



**HAL**  
open science

## Small-Angle x-ray scattering and crystallographic studies of arcelin-1: An insecticidal lectin-like glycoprotein from *Phaseolus vulgaris* L

Lionel Mourey, Jean-Denis Pedelacq, Christine Fabre, Henri Causse, Pierre Rougé, Jean-Pierre Samama

### ► To cite this version:

Lionel Mourey, Jean-Denis Pedelacq, Christine Fabre, Henri Causse, Pierre Rougé, et al.. Small-Angle x-ray scattering and crystallographic studies of arcelin-1: An insecticidal lectin-like glycoprotein from *Phaseolus vulgaris* L. *Proteins: Structure, Function, and Genetics*, 1997, 29 (4), pp.433-442. 10.1002/(sici)1097-0134(199712)29:43.0.co;2-9 . hal-03004639

**HAL Id: hal-03004639**

**<https://cnrs.hal.science/hal-03004639>**

Submitted on 20 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Small-Angle X-Ray Scattering and Crystallographic Studies of Arcelin-1: An Insecticidal Lectin-Like Glycoprotein From *Phaseolus vulgaris* L

Lionel Mourey,<sup>1</sup> Jean-Denis Pédelacq,<sup>1</sup> Christine Fabre,<sup>2</sup> Henri Causse,<sup>2</sup> Pierre Rougé,<sup>2</sup> and Jean-Pierre Samama<sup>1\*</sup>

<sup>1</sup>Groupe de Cristallographie Biologique, Institut de Pharmacologie et de Biologie Structurale, UPR 9062 CNRS, F-31077 Toulouse Cedex, France

<sup>2</sup>Groupe Lectines et Reconnaissance, Institut de Pharmacologie et de Biologie Structurale, UPR 9062 CNRS, F-31077 Toulouse Cedex, France

**ABSTRACT** Arcelin-1 and  $\alpha$ -amylase inhibitor are two lectin-like glycoproteins expressed in the seeds of the kidney bean (*Phaseolus vulgaris*). They display insecticidal activities and protect the seeds from predation by larvae of various bruchids through different biological actions. Solution-state investigations by small-angle X-ray scattering (SAXS) show the dimeric structure of arcelin-1, a requirement for its hemagglutinating properties. Anions were found to have specific properties in their effectiveness to disrupt protein aggregates, affect solubility, and improve crystallizability. The SAXS results were used to improve crystallization conditions, and single crystals diffracting beyond 1.9 Å resolution were obtained. X-ray diffraction data analysis shows that noncrystallographic symmetry-related arcelin-1 molecules form a lectin-like dimer and reveals the presence of a solvent-exposed anion binding site on the protein, at a crystal-packing interface. The solution state properties of arcelin-1 and crystal twinning may be explained by the anion specificity of this binding site. *Proteins* 29:433–442, 1997.

© 1997 Wiley-Liss, Inc.

**Key words:** glycoprotein; insecticide; bruchids; crystallization; X-ray structure

## INTRODUCTION

Lectins form a large superfamily of proteins involved in many interactions at the cell surface through specific carbohydrate recognition. They are widely spread in plants, animals, viruses, and bacteria, and many of them have been isolated and characterized.<sup>1</sup> Precise understanding of their biological function remains unclear, despite several X-ray structure determinations that provided the structural basis for sugar recognition and binding.<sup>2</sup>

Five different structural classes of plant lectins, differing in primary sequences and overall folds, can

be distinguished: legume lectins, hevein-related (cereal) lectins, monocot mannose-binding (bulb) lectins, jacalin, and ribosome-inactivating proteins (RIPs). Lectins of the *Leguminosae* share a common  $\beta$ -sandwich tertiary structure that can be described as an ellipsoidal dome. However, these closely related proteins differ in their sugar specificities and quaternary organizations.<sup>3</sup> Dimers (pea lectin,<sup>4</sup> *Lathyrus ochrus* isolectins,<sup>5,6</sup> *Erythrina corallodendron* lectin,<sup>7</sup> *Griffonia simplicifolia* lectin IV,<sup>8</sup> lentil lectin<sup>9</sup>) and tetramers (concanavalin A,<sup>10</sup> peanut agglutinin,<sup>11</sup> soybean agglutinin,<sup>12</sup> phytohemagglutinin-L<sup>13</sup>) have been described from the X-ray structures. Within this class, four glycoproteins are found in the seeds of kidney bean (*Phaseolus vulgaris*): two lectins, phytohemagglutinin-L (PHA-L) and phytohemagglutinin-E, and the related lectin-like proteins, arcelins and  $\alpha$ -amylase inhibitors ( $\alpha$ -AI). These proteins have different biochemical properties, are encoded by a cluster of physically linked genes,<sup>14</sup> and share 55–60% sequence identity. Sequence alignment suggests that the canonical monosaccharide-binding site of legume lectins, delineated by four loop regions, is altered in arcelin and  $\alpha$ -AI due to the deletion of one and two of these loops, respectively (Fig. 1).<sup>15</sup>

PHA-L binds complex-type oligosaccharides, and its three-dimensional structure pointed out a possible binding site for adenine-derived plant hormones.<sup>13</sup>  $\alpha$ -AI protects bean seeds against predation by some insect pests through the inactivation of  $\alpha$ -amylase. The molecular basis of the enzyme inhibition was recently revealed by the crystal structure of the tight 1:2 complex formed between the inhibitor and pig pancreatic  $\alpha$ -amylase.<sup>16</sup> Arcelin-1 also protects kidney bean seeds against the larvae of various

*Abbreviations:* MES, 2-Morpholinoethanesulfonic acid; PEG, polyethylene glycol; SAXS, small angle X-ray scattering.

