultural Products.

## References

1. Small, J. (1994). *Seminars in Cell Biology* 5. Academic Press Ltd.
2. Small, J.V., Isenberg, G. & Celis, J.E. (1978). Polarity of actin at the leading edge of cultured cells. *Nature* 272, 638–639.
3. Hoglund, A.-S., Karlsson, R., Arro, E., Fredriksson, B.-A. & Lindberg, U. (1980). Visualisation of the peripheral weave of microfilaments in glial cells. *J. Muscle Res. Cell Motil.* 1, 127–146.
4. Staiger, C., Yuan, M., Valenta, R., Shaw, P.J., Warn, R.M. & Lloyd, C.W. (1994). Microinjected profilin affects cytoplasmic streaming in plant cells by rapidly depolymerizing actin filaments. *Curr. Biol.* 4, 215–219.
5. Bedinger, P.A., Hardeman, K.J. & Loukides, C.A. (1994). Travelling in style: the cell biology of pollen. *Trends Cell Biol.* 4, 132–138.
6. Carlsson, L., Nystrom, L.-E., Sundkvist, I., Markey, F. & Lindberg, U. (1977). Actin polymerizability is influenced by profilin, a low molecular weight protein in non-muscle cells. *J. Mol. Biol.* 115, 465–483.
7. Machesky, L.M. & Pollard, T.D. (1993). Profilin as a potential mediator of membrane-cytoskeleton communication. *Trends Cell Biol.* 3, 381–385.
8. Pantaloni, D. & Carlier, M.-F. (1993). How profilin promotes actin filament assembly in the presence of thymosin β4. *Cell* 75, 1007–1014.
9. Tanaka, M. & Shibata, H. (1985). Poly(L-proline)-binding proteins from chick embryos are a profilin and a profilactin. *Eur. J. Biochem.* 151, 291–297.
10. Lindberg, U., Schutt, C.E., Hellsten, E., Tjäder, A.-C., Hult, T. (1988). The use of poly(L-proline) sepharose in the isolation of profilin and profilactin complexes. *Biochim. Biophys. Acta* 967, 391–400.
11. Schutt, C.E., Myslik, J.C., Rozycki, M.D., Goonesekere, N.C.W. & Lindberg, U. (1993). The structure of crystalline profilin-actin. *Nature* 365, 810–816.
12. Berridge, M.J. (1993). Inositol triphosphate and calcium signalling. *Nature* 361, 315–325.
13. Lassing, I. & Lindberg, U. (1985). Specific interaction between phosphatidyl 4,5 bisphosphate and profilactin. *Nature* 314, 472–474.
14. Lassing, I. & Lindberg, U. (1988). Specificity of the interaction between phosphatidyl-inositol 4,5 bisphosphate and the profilin:actin complex. *J. Cell. Biochem.* 37, 255–267.
15. Chinkers, M., McKanna, J.A. & Cohen, S. (1979). Rapid induction of morphological changes in human carcinoma cells A431 by epidermal growth factor. *J. Cell Biol.* 83, 260–265.
16. Mellstrom, K., Hoglund, A.-S., Nister, M., Heldin, C.-H., Westermark, B. & Lindberg, U. (1983). The effect of platelet-derived growth factor on morphology and motility of human glial cells. *J. Muscle Res. Cell Motil.* 4, 589–609.
17. Hall, A. (1994). Small GTP-binding proteins and the regulation of the actin cytoskeleton. *Ann. Rev. Cell Biol.* 10, 31–54.
18. Bar-Sagi, D., Rotin, D., Batzer, A., Mandiyan, V., Schlessinger, J. (1993). SH3 domains direct cellular localization of signalling molecules. *Cell* 74, 83–91.
19. Pawson, T. & Schlessinger, J. (1994). SH2 and SH3 domains. *Curr. Biol.* 3, 434–442.
20. Giehl, K., Valenta, R., Rothkegel, M., Ronsiek, M. & Mannherz, H.G. (1994). Interaction of plant profilin with mammalian actin. *Eur. J. Biochem.* 226, 681–689.
21. Perelroizen, I., Didry, D., Christensen, H., Chua, N.-H. & Carlier, M.-F. (1996). Role of nucleotide exchange and hydrolysis in the function of profilin in actin assembly. *J. Biol. Chem.* 271, 12302–12309.
