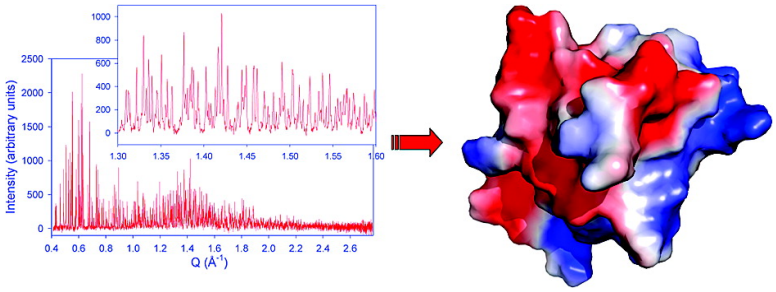
Article

Second SH3 Domain of Ponsin Solved from Powder Diffraction

Irene Margiolaki, Jonathan P. Wright, Matthias Wilmanns, Andrew N. Fitch, and Nikos Pinotsis

J. Am. Chem. Soc., **2007**, 129 (38), 11865-11871• DOI: 10.1021/ja073846c • Publication Date (Web): 05 September 2007



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 09/05/2007

Second SH3 Domain of Ponsin Solved from Powder Diffraction

Irene Margiolaki,*,† Jonathan P. Wright,† Matthias Wilmanns,‡ Andrew N. Fitch,† and Nikos Pinotsis*,‡

Contribution from the European Synchrotron Radiation Facility, ESRF, BP-220, F-38043, Grenoble, France and European Molecular Biology Laboratory, EMBL-Hamburg c/o DESY, Notkeststrasse 85, D-22603 Hamburg, Germany

Received June 11, 2007; E-mail: margiolaki@esrf.fr, pinotsis@embl-hamburg.de

Abstract: Determination of protein crystal structures is dependent on the growth of high-quality single crystals, a process that is not always successful. Optimum crystallization conditions must be systematically sought for, and microcrystalline powders are frequently obtained in failed attempts to grow the desired crystal. In materials science, structures of samples ranging from ceramics, pharmaceuticals, zeolites, etc., can nowadays be solved, almost routinely, from powdered samples, and there seems to be no fundamental reason, except the sheer size and complexity of the structures involved, why powder diffraction should not be employed to solve structures of small proteins. Indeed, recent work has shown that the high-quality powder diffraction data can be used in the study of protein crystal structures. We report the solution, model building, and refinement of a 67-residue protein domain crystal structure, with a cell volume of 64 879 Å^3, from powder diffraction. The second SH3 domain of ponsin, a protein of high biological significance due to its role in cellular processes, is determined and refined to resolution limits comparable to single-crystal techniques. Our results demonstrate the power and future applicability of the powder technique in structural biology.

Introduction

Crystal structures of proteins are primarily studied using single-crystal X-ray diffraction. Nevertheless, growth of suitable crystals remains a challenging task. A single crystal is usually obtained after optimization of crystallization protocols which initially lead to microcrystalline precipitants. Powder diffraction has long been used for characterization of a wide range of materials including pharmaceutical compounds,[1,2] zeolites,[3] and large metal−organic frameworks (MOF's).[4] The major drawback is that structure determination becomes greatly complicated because of the collapse of the three-dimensional diffraction pattern onto a single dimension as a result of the orientational averaging of the powder grains.

Current progress in powder diffraction data collection and interpretation has allowed the study of several reference proteins where a good molecular starting model of the same protein was available. Reported results include structure solution via the molecular replacement method for a new hexagonal phase of insulin[5] and lysozyme from different sources,[6,7] refinement via stereochemically restrained Rietveld analysis,[8] indicating a significant improvement in the refinement procedure when multiple powder profiles are used.[6,9] Moreover, detection of ligands in protein−ligand complexes has been accomplished by difference Fourier methods.[10]

We report the structure solution, model building, and refinement of a biologically important protein domain by means of high-angular-resolution powder diffraction. We collected powder diffraction profiles from microcrystalline samples of the second SH3 (Src homology -3) domain of ponsin (SH3.2). The anisotropic changes in unit-cell parameters induced by radiation damage were exploited to lessen—to some extent—the deleterious effects of peak overlap, extending previously reported methods.[6,9,11,12] The extracted intensities were of sufficient accuracy to employ established single-crystal molecular replacement methods from a model with about 40% sequence similarity. Maximum likelihood refinement improved the phases to a level where we could trace the main-chain alterations, build additional residues where needed, and eventually place the correct side

† European Synchrotron Radiation Facility.
‡ European Molecular Biology Laboratory.

