

# Structure comparison of human glioma pathogenesis-related protein GliPR and the plant pathogenesis-related protein P14a indicates a functional link between the human immune system and a plant defense system

T. SZYPERSKI, C. FERNÁNDEZ, C. MUMENTHALER, AND K. WÜTHRICH\*

Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zurich, Switzerland

Contributed by K. Wüthrich, December 19, 1997

**ABSTRACT** The human glioma pathogenesis-related protein (GliPR) is highly expressed in the brain tumor glioblastoma multiforme and exhibits 35% amino acid sequence identity with the tomato pathogenesis-related (PR) protein P14a, which has an important role for the plant defense system. A molecular model of GliPR was computed with the distance geometry program DIANA on the basis of a P14a–GliPR sequence alignment and a set of 1,200 experimental NMR conformational constraints collected with P14a. The GliPR structure is represented by a group of 20 conformers with small residual DIANA target function values, low AMBER-energies after restrained energy-minimization with the program OPAL, and an average rms deviation relative to the mean of 1.6 Å for the backbone heavy atoms. Comparison of the GliPR model with the P14a structure lead to the identification of a common partially solvent-exposed spatial cluster of four amino acid residues, His-69, Glu-88, Glu-110, and His-127 in the GliPR numeration. This cluster is conserved in all known plant PR proteins of class 1, indicating a common putative active site for GliPR and PR-1 proteins and thus a functional link between the human immune system and a plant defense system.

The glioma pathogenesis-related protein (GliPR) is highly expressed in the tumor glioblastoma multiforme (1), which arises from brain immune cells and accounts for over 65% of all human primary brain tumors (2). GliPR was found in all glioma cell lines and tumors studied, but was not detectable in any normal fetal or adult tissues, including normal brain, suggesting that GliPR plays an important role for tumor growth. High levels of GliPR expression can also be induced with phorbol ester in macrophages (1), which are active at the front line of the human immune system. RTVP-1, another protein that was recently found in glioblastoma multiforme (3), is almost identical to GliPR. Compared with GliPR, RTVP-1 possesses both an additional N-terminal signal sequence and a C-terminal putative transmembrane segment. Although RTVP-1 was found to be expressed also in various normal tissues, it cannot be excluded that its high expression in the tumor cells is also related to their malignant properties. GliPR and RTVP-1 exhibit high sequence homology with the plant pathogenesis-related proteins of group 1 (PR-1 proteins), which play a central role for the defense system of plants (4), for example, during the manifestation of systemic acquired resistance (5). Because proteins with a sequence identity larger than 30% are commonly believed to adopt the same fold (6), this homology is suggestive of a structural link between the human immune system and the defense system of plants (1). To

follow up on this suggestion and in view of the possible role of GliPR as a potential target for drugs interfering with tumor growth, we further investigated structural similarities between GliPR and the PR-1 proteins on the level of the three-dimensional structure.

## METHODS

The three-dimensional structure of GliPR was predicted on the basis of the high-quality NMR structure of P14a (7). Therefore, a subset of the input used for the P14a structure calculation (Table 1) was selected as follows: First, we identified those upper distance constraints that connect corresponding proton pairs in P14a and GliPR. For conserved residues, stereospecific assignments of constraints to isopropyl methyls (Val-38, Leu-54, Val-131, Val-132 in GliPR) and  $\beta$ -methylene protons (in total 18) were retained. Second, the constraints for the backbone dihedral angles  $\phi$  and  $\psi$  measured for P14a were retained, except where L-amino acid residues in P14a are either deleted or replaced by Pro or Gly in GliPR. Third, the constraints for the dihedral angles  $\chi^1$  obtained for P14a were kept for sites with identical amino acids in both proteins. Fourth, we introduced constraints for eight central hydrogen bonds (8) that were identified in the four-stranded  $\beta$ -sheet of P14a. The resulting data set, which consists of 56% of the constraints obtained with P14a (Table 1), was used as input for structure calculations of GliPR by using the program DIANA (9). Subsequent restrained energy-minimization in a water bath with the AMBER force field (10) was performed with the program OPAL (11). The same protocol was followed as described in ref. 7 for P14a. Clefts on the protein surface and their volumes were identified by using the program SURFNET (12).