22. Perelroizen, I., Carlier, M.-F. & Pantaloni, D. (1995). Binding of divalent cation and nucleotide to G-actin in the presence of profilin. *J. Biol. Chem.* 270, 1501–1508.
23. Nishida, E. (1985). Opposite effects of cofilin and profilin from porcine brain on rate of exchange of actin-bound adenosine 5′-triphosphate. *Biochemistry* 24, 1160–1164.
24. Mockrin, S.C. & Korn, E.D. (1980). *Acanthamoeba* profilin interacts with G-actin to increase the rate of exchange of actin-bound adenosine 5′-triphosphate. *Biochemistry* 19, 5359–5362.
25. Valenta, R., *et al.*, & Scheiner, O. (1991). Identification of profilin as a novel pollen allergen; IgE autoreactivity in sensitized individuals. *Science* 253, 557–560.
26. Valenta, R., *et al.*, & Scheiner, O. (1992). Profilins constitute a novel family of functional plant pan-allergens. *J. Exp. Med.* 175, 377–385.
27. Valenta, R., *et al.*, & Scheiner, O. (1992). Profilin, a novel plant pan-allergen. *Int. Arch. Allergy Immunol.* 99, 271–273.
28. Cedergren-Zeppezauer, E.S., *et al.*, & Schutt, C. E. (1994). Crystallization and structure determination of bovine profilin at 2.0 Å. *J. Mol. Biol.* 240, 459–475.
29. Federov, A.A., Magnus, K.A., Graupe, M.H., Lattman, E.E., Pollard, T.D. & Almo, S.C. (1994). X-ray structures of isoforms of the actin-binding protein profilin that differ in their affinity for phosphatidylinositol phosphates. *Proc. Natl. Acad. Sci. USA* 91, 8636–8640.
30. Metzler, W.J., *et al.*, & Mueller, L. (1993). Characterization of the three-dimensional solution structure of human profilin: $^{1}$H, $^{13}$C, and $^{15}$N NMR assignments and global folding pattern. *Biochemistry* 32, 13818–13829.
31. Vinson, V.K., Archer, S.J., Lattman, E.E., Pollard, T.D. & Torchia, D.A. (1993). Three-dimensional solution structure of *Acanthamoeba* profilin-I. *J. Cell Biol.* 122, 1277–1283.
32. Heinz, D.W., Baase, W.A., Dalhquist, F.W., Matthews, B.W. (1993). How amino-acid insertions are allowed in an α helix of T4 lysozyme. *Nature* 361, 561–564.
33. Haffner, C., Jarchau, T., Reinhard, M., Hoppe, J., Lohmann, S.M. & Walter, U. (1995). Molecular cloning, structural analysis and functional expression of the proline-rich focal adhesion and microfilament-associated protein VASP. *EMBO J.* 14, 19–27.
34. Sohn, R.H., Chen, J., Koblan, K.S., Bray, P.F. & Goldschmidt-Clermont, P.J. (1995). Localization of a binding site for phosphatidylinositol 4,5-bisphosphate on human profilin. *J. Biol. Chem.* 270, 21114–21120.
35. Bjorkegren, C., Rozycki, M., Schutt, C.E., Lindberg, U. & Karlsson, R. (1993). Mutagenesis of human profilin locates its poly(L-proline)-binding site to a hydrophobic patch of aromatic amino acids. *FEBS Lett.* 333, 123–126.
36. Haarer, B.K., Petzold, A.S. & Brown, S.S. (1993). Mutational analysis of yeast actin. *Mol. Cell. Biol.* 13, 7864–7873.
37. Archer, S.J., Vinson, V.K., Pollard, T.D. & Torchia, D.A. (1994). Elucidation of the poly-L-proline binding site in *Acanthamoeba* profilin I by NMR spectroscopy. *FEBS Lett.* 337, 145–151.
38. Metzler, W.J., Bell, A.J., Ernst, E., Lavoie, T.B. & Mueller, L. (1994). Identification of the poly-L-proline-binding site on human profilin. *J. Biol. Chem.* 269, 4620–4625.
39. Perelroizen, I., Marchand, J.-B., Blancchoin, L., Didry, D. & Carlier, M.-F. (1994). Interaction of profilin with G-actin and poly(L-proline). *Biochemistry* 33, 8472–8478.
40. Machesky, L.M., Cole, N.B., Moss, B. & Pollard, T.D. (1994). Vaccinia virus expresses a novel profilin with a higher activity for polyphosphoinositides than actin. *Biochemistry* 33, 10815–10824.