(1) David, W. I. F.; Shankland, K.; Shankland, N. *Chem. Commun.* **1998**, 931−932.
(2) Harris, K. D. M.; Tremayne, M. *Chem. Mater.* **1996**, *8*, 2554−2570.
(3) Baerlocher, C.; Gramm, F.; Massuger, L.; McCusker, L. B.; He, Z. B.; Hovmoller, S.; Zou, X. D. *Science* **2007**, *315*, 1113−1116.
(4) Ferey, G.; Merlot-Draznieks, C.; Serre, C.; Millange, F.; Dutour, J.; Surble, S.; Margiolaki, I. *Science* **2005**, *309*, 2040−2042.
(5) Von Dreele, R. B.; Stephens, P. W.; Smith, G. D.; Blessing, R. H. *Acta Crystallogr.* **2000**, *D56*, 1549−1553.
(6) Basso, S.; Fitch, A. N.; Fox, G. C.; Margiolaki, I.; Wright, J. P. *Acta Crystallogr.* **2005**, *D61*, 1612−1625.
(7) Margiolaki, I.; Wright, J. P.; Fitch, A. N.; Fox, G. C.; Von Dreele, R. B. *Acta Crystallogr.* **2005**, *D61*, 423−432.
(8) Von Dreele, R. B. *Methods Enzymol.* **2003**, *368*, 254−267.
(9) Von Dreele, R. B. J. *Appl. Crystallogr.* **2007**, *40*, 133−143.
(10) Von Dreele, R. B. *Acta Crystallogr.* **2005**, *D61*, 22−32.
(11) (a) Zachariasen, W. H.; Ellinger, F. H. *Acta Crystallogr.* **1963**, 369−375. (b) Shankland, K.; David, W. I. F.; Sivia, D. S. J. *Mater. Chem.* **1997**, *7*, 569−572. (c) Brunelli, M.; Wright, J. P.; Vaughan, G. R. M.; Mora, A. J.; Fitch, A. N. *Angew. Chem. Int. Ed.* **2003**, *42* (18), 2029−2032.
(12) Wright, J. P.; Markvardsen, A. J.; Margiolaki, I. *Z. Kristallogr.* **2007**, in press.

chains along the sequence. We further exploited an approach combining multiple-data-set Rietveld analysis and periodical OMIT map computation to reduce the bias of the final model, extend the resolution limits to levels comparable to single-crystal measurements, and even detect several water molecules bound to the protein.

Since the primary limitation of the powder diffraction technique is the complexity of the system, the 67-residue SH3 motif is an appropriate protein domain to study. Protein fragments at the size of an SH3 domain are widespread in a large number of modular proteins, involved in highly specific and reversible protein−protein signaling processes. They are essential for functions like cell growth, differentiation, motility, and apoptosis.[13,14] The second SH3 domain of ponsin (SH3.2) binds to proline-rich motifs of the cytoskeletal proteins paxillin and vinculin at the cell−extracellular matrix adhesion sites,[15] while specifically the interaction to paxillin is also linked to muscle differentiation processes forming the lateral cell−matrix contacts of cross-striated muscle cells called costamers.[16]