Amino acid sequence alignments were performed with the program CLUSTALW (13). To identify sequence homologues of GliPR, the databases Swiss-Prot, Protein Identification Resource (PIR), and GenBank were screened by using the program BLAST (14) with default parameters (National Center for Biotechnology Information's BLAST WWW server; <http://www.dot.imgen.bom.tmc.edu:9331/seq-search/protein-search.html>). The cutoff for the score was set to 70, and protein fragments and nearly identical sequences were eliminated. A genealogical tree was subsequently constructed for the sequence homologues with the program ALLALL (15). For secondary structure prediction we employed the program PHDsec (16), and Fig. 2 was generated by using the molecular graphics program MOLMOL (17).

Abbreviations: PR, pathogenesis-related; GliPR, glioma PR protein; PAM, accepted point mutation; RTVP-1, related to testes-specific, vespid, and pathogenesis proteins; PIR, Protein Identification Resource.

\*To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/952262-5\$2.00/0  
PNAS is available online at <http://www.pnas.org>.

Table 1. Structure calculation for human GliPR, using an input of conformational constraints obtained with the NMR- and homology-based approach described in the text and comparison with the NMR structure of P14a (7)

	GliPR	P14a
Conformational constraints		
Total number of constraints	1,216	2,021
Upper distance limits	932	1,692
Disulfide bond constraints*	18	18
$\beta$ -Sheet hydrogen bond constraints*	32	—
$\psi, \phi$ -dihedral angle constraints	202	228
$\chi^1$ -dihedral angle constraints	32	83
20 DIANA conformers used to represent the structure, after energy minimization (average value $\pm$ SD)		
DIANA target function, $\text{\AA}^2$ †	6.5 $\pm$ 0.7	3.1 $\pm$ 0.7
NOE constraint violations, $\text{\AA}$ :		
Sum	7.86 $\pm$ 0.49	9.20 $\pm$ 0.64
Max	0.11 $\pm$ 0.14	0.10 $\pm$ 0.01
Dihedral angle violations, $^\circ$ :		
Sum	25.6 $\pm$ 4.3	45.0 $\pm$ 5.1
Max	2.4 $\pm$ 0.3	2.3 $\pm$ 0.2
AMBER energies, kcal·mol <sup>-1</sup> :		
Van der Waals	-358 $\pm$ 21	-570 $\pm$ 12
Electrostatic	-4,203 $\pm$ 182	-4,576 $\pm$ 74
Average rms deviations for different atom selections (average value $\pm$ SD)		
Backbone N, C $^\alpha$ , C' of constrained segments‡	1.66 $\pm$ 0.26	0.88 $\pm$ 0.08
All heavy atoms of constrained segments‡	2.37 $\pm$ 0.24	1.30 $\pm$ 0.09
Backbone of the regular secondary structures§	0.98 $\pm$ 0.19	0.45 $\pm$ 0.08
Backbone of constrained segments‡ + core side chains¶	1.60 $\pm$ 0.21	0.91 $\pm$ 0.07

\*In GliPR, the disulfide bonds between C65–C146, C121–C125, and C141–C163, and the hydrogen bonds W91...F161, A142...I162, Q144...H160, H160...Q144, F161...W91, I162...A142, C163...N89, and N164...G140 in the  $\beta$ -sheet were constrained following ref. 8.

†Before energy minimization.

‡“Constrained segments” of GliPR are those parts of the polypeptide chain for which conformational constraints were available from the NMR measurements with P14a, which includes the residues 21–38, 44–72, 87–146, 148–151, and 159–176 (see also Fig. 1).

§The secondary structure elements in GliPR are as follows:  $\alpha$ -helix I, residues 24–37;  $\alpha$ -II, 52–64;  $\alpha$ -III, 100–111;  $\alpha$ -IV, 127–132;  $\beta$ -strand A, 49–50;  $\beta$ -B, 87–92;  $\beta$ -C, 138–145;  $\beta$ -D, 159–166 (see Fig. 1).