\*Correspondence to: Jean-Pierre Samama, IPBS-CNRS, 205 route de Narbonne, 31077 Toulouse Cedex, France.  
E-mail: samama@ipbs.fr

Received 27 March 1997; Accepted 26 June 1997

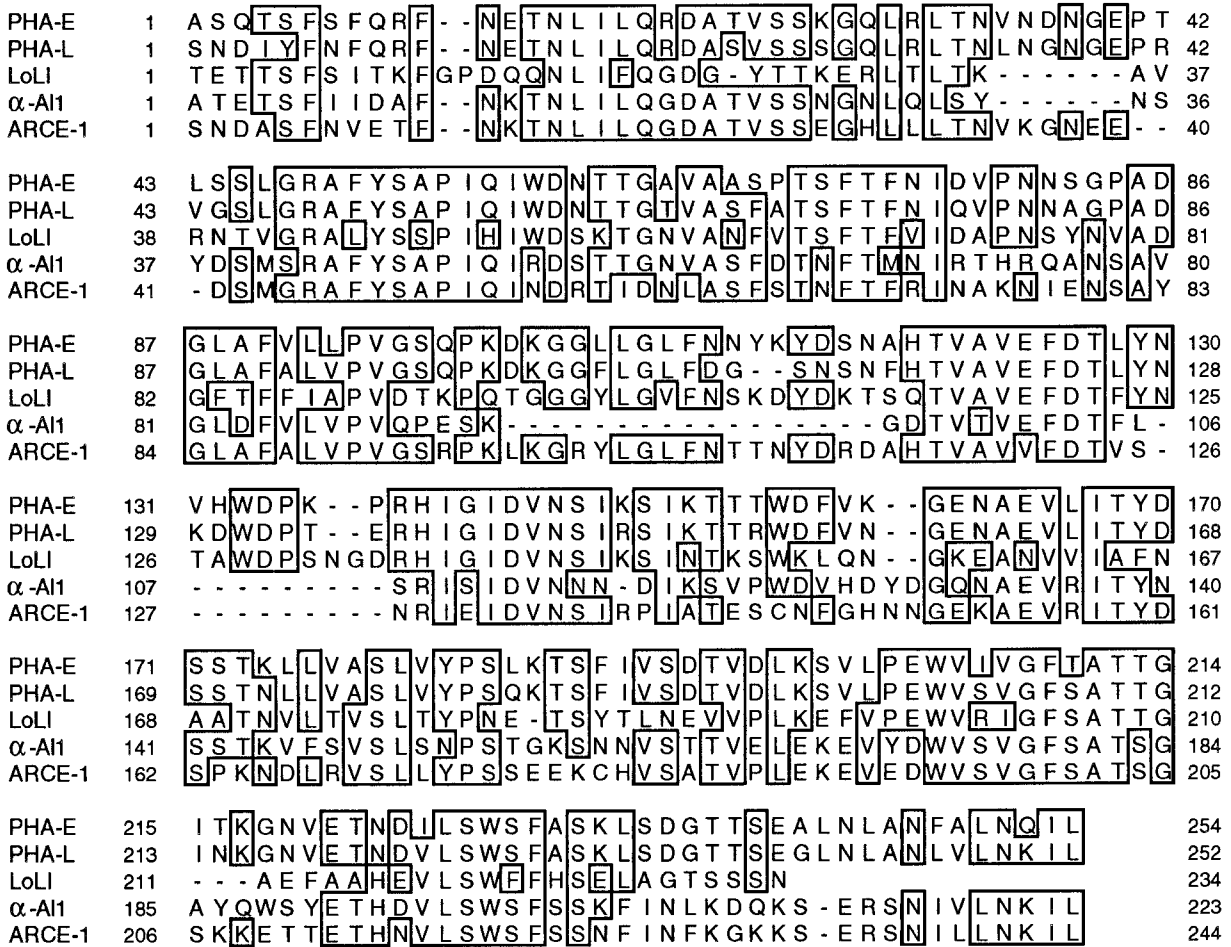


Fig. 1. Sequence alignment of *Lathyrus ochrus* isolectin I (LoLI), erythroagglutinating (PHA-E), and leucoagglutinating (PHA-L) subunits of phytohemagglutinin,  $\alpha$ -amylase inhibitor 1 ( $\alpha$ -AI1), and arcelin-1 (ARCE-1). The sequence identities between

arcelin-1 and LoLI, PHA-E, PHA-L, and  $\alpha$ -AI1 are 45, 55, 58, and 60%, respectively. The four loop regions implicated in monosaccharide binding in LoLI are: 73–81, 95–102, 123–135, and 210–215.

bruchids, such as the Mexican bean weevil (*Zabrotes subfasciatus*).<sup>17</sup> However, this protection occurs through a different mechanism than that of  $\alpha$ -AI and is still not well understood. A step toward the understanding of the structure-function relationships of arcelin-1 should be provided by solution-state investigations and high resolution X-ray structure determination.

In this article, we describe the first results of the crystallographic study of arcelin-1 and emphasize the use of small-angle X-ray scattering as a diagnostic tool for producing crystals of this glycoprotein, diffracting beyond 1.9 Å resolution.

## MATERIALS AND METHODS

### Solution X-Ray Scattering

Arcelin-1 was purified from seed flour by using a two-step procedure based on ion exchange chromatography and gel chromatofocusing (P. Rougé, unpublished data). All SAXS and crystallization experiments were performed by using the same purification

batch. Stock protein solutions (10 mg/ml) were prepared by dissolving lyophilized arcelin-1 in water. They were then centrifuged for 1 h at 4°C and 15,000 rpm to discard any insoluble material. The protein solutions used for SAXS measurements were prepared by mixing equal volumes of stock arcelin-1 solution and twice-concentrated solutions of the different media (Table I). Ionic strength values were calculated by using  $I = \frac{1}{2} \sum c_i Z_i^2$  with  $c_i$  and  $Z_i$  the concentration and valency of all ionic species present in the solution. In the case of the polyacid sodium citrate, the concentrations of each anion were calculated by using the program EQS4WIN (Mathrek Systems) and the different pK values<sup>18</sup> at the given pH. All samples contained 1.3 mM DTT to eliminate the free radicals formed in solution under X-ray irradiation.

X-ray scattering curves were recorded on the small-angle-scattering instrument D24<sup>19</sup> by using synchrotron radiation at LURE-DCI (Orsay, France). The instrument and the data acquisition system

**TABLE I. Conditions Used for Arcelin-1 SAXS Experiments\***

Buffer	Added salt		pH	<i>I</i> (M)	$R_g$ (Å ± 0.2 Å)
	Type	Conc. (M)			
MES	—		4.9	0.10	—
MES	NaSCN	0.3	4.9	0.40	—
MES	NaCl	0.3	4.9	0.40	28.6
MES	Na citrate	0.3	4.9	0.52	28.8
MES	NaSCN	1.0	4.9	1.10	28.7
MES	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	0.3	4.9	1.00	29.3
Na citrate	—		5.0	0.15	28.8
Na citrate	NaSCN	0.3	5.0	0.45	29.0
Na citrate	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	0.3	5.0	1.05	29.4

\*All protein and buffers concentrations were 5 mg/ml and 0.1 M, respectively.

were described previously.<sup>20</sup> The wavelength of the X-rays was 1.488 Å (K-edge of Ni), and the sample to detector distance was set to 1827 mm. Scattered intensities were recorded in the angular range  $0.00132 < s \text{ (Å}^{-1}\text{)} < 0.02976$ , where  $s$  is the scattering vector ( $2 \sin \theta / \lambda$ ), and  $2\theta$  and  $\lambda$  are the scattering angle and wavelength of the X-rays, respectively. All experiments were performed at 4°C by using a temperature-controlled cell.<sup>21</sup> Eight successive frames of 100 s each were collected for each sample. The scattering intensity of a reference sample of carbon black, recorded immediately before and after each sample, was used to normalize all data to the transmitted intensity. The scattering contribution of the buffer was subtracted before further analysis.