Experimental Section

Protein Expression and Purification. The DNA sequence coding for PS-FL (accession no: AM260536) residues 824−884 corresponding to the second SH3 domain (SH3.2) was amplified by PCR using as template the construct contained in the pMypG vector[17] and primers suitable for cloning into the pET151/D-TOPO vector (Invitrogen). The vector overexpresses additionally an N-terminal His-tag sequence, a V5 epitope sequence, and a TEV protease cleavage site and is specifically designed for directional blunt-end cloning. For that reason the overexpressed product contains six non-native amino acids (GIDPFT) after the TEV Cleavage site. The recombinant plasmid was used to transform *E. coli* BL21(DE3) cells. The cell culture was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at an OD600 of approximately 0.6 at 310 K. Following induction, the culture was incubated for about 6 h. The cells were harvested with a lysis buffer containing 25 mM Tris pH 7.8, 300 mM NaCl, 10 mM imidazole, and 2 mM β-mercapto-ethanol and lysed by sonication. The cell debris was pelleted by centrifugation for 45 min at 277 K and 20 000 rev min^{-1}. The crude lysate was filtrated through a 0.22 μm membrane and loaded onto a Ni^{2+}-NTA column (QIAGEN). The protein was eluted with the lysis buffer containing 400 mM imidazole and subsequently dialyzed overnight by adding TEV-protease. The cleaved protein fragment was separated using the same Ni^{2+}-NTA column as before and then loaded into a gel-filtration column (Superdex 75, 16/60, Amersham Pharmacia Biotech) running on a 25 mM Tris pH 7.8, 150 mM NaCl, 2 mM EDTA buffer. The protein eluted with an apparent molecular weight of approximately 8 kDa, which is consistent with a monomer. The peak fractions were analyzed by SDS−PAGE to assess the chemical purity and pooled, and then dynamic light scattering was used to assess the conformational purity of the protein.

Crystallization. After purification and concentration, the SH3.2 domain spontaneously formed a microcrystalline material after about 12 h. This material was suitable only for powder diffraction measurements, and therefore, two batches of protein were crystallized in a similar manner. In order to verify the structural information obtained from the microcrystalline samples, a single crystal of the same domain was desired. In order to obtain a suitable single crystal, immediately after purification using a buffer of lower ionic strength, the protein was concentrated at about 10 mg mL^{-1} and set for crystallization at the high-throughput crystallization facility of the EMBL-Hamburg Outstation.[18] The Index kit from Hampton (http://www.hamptonresearch.com/) was employed, and the highest quality single crystal of dimensions 300 × 110 × 90 μm was obtained at condition 36 (15% tacsimate, pH 7.0, 0.1 M HEPES pH 7.0, 2% w/v polyethylene glycol 3,350) by mixing equal volumes of 300 nL of precipitant and protein using the sitting drop vapor diffusion technique. Additional diffraction measurements verified the identical space group and unit cell dimensions for both powder and single-crystal samples.

Table 1. Data Collection and Refinement Statistics of the Single-Crystal SH3.2 Model

crystal data	
space group	$P2_12_12_1$
unit cell parameters (Å)	$a = 24.661$, $b = 35.499$, $c = 71.079$
solvent content (%)	38.54
mosaicity (deg)	0.75
data collection	
beam line	X11, EMBL/DESY
wavelength (Å)	0.8125
resolution range	30.0−1.49
R_{sym} (%)[a]	6.5 (22.9)
$I/\sigma(I)$[a]	23.1 (7.9)
reflns measd/unique	24,9775/10,707
redundancy	5.9
completeness (%)[a]	99.6 (100.0)
refinement	
program	REFMAC5
R/R_{free} (%)	13.1/16.2
reflns (work/test sets)	10,124/311
no. of protein atoms	554
no. of water molecules	101
no. of ligands	2 formate
rmsd bonds (Å)	0.017
rmsd angles (deg.)	1.611
model quality (Ramachandran plot)	
most favored (%)	94.5
additionally allowed (%)	5.5

[a] Higher resolution cell values in parenthesis.

Synchrotron X-ray Powder Diffraction Measurements. For the large unit cells typical of proteins, the choice of the instrumental configuration is a key issue.[19] The best-quality diffraction profiles, in terms of minimum peak widths to reduce peak overlap, data resolution (d-spacing), and signal-to-noise ratio, were obtained at the high-resolution powder diffraction beam line ID31[20] at the European Synchrotron Radiation Facility (ESRF, France). In order to improve the information content of the diffracted profiles, we performed a series of measurements on two batches of sample prepared in a similar manner (samples A and B), albeit with slightly different cell-parameter ratios. In addition, different wavelengths were employed to decrease radiation damage effects and optimize instrumental characteristics. Precisely, diffraction data were collected for sample A at a wavelength of 1.252481(32) Å. Additional measurements for sample B were performed at two wavelengths, 1.251209(40) and 0.8012034(76) Å, in order to observe possible wavelength dependence of radiation damage effects and achieve optimum data quality (angular and d-spacing resolution). The samples were contained in spinning glass capillaries, 1.5 mm in diameter, mounted on the axis of the diffractometer, and patterns were measured with a period of 2.0 min using a beam size of 1.5 mm^2 (1.5