¶The “molecular core” of GliPR includes the residues 27, 28, 30, 31, 34, 35, 38, 45, 48, 50, 54, 55, 58, 61, 62, 65, 67, 89–91, 98, 103, 104, 106–108, 114, 116, 121, 125, 128–131, 133, 136, 139–142, 144, 146, 160, 161, 163–165, 170, 175, 176, which are indicated with “1” or “2” in Fig. 1. The displacements calculated for all these residues except C65, F67, C121, and C146, after superimposing the backbone heavy atoms N, C $^\alpha$ , and C' of the “constrained segments” and the side chain heavy atoms of the core residues for minimal rms deviation, are smaller than 2.3  $\text{\AA}$ . The corresponding displacements in the NMR solution structure of P14a (7) were all smaller than 1.9  $\text{\AA}$ .

## RESULTS AND DISCUSSION

The sequence alignment of the 219-residue protein GliPR and the 135-residue protein P14a of Fig. 1 shows that GliPR possesses a 20-residue N-terminal extension and a 43-residue C-terminal extension of the polypeptide chain. Moreover, there are four peptide segments in GliPR, comprising residues 39–43, 73–86, 147, and 152–158, that are inserted into the polypeptide chain of P14a, whereas the segments 40–43, 86–87, and 133 of P14a are not present in GliPR. P14a is related with other PR-1 proteins by numerous highly conserved residues that have essential structural roles in the architecture of the P14a core (7). Although the sequence homology of P14a with GliPR is significantly lower (35% identity) than with other PR-1 proteins (>54% identity), similar conservation is readily apparent in Fig. 1, and the additions to the sequence of P14a are all located either as insertions between regular secondary structures or as extensions at the chain ends (Fig. 1). Hence, the sequence alignment of GliPR and P14a in Fig. 1 indicates similarity also on the level of the three-dimensional structures of the two proteins.

Direct three-dimensional structure comparison of GliPR and P14a was based on the high-quality NMR structure of P14a (7). The structure calculations performed for GliPR with the input adapted from the experiments with P14a (see *Methods*) converged with low residual target function values and constraint violations (Table 1), showing that the distance geometry algorithm found sterically allowed three-dimensional

folds as well as good side chain packing in the molecular core of GliPR. This is further evidenced by low van der Waals and electrostatic AMBER energies (10), which are comparable to those obtained for the high-quality NMR solution structure of P14a. The average rms deviations among the conformers selected to represent the structure of GliPR (Table 1) correspond to those of a good-quality NMR structure determination. The presently employed approach for homology modeling based on a high-quality NMR solution structure has the obvious advantage that the sampling of conformation space performed for the NMR structure is transferred to the target sequence, which avoids steric clashes that might arise if a single conformer of the template molecule were used to derive a set of distance restraints (for a recent review of sequence homology-based structure prediction, see ref. 18).

In the molecular model of GliPR (Fig. 2a) obtained with this NMR- and homology-based approach the regular secondary structure elements form the same, so far unique,  $\alpha$ - $\beta$ - $\alpha$  sandwich as P14a (Fig. 2b), where the intermediary layer is formed by a four-stranded mixed  $\beta$ -sheet of topology  $+3x, -2x, +1$ . The visual impression of near-identity of the polypeptide folds of GliPR and P14a is substantiated by a rms deviation of 0.88  $\text{\AA}$  calculated between the mean coordinates of the two sets of 20 DIANA conformers for the backbone heavy atoms of the 128 residues that are aligned in identical positions in Fig. 1. The spatial arrangement of regular secondary structure elements leads to a bipartite hydrophobic core