The radius of gyration ( $R_g$ ) was deduced from Guinier's law:

$$I(s) = I(0) \exp(-4\pi^2 R_g^2 s^2 / 3)$$

where  $I(s)$  is the SAXS intensity,  $I(0)$  is the value of the extrapolated scattered intensity at  $s = 0$ .<sup>22</sup> For monodisperse solutions, the Guinier plot of  $\ln I(s)$  versus  $s^2$  approximates a straight line with a slope proportional to  $R_g^2$  in the small-angle region. After normalization to protein concentration ( $c$ ), the value of  $I(0)/c$  is proportional to the molecular weight of the scattering entity. To calibrate this value, we used a monodisperse solution of 3-phosphoglycerate kinase (PGK, MW 44,500 Da). Measurements with PGK were performed in 20 mM Tris buffer pH 7.5 by using identical experimental settings.

### Crystallization

All crystallization trials were performed at 4°C by vapor diffusion using the hanging drop method and freshly prepared aqueous protein solutions (10 mg/ml). Crystals diffracting to 1.9 Å resolution were obtained by mixing 2 μl of arcelin-1 solution and 1 μl of reservoir solution containing 25% PEG400 and 300 mM ammonium sulfate in 100 mM MES buffer at pH 4.8 or 4.9.

### Data Collection and Processing: Molecular Replacement

Crystals were characterized in the laboratory with a Rigaku RU300 rotating anode producing Cu-K<sub>α</sub> X-rays collimated with the Yale double-mirror focusing (Molecular Structure Corporation) and an R-Axis II imaging plate detector (Rigaku).

A 1.9 Å native data set was collected on the W32 wiggler beam line at LURE-DCI (Orsay, France), tuned at a wavelength of 0.938 Å, and equipped with a large MarResearch scanner. The crystal to detector distance was set to 270 mm, and the crystal temperature was maintained at 4°C. A total of 58 frames (1.5° oscillation per frame and 3-min time exposure) was collected from a single arcelin-1 crystal (660 × 400 × 180 μm<sup>3</sup>). Data were processed with MOSFLM<sup>23</sup> and the CCP4 suite of programs.<sup>24</sup>

Molecular replacement calculations were performed with the AMoRe program<sup>25</sup> as implemented in the CCP4 package. The *Lathyrus ochrus* isolectin I crystallographic structure<sup>5</sup> (PDB<sup>26</sup> access code 1LOE) was used as a search model for the rotation and the translation functions. It was truncated as a polyaniline with  $B$  factors set to 20.0 Å<sup>2</sup>. In addition, residues 125–134, which are absent in arcelin-1, according to sequence alignment, were removed.

## RESULTS

### SAXS Experiments

The curvature of the Guinier plot at small  $s$  values indicated the presence of protein aggregates when experiments were conducted in MES buffer pH 4.9,  $I = 0.1$  M (Fig. 2A). The effect of anions on the monodispersity of arcelin-1 solutions was investigated with SAXS measurements performed in the presence of thiocyanate, chloride, and citrate ions (Fig. 2B–D), in the same buffer and at low ionic strength ( $I = 0.4$ – $0.5$  M). We found that the improvement of monodispersity followed the lyotropic character of the anions (citrate > chloride > thiocyanate).<sup>\*</sup> The most chaotropic thiocyanate ions, used at higher concentration ( $I = 1.1$  M), further decreased aggregation (Fig. 3B versus 2B), suggesting that the ionic strength was a second parameter that improved monodispersity. Accordingly, the Guinier plot indicated that aggregation was nearly abolished when the most lyotropic sulfate ions were used at  $I = 1.0$  M (Fig. 3C).

With regard to protein crystallization, the relative contributions of the ionic strength and of the lyotropic properties of anions were approached by comparing experiments performed in 100 mM sodium

\*The lyotropic to chaotropic classification of anions has been related to several properties of biological macromolecules, such as folding, stability, and solubility.<sup>27,28</sup>