41. Kaiser, D.A., Goldschmidt-Clermont, P.J., Levine, B.A. & Pollard, T.D. (1989). Characterization of renatured profilin purified by urea elution from poly-L-proline agarose columns. *Cell Motil. Cytoskeleton* 14, 251–262.
42. Rozycki, M.D., Schutt, C.E. & Lindberg, U. (1991). Affinity chromatography based purification of profilin-actin. *Methods Enzymol.* 196, 100–118.
43. Wu, X., *et al.*, & Kuriyan, J. (1995). Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure* 3, 215–226.
44. Noble, M.E.M., Musaccino, A.M., Saraste, M., Courtneidge, S.A. & Wierenga, R.K.W. (1993). Crystal structure of the SH3 domain in human FYN; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. *EMBO J.* 12, 2617–2624.
45. Eck, M.J., Atwell, S.K., Shoelson, S.E., Harrison, S.C. (1994). Structure of the regulatory domains of the src-family tyrosine kinase LCK. *Nature* 368, 764–769.

46. Lim, W.A., Richards, F.M. & Fox, R.O. (1994). Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* 372, 375–379.
47. Ralston, A. (1965). *A First Course in Numerical Analysis*. pp. 329–333, McGraw-Hill, New York.
48. Reinhard, M., *et al*., & Walter, U. (1995). The proline-rich focal adhesion and microfilament protein VASP is a ligand for profilins. *EMBO J.* 14, 1583–1589.
49. Shelterline, P. & Sparrow, J.C. (1994). Actin. In *Protein Profile*. Appendix, Acad. Press.
50. Malby, R.L., *et al*., & Colman, P.M. (1994). The structure of a complex between the NC10 antibody and influenza virus neuraminidase and comparison with the overlapping binding site of the NC41 antibody. *Structure* 2, 733–746.
51. Bently, G.A., Boulot, G., Riottot, M.M. & Poljak, R.J. (1990). Three-dimensional structure of an idiotope-anti-idiotope complex. *Nature* 348, 254–257.
52. de Vos, A., Ultsch, M. & Kossiakoff, A.A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* 255, 306–312.
53. Chik, J.K., Lindberg, U. & Schutt, C.E. (1996). The structure of an open state of β-actin at 2.65 Å resolution. *J. Mol. Biol.* 263, 607–623.
54. Sudol, M. (1996). The WW module competes with the SH3 domain? *Trends Biol. Sci.* 21, 161–163.
55. Marcias, M.J., *et al*., & Oschkinat, H. (1996). Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* 382, 646–649.
56. Christensen, H.E.M., Ramachandran, S., Tan, C.-T., Surana, U., Dong, C.-H. & Chua, N.-H. (1996). *Arabidopsis* profilins are functionally similar to yeast profilins: identification of a vascular bundle specific profilin and a pollen specific profilin. *Plant J*. 10, 269–279.
57. Chait, B.T. (1994). Mass spectrometry — a useful tool for the protein X-ray crystallographer and NMR spectroscopist. *Structure* 2, 465–467.
58. Doublie, S. (1996). Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol*., in press.
59. Jancarik, J. & Kim, S.-H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* 24, 409–411.
60. Sweet, R.M. & Skinner, J. (1995). An integrated graphical user interface for synchrotron beamline control and data collection. Montreal.
61. Rodgers, D.W. (1994). Cryocrystallography. *Structure* 2, 1135–1140.
62. Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*. (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
63. Collaborative Computational Project No. 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* 50, 760–763.
64. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. A* 50, 157–167.
65. Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. & Sweet, R.M. (1993). Crystal structure of the globular domain of histone H5 and its implications for nucleosome binding. *Nature* 362, 219–223.
66. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* 47, 110–119.
67. Brünger, A.T. (1993). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT, USA.
68. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355, 472–475.
69. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24, 946–950.
70. Barton, G. (1990). Protein multiple sequence alignment and flexible pattern matching. *Methods Enzymol.* 183, 403–428.
71. Barton, G.J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* 6, 37–40.
72. Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11, 281–296.
73. Merritt, E.A. & Murphy, M. (1994). Raster3D Version 2.0: a program for photorealistic molecular graphics. *Acta Cryst. D* 50, 869–873.
74. Blow, D.M. & Crick, F.H.C. (1959). The treatment of errors in the isomorphous replacement method. *Acta Cryst.* 12, 794–802.