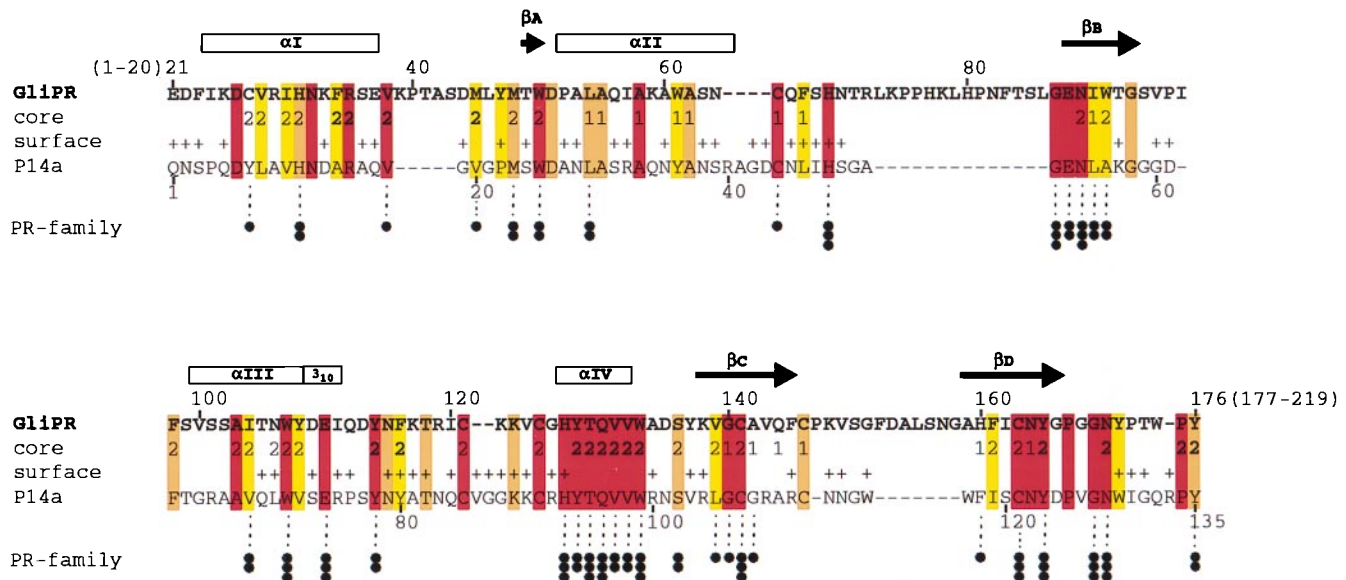


FIG. 1. Sequence alignment of the proteins GliPR and P14a. The one-letter amino acid code is used, and the numbering for GliPR is indicated above its sequence. The locations of the regular secondary structure elements identified for P14a (7) are depicted at the top, with the four  $\alpha$ -helices and the  $3_{10}$ -helix represented by boxes marked  $\alpha$ I,  $\alpha$ II,  $\alpha$ III,  $\alpha$ IV, and  $3_{10}$ , and the four  $\beta$ -strands, by the arrows  $\beta$ A,  $\beta$ B,  $\beta$ C, and  $\beta$ D. In the row "core" the residues forming a small hydrophobic cluster associated with the  $\beta$ -sheet and helix II (Fig. 2 *a* and *b*) are indicated 1, those constituting the larger core associated with the  $\beta$ -sheet and helices I, III, and IV (Fig. 2 *a* and *b*) with 2, and residues of helix IV and those directly contacting helix IV in P14a (Fig. 2 *a* and *b*) with 2. In the row "surface," + identifies residues that exhibit a solvent accessibility larger than 25% in at least one of the 20 energy-minimized DIANA conformers of P14a. Color code: residues conserved between GliPR and P14a are orange, and those conserved also within the whole group of known plant PR-1 proteins (see figure 10 in ref. 7) are red. Conserved hydrophobicity of Ala, Val, Leu, Ile, Tyr, Phe, uncharged His, Trp, Pro, Met, or Cys in GliPR and all PR-1 proteins is highlighted in yellow. Conservation within the "PR-protein superfamily" (see text and Fig. 3), derived from multiple alignment of all 26 representative members, is indicated at the bottom: three dots label residues that are strictly conserved within the superfamily, two dots denote conservation of side chain character (hydrophobic: Ala, Val, Leu, Ile, Tyr, Phe, Trp, Met, Cys, and Pro; hydroxyl-bearing: Ser and Thr; or hydrogen bond acceptors: uncharged His, Asp, Asn, Glu, and Gln), and one dot indicates mutations from hydrophobic amino acids to Gly, Thr, or His.

associated with the two layer interfaces of the sandwich, which is further stabilized by the three disulfide bonds. The  $\alpha$ -helix IV is completely buried between the  $\beta$ -sheet and the two-helix bundle of helices I and III, and it is fully integrated into the larger of the two clusters. For the insertions and extensions of the polypeptide chain relative to P14a (Fig. 1), for which the presently used approach does not yield conformational constraints, secondary structure prediction did not indicate any regular secondary structures.