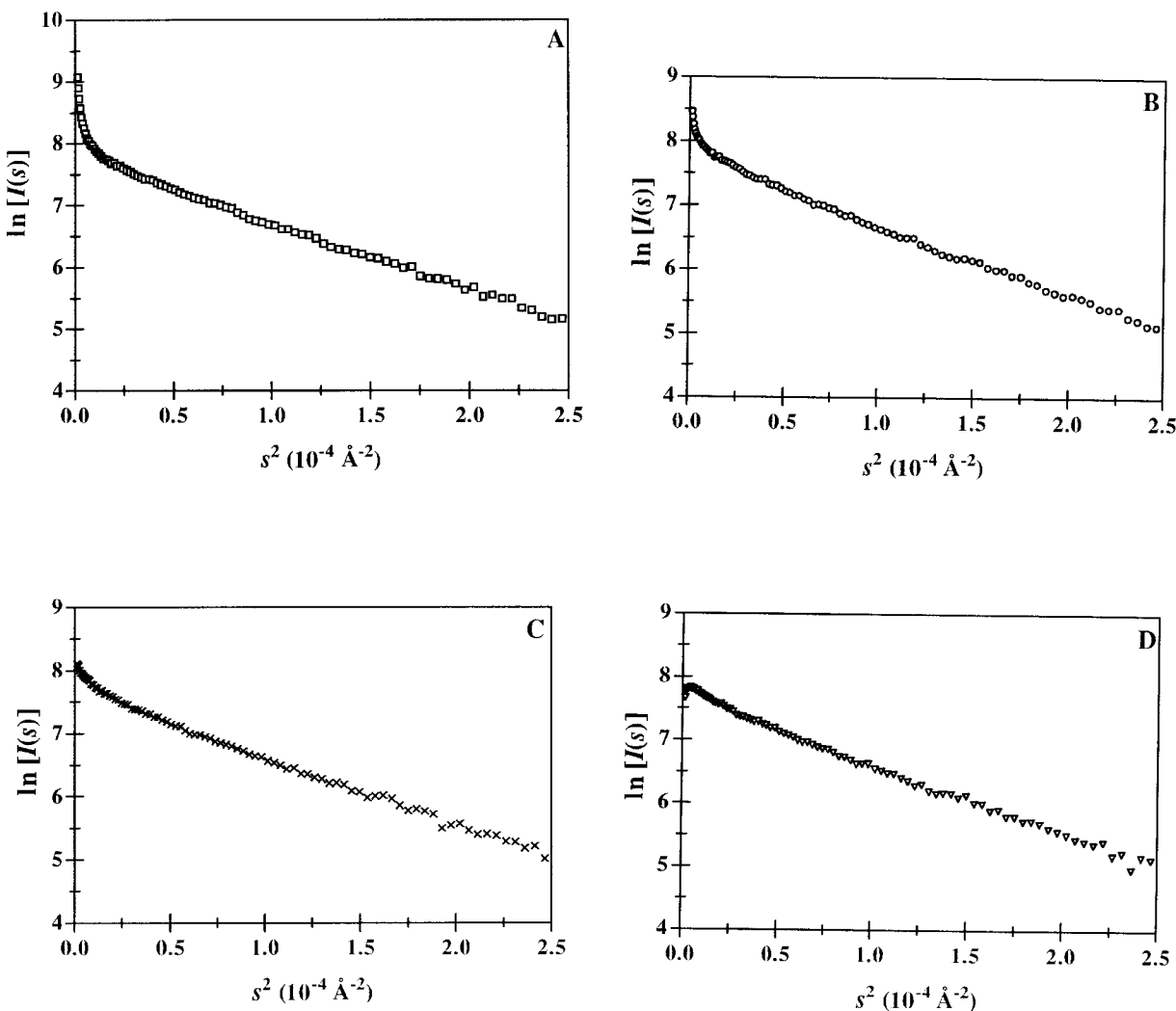


Fig. 2. Guinier plots of 5 mg/ml arcelin-1 solutions in (A) 0.1 M MES pH 4.9 ( $I = 0.1$  M), and in the same buffer supplemented with (B) 0.3 M sodium thiocyanate ( $I = 0.4$  M). (C) 0.3 M sodium chloride ( $I = 0.4$  M). (D) 0.3 M sodium citrate ( $I = 0.5$  M).

citrate buffer alone ( $I = 0.15$  M) and in this buffer supplemented with 300 mM sodium thiocyanate ( $I = 0.45$  M) or 300 mM ammonium sulfate ( $I = 1.05$  M). The nearly monodispersed character of the protein solution in citrate buffer (Fig. 4A) was slightly altered in the presence of the chaotropic thiocyanate ions (Fig. 4B), even though the ionic strength was increased. On the contrary, sulfate ions had no significant effect (Fig. 4C).

The Guinier plots for arcelin-1, computed for all monodispersed conditions, gave a similar radius of gyration. The average  $R_g$  value is  $29.0 \pm 0.3$  Å. The extrapolated intensities  $I(0)$  were brought to absolute scale by measuring the small-angle X-ray scattering of a 3-phosphoglycerate kinase solution of known concentration. The calculated average molecular weight of the arcelin-1 scattering species was  $60.1 \pm 3.1$  kDa, which according to the molecular

weight of the monomer from SDS-PAGE electrophoresis, indicated that arcelin-1 exists as dimeric species in solution.

### Crystallization

Thin plates were obtained from a large screening of the crystallization parameters, among which acidic pH and PEG400 appeared to be favorable for nucleation. The crystallization media were then chosen to also take into account the requirements for monodispersed protein solutions. They were conducted in the presence of 300 mM ammonium sulfate, added to three different buffers: 100 mM MES, 100 mM sodium acetate, or 100 mM sodium citrate. The first and the last conditions correspond to the SAXS experiments shown in Figures 3C and 4C. In all three conditions, the ionic strength and the pH were similar ( $I \approx 1.0$  M,  $\text{pH} \approx 4.9$ ). Sulfate ions contrib-

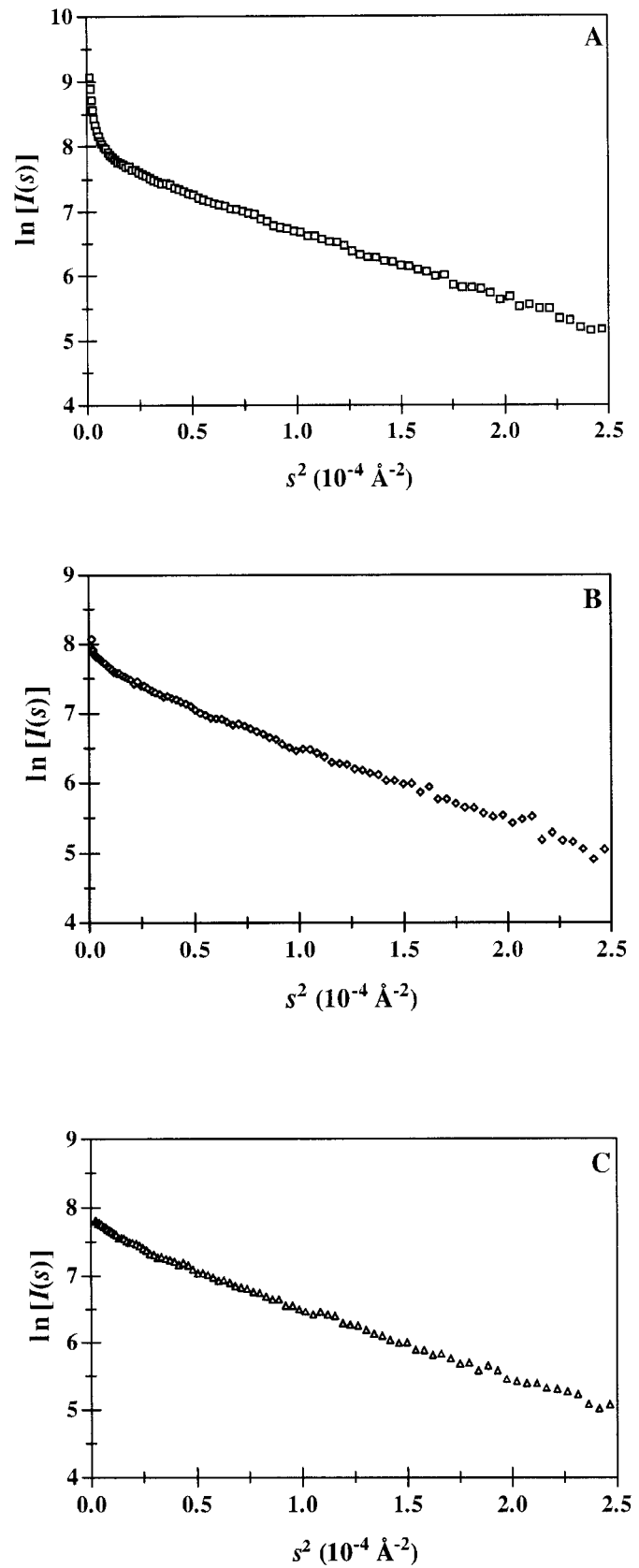


Fig. 3. Guinier plots of 5 mg/ml arcelin-1 solutions in (A) 0.1 M MES pH = 4.9 and in the same buffer supplemented with (B) 1.0 M sodium thiocyanate ( $I = 1.1$  M). (C) 0.3 M ammonium sulfate ( $I = 1.0$  M).