Inspection of the three-dimensional structures shows that only His-69 and His-127 are solvent exposed and strictly conserved in GliPR and the family of plant PR-1 proteins (Fig. 1), which makes these residues prime candidates for a role in functionally active sites (19). These histidyls are in close spatial proximity to the highly conserved glutamates 88 and 110, which also possess no obvious structural role. Hence, the cluster comprising these two histidines and two glutamates is a likely candidate for the so-far-unknown active site of both GliPR (Fig. 2*c*) and the plant PR-1 proteins represented here by P14a (Fig. 2*d*). There are no other conserved hydrophobic, polar, or charged amino acid combinations on the protein surface that could account for a preserved interaction of the functionally related plant PR-1 proteins (4, 5) with their biomolecular targets. To further evaluate the likelihood that the clusters identified in Fig. 2 *c* and *d* are related to a functional site, we identified the largest surface clefts of P14a and GliPR. The calculation showed that the cleft containing the putative active site residues is the largest cleft on both protein surfaces (Fig. 2 *c* and *d*). Because statistically enzyme active sites are typically located in the largest surface cleft of a protein (20, 21) this suggests that GliPR and plant PR-1 proteins might function as enzymes with the putative active sites shown in Fig. 2 *c* and *d*. In these, His-127 is the C-terminal residue of  $\alpha$ -helix IV, which is completely buried within the

molecular core (Fig. 2 *a* and *b*). All residues of this helix and the directly succeeding Trp are strictly conserved in GliPR and the plant PR-1 proteins and the residues that contact helix IV are also highly conserved (Fig. 1), so that the distinct hydrogen bonding network identified within the core of P14a (see figure 8 of ref. 7) is maintained in GliPR. The embedding of helix IV within the core and its positional fine adjustment with a network of hydrogen bonds (22) could thus be the structural basis of the function of PR proteins. The dipole of helix IV may even serve for proper tuning of the pK<sub>a</sub> value of His-127 (23). In P14a, His-48, Glu-53, and His-93 (Fig. 2*d*) are flexibly disordered in solution (7), but they could adopt the conformation found for the active site histidyl and glutamyl residues in several Zn proteases. However, no Zn<sup>2+</sup> ion could be unambiguously identified in an x-ray crystal structure of P14a obtained with crystals that were grown in the presence of Zn<sup>2+</sup> (V. Mikol, personal communication). The two histidines could also represent the active site of a ribonuclease, but biochemical assays performed to detect ribonuclease activity were negative (E. Mösinger, personal communication).

Overall, the strict conservation of the putative active site residues (Fig. 2 *c* and *d*) strongly suggests that human GliPR and plant PR-1 proteins operate according to the same molecular mechanism, which establishes a possible functional link between the human immune system and a plant defense system. In this context the evolutionary origin of these PR proteins is of keen interest. The close structural similarity of the molecular cores of P14a and GliPR (Fig. 2 *a* and *b*) speaks against the convergent evolution of two independent ancestors, because even if this had led to the same fold and active site geometry of the two proteins it would very likely have resulted in different architectures for the molecular cores. The alternative assumption that P14a and GliPR arose from a common ancestor raises questions concerning the role of the



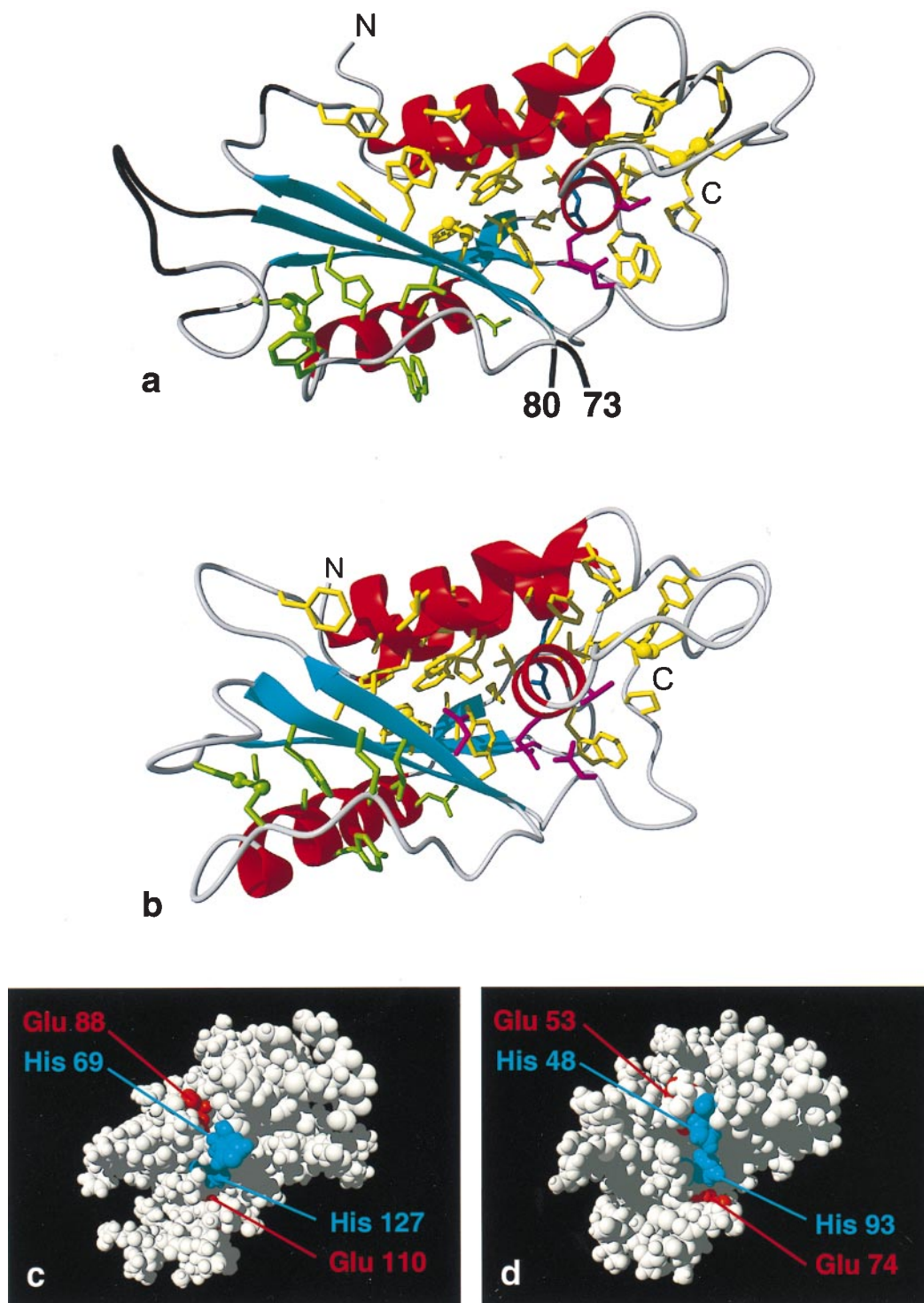


FIG. 2. (a) Ribbon drawing of the DIANA conformer of GliPR with the lowest residual target function after energy minimization. The input was obtained from a NMR- and homology-based approach (see text). The helices I to IV are shown in red, the  $\beta$ -strands A to D in cyan, and other polypeptide segments in gray, and the polypeptide chain ends are indicated with N and C. Polypeptide segments that represent insertions into the amino acid sequence of P14a (see Fig. 1) are depicted in black. Amino acid side chains that form a large hydrophobic cluster associated with the  $\beta$ -sheet and the  $\alpha$ -helices I, III, and IV are shown in yellow, with the two disulfide bonds C121–125 and C141–C163 drawn as ball and stick models, and side chains that belong to a smaller hydrophobic cluster associated with the  $\beta$ -sheet,  $\alpha$ -helix II, and the disulfide bond C65–C146 are green. Polar side chains in the hydrophobic core are identified by drawing His-31, Thr-129, Gln-130, and Asn-170 in magenta and Arg-35 in blue. (b) Best DIANA conformer from the NMR structure determination of P14a (7) in the same orientation and with the same color code as used for GliPR in a. The magenta-colored polar side chains are His-11, Thr-95, Gln-96, and Asn-128, and the blue side chain is Arg-15. (c and d) Space-filling representations of GliPR and P14a, respectively, showing the putative active site (see text), with histidiny and glutamyl residues depicted in blue and red, respectively. For clarity, in GliPR the insertions relative to the sequence of P14a are omitted (Fig. 1).

ancestor protein, which presumably functioned at a very early stage in evolution when living organisms separated into different kingdoms and the presently known defense systems

probably did not yet exist. A possible solution would be that these proteins have been horizontally transferred (24) at a much later stage of evolution—e.g., PR-1 proteins might

originally have evolved in plants and subsequently been recruited by mammals.