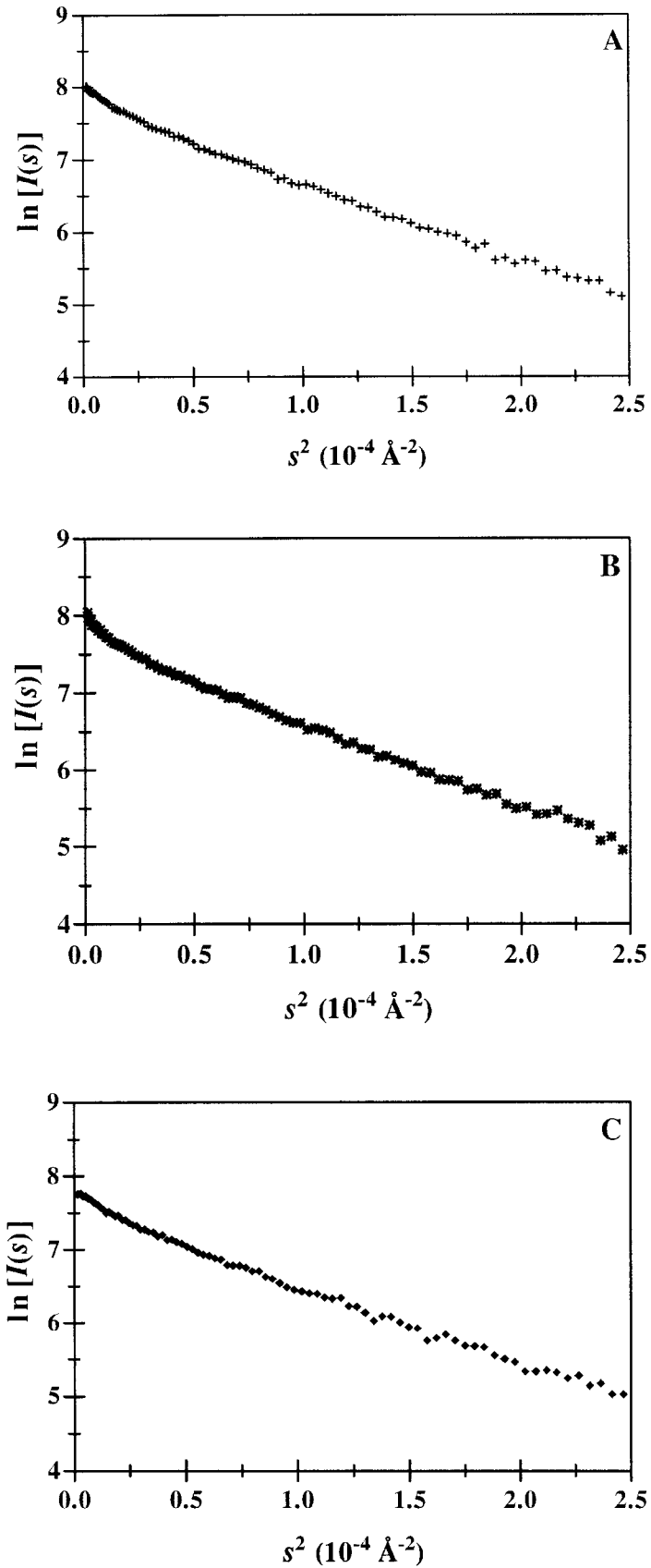


Fig. 4. Guinier plots of 5 mg/ml arcelin-1 solutions in (A) 0.1 M sodium citrate pH 5.0 ( $I = 0.15$  M) and in the same buffer supplemented with (B) 0.3 M sodium thiocyanate ( $I = 0.45$  M). (C) 0.3 M ammonium sulfate ( $I = 1.05$  M).



uted as both ionic strength enhancer and as the most lyotropic agent. In the presence of citrate buffer, the protein solution remained clear during the whole crystallization process, and crystals appeared within 1 week. Skin formation was observed in the case of acetate buffer, and crystals were obtained after 1 month. However, in these two conditions, polynucleation was substantial, and the crystals were often twinned. By using MES buffer, single crystals appeared after some precipitation of the protein, and 2 months were needed to reach equilibrium. It thus appeared that acetate and citrate ions increasingly counteract the effectiveness of sulfate ions in lowering protein solubility.

### Crystallographic Data and Molecular Replacement

Crystals belong to the orthorhombic space group  $P2_12_12$  with cell parameters  $a = 85.6 \text{ \AA}$ ,  $b = 92.6 \text{ \AA}$ ,  $c = 67.3 \text{ \AA}$ . The asymmetric unit contains two molecules of arcelin-1, which gave a calculated crystal volume per unit of molecular weight  $V_M$  of  $2.2 \text{ \AA}^3/\text{Da}$  and a solvent content of 44%. These values are in the range generally found for protein crystals.<sup>29</sup> Diffraction data were collected to 1.9 Å resolution from a single crystal of arcelin-1 (Table II).

Molecular replacement was performed by using one promoter of isolectin I (LoLI) as a model. The sequence identity between arcelin-1 and LoLI is 45% (Fig. 1). The cross-rotation function, with an outer radius of integration of 25 Å, and the translation function gave the best results between 8.0 and 3.0 Å resolution. The highest peak of the rotation function had a value of 16.5 (the r.m.s. deviation from mean density is 2.2), which, when compared with the height of the second peak (10.0), indicated a possible solution. It indeed gave, after a translation search, the best correlation coefficient ( $C = 0.173$ ) and crystallographic  $R$ -factor ( $R_f = 0.685$ ), whereas other solutions of the cross-rotation function had correlation coefficients not higher than 0.132. The position of this molecule was fixed, and the search was pursued to locate a second molecule in the asymmetric unit. One solution was apparent according to the crystallographic criteria ( $C = 0.537$  and  $R_f = 0.529$ ). It should be mentioned that the unambiguous assignment of the crystal space group, among the eight possible ones in the orthorhombic system, was provided by the translation function. Final  $C$  and  $R_f$  with other space groups than  $P2_12_12$  were not better than 0.499 and 0.552, respectively. The two molecules in the asymmetric unit were found in the same orientation and related by the translation vector (0.5976, 0.4999, 0.5005). These results were in agreement with the Patterson map computed from the native data set, which displayed a strong peak (0.52 of the origin peak) at (0.5972, 0.5, 0.5), indicating pseudocentering of the primitive lattice.