To follow up this possibility we constructed a genealogical tree (Fig. 3) for a set of 26 sequence homologues of GliPR. The accepted point mutation (PAM) distances between the sequences and the resulting topology of the tree suggest that GliPR and P14a arose from a common ancestor that has evolved into a large "PR-protein superfamily." Its members include human GliPR, mammalian sperm coating proteins (25), plant PR-1 proteins, allergens of insect venoms (26), and snake or lizard toxins, and are thus found in the three kingdoms of animals, plants, and fungi. The underlying molecular mechanism for the action of these proteins is unknown. The alignment of their amino acid sequences with GliPR revealed a striking conservation of amino acid residues corresponding to the buried helix IV of GliPR and P14a (Figs. 1 and 2 *a* and *b*) within the superfamily (see also refs. 3, 25, and

26). In addition, numerous residues aligned with residues that form the molecular core show significant conservation (Fig. 1), in particular those that participate in the hydrogen bonding network in the core of GliPR (His-33, Asn-89, Tyr-165; see also figure 8 in ref. 7). The evolutionary distances between GliPR and P14a (PAM = 120), and between GliPR and the other members of the superfamily are comparable (Fig. 3). It may thus well be that the  $\alpha$ - $\beta$ - $\alpha$  sandwich fold of P14a and GliPR is characteristic for the entire superfamily. Moreover, the putative active site residues His-69, Glu-110, and His-127 of GliPR (Fig. 2 *c* and *d*) are conserved in all but one sequence, and Glu-88 of GliPR is conserved in 19 of the 26 sequences, while the other 7 possess a Gln in the corresponding position (Fig. 3), indicating that all proteins of the PR-protein superfamily could, on the basis of present knowledge of the molecular structure, operate according to the same molecular mechanism.

We thank Drs. V. Mikol and E. Mösinger for communicating unpublished results. Financial support was obtained from the Schweizerischer Nationalfonds (project 31.49047.96) and a fellowship to C.F. from the Schweizerische Bundesstipendienkommission.