**TABLE II. Data Processing Statistics for Native Arcelin-1 Crystal\***

	33.70–1.90 Å	1.95–1.90 Å
Number of measurements	126,314	7,731
Number of unique reflections	41,612	2,913
Completeness (%)	97.7	93.9
$R_{\text{merge}}$ (%)	6.4	15.7
$\langle I/\text{sd}(I) \rangle$	11.9	7.5

\*Values are given for the entire resolution range and for the highest resolution shell.

A rigid body fitting dropped the  $R_f$  value to 0.494 in the 8.0–3.0 Å resolution range and increased  $C$  to 0.624. Crystal packing showed no bad contact between crystallographically and noncrystallographically related molecules (Fig. 5).

### DISCUSSION

Crystallization of glycoproteins is sometimes difficult to achieve due to the heterogeneity and the conformational flexibility of their carbohydrate moieties. Despite a total carbohydrate content of 10% (w/w), single crystals of arcelin-1 diffracting to 1.9 Å resolution were obtained when taking into account the results of the physicochemical investigations on the protein solution.

SAXS is well suited to investigate the macromolecular organization and the properties of proteins in solution,<sup>30</sup> which may, in turn, provide insights into their function.<sup>31</sup> The concentrations of arcelin-1 used in these experiments (5 mg/ml, 0.1 mM) are close to the initial protein concentration in the crystallization trials and may not be unrealistic with respect to the situation in vivo, because arcelin-1 is expressed at a high level in the seeds.<sup>17</sup> SAXS experiments showed that arcelin-1, at low ionic strength in zwitterionic buffer, and in undersaturating conditions, was prone to aggregation. This behavior might be surprising in view of the accepted assumption that electrostatic interactions dominate, at low ionic strength, for protein bearing a net charge different from zero. Indeed, according to the calculation made from the number and the type of charged residues,<sup>32</sup> arcelin-1 ( $\text{pI}^{\text{exp}} = 6.8$ ,  $\text{pI}^{\text{calc}} = 6.3$ ) bears nine positive charges at pH 4.9, and attractive protein-protein interactions should not be expected. At the protein concentration used in these experiments, the form factor dominates the scattered intensities, and the curvature near the origin seems best explained by the presence of oligomers. Their formation could be mediated by the glycan chains, because arcelin-1 also displays complex-glycan binding properties, or due to nonspecific electrostatic interactions between clusters of positive and negative regions at the protein surface, if the charged residues are not randomly distributed. Monodispersity of the arcelin-1 solution, which seems to be a prerequisite to crystallization,<sup>33,34</sup> was dependent on two param-



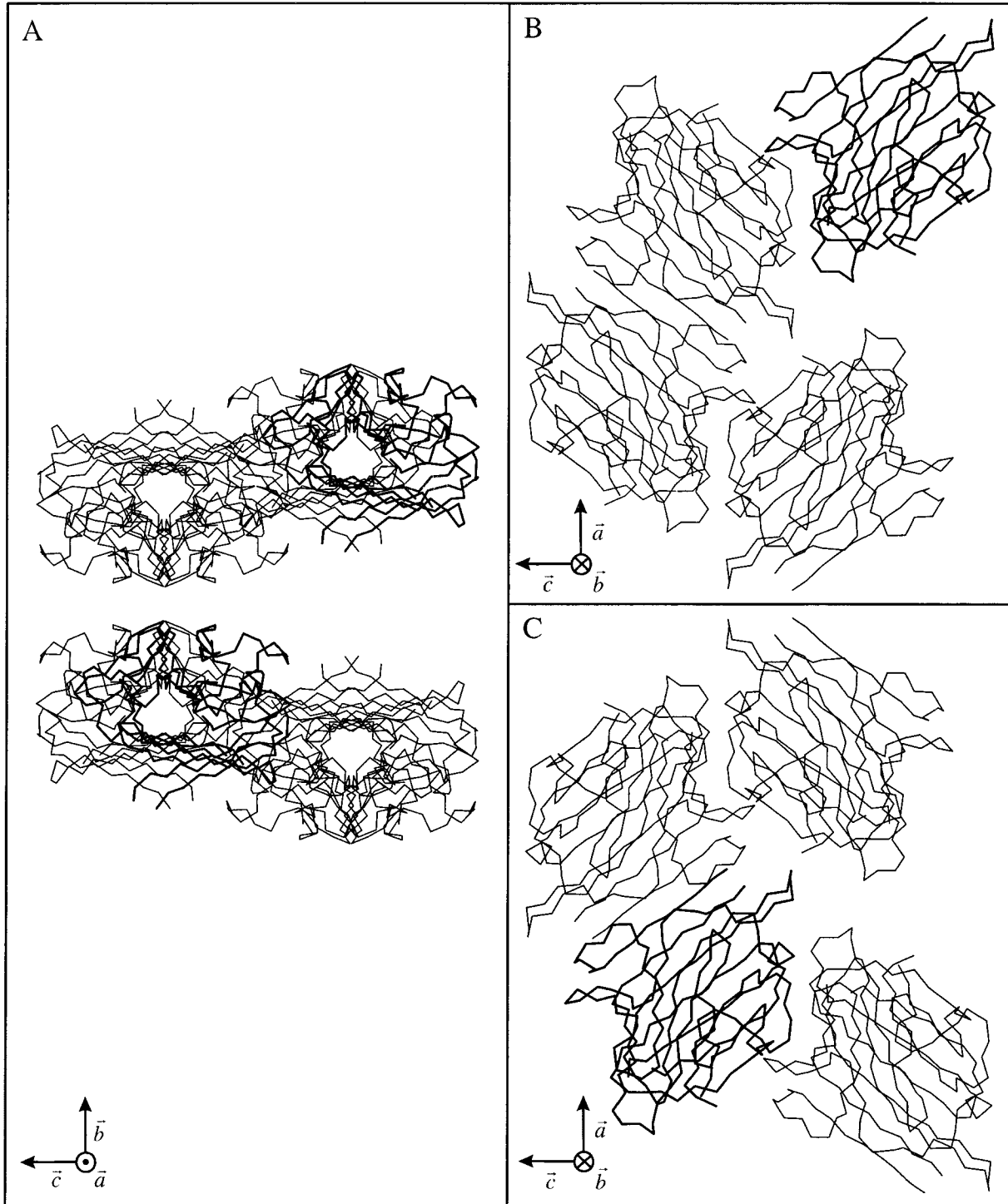


Fig. 5. Crystal packing of arcelin-1. **A**: The eight molecules in the unit cell are shown with the crystallographic  $b$  and  $c$  axes running vertically and horizontally in the plane of the figure, respectively. Each motif corresponds to four arcelin-1 molecules.

**B,C**: Orthogonal view of each of these motifs, looking down the  $b$  axis. The lectin-type dimers are now apparent. The two molecules that were found by molecular replacement to be related by a translation vector are displayed with thicker lines.

eters. The first one was the ionic strength, which likely contributes through the nonspecific screening effect of anions on the excess of positive charges.<sup>35</sup>

The second parameter was the type of anions. Their effectiveness against aggregation followed the Hofmeister series (sulfate  $\approx$  citrate  $>$  chloride  $>$

thiocyanate). Interestingly, the effectiveness of anions to lower arcelin-1 solubility (sulfate > acetate > citrate) also followed this classification.

The aggregation tendency of arcelin-1 and the effectiveness of anions with respect to protein solubility, are opposite to what has been described for other positively charged proteins.<sup>32</sup> With lysozyme, it was shown that the presence of 13 positive charges induced repulsive protein-protein electrostatic interactions in undersaturating and low ionic strength conditions and that the effectiveness of anions to lower protein solubility followed the reverse order of the Hofmeister series.<sup>30</sup> However, the physicochemical properties of arcelin-1 might be explained by the presence of a surface-exposed anion binding site and preferential protein-sulfate interactions. Anion binding would disrupt aggregation by neutralizing this specific positively charged region, but the binding efficiency and thus the effect on aggregation, will depend on the geometry and the number of charges on the anion. Citrate and sulfate ions seem most efficient in that respect. By decreasing the positive charge of the protein, anion binding would decrease its solubility. In that respect, sulfate seems much more efficient than citrate. Surface plasmon resonance experiments could not provide firm conclusions with regard to protein-sulfate specific interactions. However, a  $(3F_{\text{obs}} - 2F_{\text{calc}}) \exp(i\alpha_{\text{calc}})$  map, computed in the course of structure refinement, shows the typical tetrahedral electron density of a sulfate ion, in the immediate vicinity of an arginine and a histidine residue at the protein surface. This may explain why acidic pH is required for crystallization. This sulfate ion also defines a packing interface between two arcelin-1 molecules. Citrate, acetate (and thiocyanate) ions binding to this preferential sulfate binding site would influence crystal packing and may explain the formation of twinned crystals observed in these conditions.

Two molecules of arcelin-1, related by a pure translation vector, were found from molecular replacement. However, examination of the crystal packing allowed a second description of noncrystallographically related molecules leading to a compact dimer. In this case, the two molecules in the asymmetric unit are related by a noncrystallographic twofold axis lying parallel to the *b* crystallographic axis (Fig. 5B,C). This dimer closely resembles the canonical dimeric lectins organization in which the two monomers are associated to form a dimerwide  $\beta$ -sheet. This dimer likely corresponds to the molecular species observed in SAXS experiments. The experimental  $R_g$  value for arcelin-1 (29.0 Å) was higher than that calculated from the crystallographic coordinates of LoLI (24.9 Å) using the program CRY SOL.<sup>36</sup> This may arise from the contribution of the two or three glycan chains linked to each arcelin-1 monomer (P. Rougé, unpublished data), which have no

equivalent in the legume lectin used for the comparison.

The dimeric structure of arcelin-1 has implications for its biological function. The presence of oligomeric forms is a prerequisite for the bridging properties associated with the hemagglutinating activity of the protein.<sup>37</sup> Surprisingly, arcelin-5, which displays 62% sequence identity to arcelin-1, crystallized as a monomer, and crystal packing did not reveal dimeric species.<sup>38</sup> Such observations are puzzling with respect to the previously reported hemagglutination properties of this protein.<sup>39</sup> The 1.9 Å resolution three-dimensional structure of arcelin-1 should provide a better understanding of the insecticidal properties of this protein and will help to relate the molecular features of arcelin-1,  $\alpha$ -amylase inhibitor, and PHA-L to their different biochemical properties.

#### ACKNOWLEDGMENTS

We thank Roger Fourme and Patrice Vachette (LURE, Orsay, France) for excellent data collection facilities. We are grateful to Joël Janin for providing us with pure 3-phosphoglycerate kinase. We also thank Martin Welch for critical reading of the manuscript.

#### REFERENCES

- Goldstein, I.J., Poretz, R.D. Isolation, physicochemical characterization and carbohydrate-binding specificity of lectins. In: "The Lectins: Properties, Functions and Applications in Biology and Medicine." Liener, I.E., Sharon, N., Goldstein, I.J. (eds.). New York: Academic Press Inc., 1986:33–247.
- Weis, W.I., Drickamer, K. Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* 65:441–473, 1996.
- Rougé, P., Cambillau, C., Bourne, Y. The three-dimensional structure of legume lectins. In: "Lectin Reviews." Kilpatrick, D.C., Van Driessche, E., Bog-Hansen, T.C. (eds.). Vol. 1. St. Louis, MO: Sigma Chemical Company, 1991:143–159.
- Einspahr, H., Parks, E.H., Suguna, K., Subramanian, E., Suddath, F.L. The crystal structure of pea lectin at 3.0 Å resolution. *J. Biol. Chem.* 261:16518–16527, 1986.
- Bourne, Y., Abergel, C., Cambillau, C., Frey, M., Rougé, P., Fontecilla-Camps, J.C. X-Ray crystal structure determination and refinement at 1.9 angstroms resolution of isolectin I from the seeds of *Lathyrus ochrus*. *J. Mol. Biol.* 214:571–584, 1990.
- Bourne, Y., Mazurier, J., Legrand, D., Rougé, P., Montreuil, J., Spik, G., Cambillau, C. Structures of a legume lectin complexed with the human lactotransferrin N2 fragment, and with an isolated biantennary glycopeptide: Role of the fucose moiety. *Structure* 2:209–219, 1994.
- Shaanan, B., Lis, H., Sharon, N. Structure of a legume lectin with an ordered N-linked carbohydrate in complex with lactose. *Science* 254:862–866, 1991.
- Delbaere, L.T.J., Vandonselaar, M., Prasad, L., Quail, J.W., Wilson, K.S., Dauter, Z. Structures of the lectin IV of *Griffonia simplicifolia* and its complex with the Lewis B human blood group determinant at 2.0 angstroms resolution. *J. Mol. Biol.* 230:950–965, 1993.
- Loris, R., Steyaert, J., Maes, D., Lisgarten, J., Pickersgill, R., Wyns, L. Crystal structure determination and refinement at 2.3 angstroms resolution of the lentil lectin. *Biochemistry* 32:8772–8781, 1993.
- Reeke, G.N., Jr., Becker, J.W., Edelman, G.M. The covalent and three-dimensional structure of concanavalin A. IV.