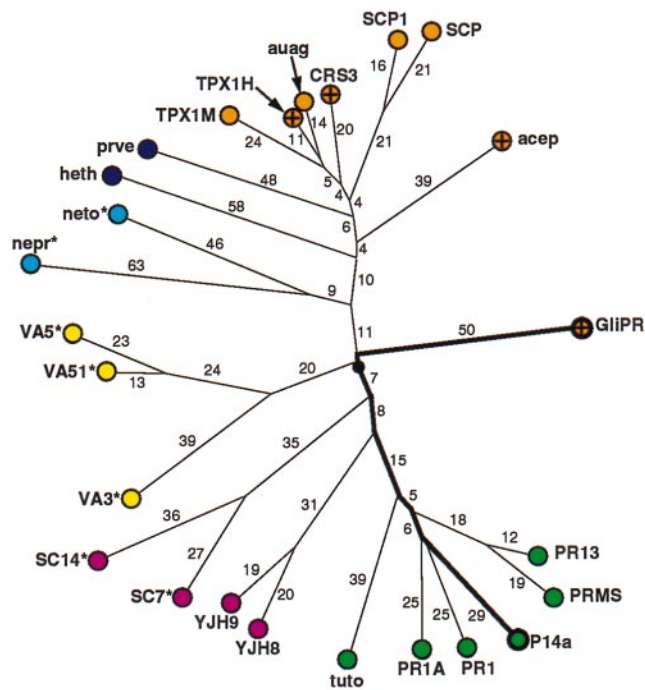


FIG. 3. Genealogical tree of the "PR-protein superfamily," where the distances are given in PAM units. For clarity, distances with PAM = 1 or 2 are not indicated. Color code: mammals, orange (human proteins are further marked with a cross); plants, green; fungi, magenta; insects, yellow; nematodes, light blue; and other vertebrates, dark blue. The branches are labeled either with the first four letters of the Swiss-Prot database code in capital letters, or with an abbreviation derived from the GenBank and PIR database codes. Proteins for which the putative active site (Fig. 2 *c* and *d* and text) is not strictly conserved are identified as follows: where Glu-88 of GliPR is exchanged with Gln there is an asterisk. For "heth" the residue corresponding to His-69 of GliPR is exchanged with Glu. Abbreviations: from GenBank and PIR, acep, acidic epididymal glycoprotein, S80310; auag, autoantigen 1 of the guinea pig sperm acrosome, U35712; heth, helothermine precursor (Mexican beaded lizard), U13619; nemt, gene sequence from chromosome III of nematode *Caenorhabditis elegans*, Z68108; nepr, nematode protective antigen, A28112; prvn, Prepro-Cys-rich venom protein, U59447; tuto, tumor-related protein from tobacco; from Swiss-Prot, CRS3\_HUMAN, PR13\_HORVU, PR1A\_TOBAC, PR1\_ARATH, PRMS\_MAIZE, SC14\_SCHCO, SC7\_SCHCO, SCP1\_MOUSE, SCP\_RAT, SGP28\_X94323, TPX1\_HUMAN, TPX1\_MOUSE, VA3\_SOLIN, VA5\_VESCR, VA5\_VESPE, VJH8\_YEAST, and VJH9\_YEAST.

- Murphy, E. V., Zhang, Y., Zhu, W. & Biggs, J. (1995) *Gene* **159**, 131–135.
- Morris, J. H. & Schoene, W. C. (1984) in *The Pathological Basis of Disease*, eds. Robbins, S. L., Cotran, R. F. & Kumar, V. (Saunders, Philadelphia), pp. 1401–1456.
- Rich, T., Chen, P., Furman, F., Huynh, N. & Israel, M. A. (1996) *Gene* **180**, 125–130.
- Bol, J. F., Linthorst, H. J. M. & Cornelissen, B. J. C. (1990) *Annu. Rev. Phytopathol.* **28**, 113–138.
- Ryals, J., Uknes, S. & Ward, E. (1994) *Plant Physiol.* **104**, 1109–1112.
- Orengo, C. A., Jones, D. T. & Thornton, J. M. (1994) *Nature (London)* **372**, 631–634.
- Fernández, C., Szyperski, T., Bruyère, T., Ramage, P., Mösinger, E. & Wüthrich, K. (1997) *J. Mol. Biol.* **266**, 576–593.
- Williamson, M. P., Havel, T. F. & Wüthrich, K. (1985) *J. Mol. Biol.* **182**, 295–315.
- Güntert, P., Braun, W. & Wüthrich, K. (1991) *J. Mol. Biol.* **217**, 517–530.
- Weiner, P. K., Kollman, P. A., Nguyen, D. & Case, D. A. (1986) *J. Comp. Chem.* **7**, 230–252.
- Luginbühl, P., Güntert, P., Billeter, M. & Wüthrich, K. (1996) *J. Biomol. NMR* **8**, 136–146.
- Laskowski, R. A. (1995) *J. Mol. Graph.* **13**, 323–330.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
- Gonnet, G. H., Cohen, M. A. & Benner, S. A. (1992) *Science* **256**, 1443–1445.
- Rost B. & Sander, C. (1994) *Proteins* **19**, 55–72.
- Koradi, R., Billeter, M. & Wüthrich, K. (1996) *J. Mol. Graph.* **14**, 51–55.
- Sali, A. (1995) *Curr. Opin. Biotechnol.* **6**, 437–451.
- Schulz, G. E. & Schirmer, R. H. (1979) *Principles of Protein Structure* (Springer, Berlin).
- Laskowski, R. A., Luscombe, N. M., Swindells, M. B. & Thornton, J. M. (1996) *Protein Science* **5**, 2438–2452.
- Peters, K. P., Fauck, J. & Frömmel, C. (1996) *J. Mol. Biol.* **256**, 201–213.
- Sauer, R. T. (1996) *Folding Design* **1**, R27–R30.
- Sancho, J., Serrano, L. & Fersht, A. R. (1992) *Biochemistry* **31**, 2253–2258.
- Syvanen, M. (1994) *Annu. Rev. Genet.* **28**, 237–261.
- Kjeldsen, L., Cowland, J. B., Johnson, A. H. & Borregaard, N. (1996) *FEBS Lett.* **380**, 246–250.
- Lu, G., Villalba, M., Coscia, M. R., Hoffman, D. R. & King, T. P. (1993) *J. Immunol.* **150**, 2823–2830.