- Atomic coordinates, hydrogen bonding, and quaternary structure. *J. Biol. Chem.* 250:1525–1547, 1975.
11. Banerjee, R., Mande, S.C., Ganesh, V., Das, K., Dhanaraj, V., Manhanta, S.K., Suguna, K., Suroolia, A., Vijayan, M. Crystal structure of peanut lectin, a protein with an unusual quaternary structure. *Proc. Natl. Acad. Sci. USA* 91:227–231, 1994.
  12. Dessen, A., Gupta, D., Sabesan, S., Brewer, C.F., Sacchettini, J.C. X-ray crystal structure of the soybean agglutinin cross-linked with a biantennary analog of the blood group I carbohydrate antigen. *Biochemistry* 34:4933–4942, 1995.
  13. Hamelryck, T.W., Dao-Thi, M.H., Poortmans, F., Chrispeels, M.J., Wyns, L., Loris, R. The crystallographic structure of phytohemagglutinin-L. *J. Biol. Chem.* 271:20479–20485, 1996.
  14. Chrispeels, M.J., Raikhel, N.V. Lectins, lectin genes, and their role in plant defense. *Plant Cell* 3:1–9, 1991.
  15. Rougé, P., Barre, A., Causse, H., Chatelain, C., Porté, G. Arcelin and  $\alpha$ -amylase inhibitor from the seeds of common bean (*Phaseolus vulgaris* L.) are truncated lectins. *Biochem. Syst. Ecol.* 21:695–703, 1993.
  16. Bompard-Gilles, C., Rousseau, P., Rougé, P., Payan, F. Substrate mimicry in the active center of a mammalian  $\alpha$ -amylase: Structural analysis of an enzyme-inhibitor complex. *Structure* 4:1441–1452, 1996.
  17. Osborn, T.C., Alexander, D.C., Sun, S.S.M., Cardona, C., Bliss, F.A. Insecticidal activity and lectin homology of arcelin seed protein. *Science* 240:207–210, 1988.
  18. Dawson, R.M.C., Elliott, D.C., Elliott, W.H., Jones, K.M. In: "Data for Biochemical Research." Oxford: Oxford University Press, 1986.
  19. Depautex, C., Desvignes, C., Leboucher, P., Lemonnier, M., Dagneaux, D., Benoit, J.P., Vachette, P. LURE Annual report 1985–1987. Doc. CEN Saclay, 1987.
  20. Bordas, J., Koch, M.H.J., Clout, P.N., Dorrington, E., Boulin, C., Gabriel, A. A synchrotron radiation camera and data acquisition system for time resolved x-ray scattering studies. *J. Phys., E: Sci. Instrum.* 13:938–944, 1980.
  21. Dubuisson, J.M., Decamps, T., Vachette, P. Improved signal-to-background ratio in small-angle X-ray scattering experiments with synchrotron radiation using an evacuated cell for solutions. *J. Appl. Crystallogr.* 30:49–54, 1997.
  22. Guinier, A., Fournet, G. In: "Small Angle X-ray Scattering." New York: Wiley, 1955.
  23. Leslie, A.G.W. Recent changes to the MOSFLM package for processing film and image plate data. CCP4 and ESF-EACMB Newsletter on Protein Crystallography, Number 26, 1992.
  24. Collaborative Computational Project 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* 50:760–763, 1994.
  25. Navaza, J. AMoRe: An automated package for molecular replacement. *Acta Crystallogr. A* 50:157–163, 1994.
  26. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr, Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., Tasumi, M. The protein data bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542, 1977.
  27. von Hippel, H., Schleich, T. The effects of neutral salts on the structure and conformational stability of macromolecules in solution. In: "Structure and Stability of Biological Macromolecules." Timasheff, S.N., Fashman, G.D. (eds.). New York: Dekker, 1969:417–574.
  28. Collins, K.D., Washabaugh, M.W. The Hofmeister effect and the behaviour of water at interfaces. *Q. Rev. Biophys.* 18:323–422, 1985.
  29. Matthews, B.W. Solvent content of protein crystals. *J. Mol. Biol.* 33:491–497, 1968.
  30. Ducruix, A., Guilloteau, J.P., Riès-Kautt, M., Tardieu, A. Protein interactions as seen by solution X-ray scattering prior to crystallization. *J. Crystal Growth* 168:28–39, 1996.
  31. Birck, C., Vachette, P., Welch, M., Swarén, P., Samama, J.P. Is the function of the cdc2 kinase subunit proteins tuned by their propensities to oligomerize? Conformational states in solution of the cdc2 kinase partners p13<sup>suc1</sup> and p9<sup>ckspby</sup>. *Biochemistry* 35:5577–5585, 1996.
  32. Riès-Kautt, M., Ducruix, A. Inferences drawn from physicochemical studies of crystallization and precrystalline state. *Methods Enzymol.* 276:23–59, 1997.
  33. Zulauf, M., D'Arcy, A. Light scattering of proteins as a criterion for crystallization. *J. Crystal Growth* 122:102–106, 1992.
  34. Veessler, S., Marcq, S., Lafont, S., Astier, J.P., Boistelle, R. Influence of polydispersity on protein crystallization: A quasi-elastic light scattering study applied to  $\alpha$ -amylase. *Acta Crystallogr. D* 50:355–360, 1994.
  35. Melander, W., Horváth, C. Salt effects on hydrophobic interactions in the precipitation and chromatography of proteins: An interpretation of the lyotropic series. *Arch. Biochem. Biophys.* 183:200–215, 1977.
  36. Svergun, D., Barberato, C., Koch, M.H.J. CRYSOLE—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* 28:768–773, 1995.
  37. Hartweck, L.M., Vogelzang, R.D., Osborn, T.C. Characterization and comparison of arcelin seed protein variants from common bean. *Plant Physiol.* 97:204–211, 1991.
  38. Hamelryck, T.W., Poortmans, F., Goossens, A., Angenon, G., Van Montagu, M., Wyns, L., Loris, R. Crystal structure of arcelin-5, a lectin-like defense protein from *Phaseolus vulgaris*. *J. Biol. Chem.* 271:32796–32802, 1997.
  39. Goossens, A., Geremia, R., Bauw, G., Van Montagu, M., Angenon, G. Isolation and characterisation of arcelin-5 proteins and cDNAs. *Eur. J. Biochem.* 225:787–795, 1994.