(13) Li, S. S. *Biochem. J.* **2005**, *390*, 641−653.
(14) Ball, L. J.; Kuhne, R.; Schneider-Mergener, J.; Oschkinat, H. *Angew. Chem., Int. Ed.* **2005**, *44*, 2852−2869.
(15) Zhang, M.; Liu, J.; Cheng, A.; DeYoung, S. M.; Chen, X. W.; Dold, L. H.; Saltiel, A. R. *EMBO J.* **2006**, *25*, 5284−5293.
(16) Gehmlich, K.; Pinotsis, N.; Hayess, K.; Van der Ven, P. F. M.; Milting, H.; Banayosy, A. E.; Körfer, R.; Wilmanns, M.; Ehler, E.; Fürst, D. O. *J. Mol. Biol.* **2007**, in press.
(17) Salgia, R.; Li, J. L.; Lo, S. H.; Brunkhorst, B.; Kansas, G. S.; Sobhany, E. S.; Sun, Y. P.; Pisick, E.; Hallek, M.; Ernst, T.; Tantravahi, R.; Chen, L. B.; Griffin, J. D. *J. Biol. Chem.* **1995**, *270*, 5039−5047.
(18) Mueller-Dieckmann, J. *Acta Crystallogr.* **2006**, *D62*, 1446−1452.
(19) Margiolaki, I.; Wright, J. P.; Fitch, A. N.; Fox, G. C.; Labrador, A.; Von Dreele, R. B.; Miura, K.; Gozzo, F.; Schiltz, M.; Besnard, C.; Camus, F.; Pattison, P.; Beckers, D.; Degen, T. *Z. Kristallogr.* **2007**, in press.
(20) Fitch, A. N. *J. Res. Natl. Inst. Stand. Tech.* **2004**, *109*, 133−142.

Table 2. Data Collection and Refinement Statistics Using the Four High-Resolution Profiles Collected for Samples A and B

crystal data	data set 1	data set 2	data set 3	data set 4
sample	A	A	B	B
space group	$P2_12_12_1$			
lattice parameters (Å)	a =24.70420(9)		a = 24.79017(7)	
	b = 36.42638(14)		b = 36.35407(12)	
	c = 72.09804(26)		c = 72.22940(32)	
	data collection			
exposure time (min)	2	4	2	2
wavelength (Å)	1.252481(32)	1.252481(32)	1.251209(40)	0.8012034(76)
$\Delta 2\theta$ (deg)	0.00803(2)	0.00798(3)	0.00199(5)	0.00073(3)
	Rietveld refinement			
program	GSAS			
d-spacing range (Å)	15.27−2.273	15.27−2.273	15.26−2.79	15.3−2.864
N_{reflns}	3317	3348	1865	1719
$N_{\text{restraints}}$	1980	1980	1980	1980
N_{steps}	9100	10666	6154	4359
$N_{\text{obs}} = N_{\text{restraints}} + N_{\text{steps}}$ for all patterns: 32 259				
$N_{\text{parameters}}$	1767	1767	1767	1767
profile scale factors	2.503(7)	2.108(6)	2.198(11)	1.826(5)
R_{wp}(%)	3.90	3.69	3.69	3.94
R_{p}(%)	2.80	2.74	2.80	3.01
R_{F^2} (%)	21.53	24.01	28.06	14.02
*T*otal R factors for the 4 histograms: R_{wp} = 3.82%, R_{p} = 2.86%				
profile parameters (No. 5)[33]				
W	0.0041(2)	0.0038(2)	0.0085(4)	0.0057(2)
Y	7.698(6)	10.005(4)	10.79(2)	13.70(1)
Y_e	−2.27(1)	−3.12(1)	2.16(2)	−0.51(2)
D11	0	0.01937(13)	0.0650(2)	0.0650(2)
D22	0	0.00198(6)	−0.0172(7)	−0.0172(7)
D33	0	0.000734(13)	0.0040(2)	0.0040(2)
solvent scattering coefficients via Babinet's principle model				
A_s	4.978(13)	4.740(14)	3.916(21)	3.987(12)
B_s	1.128(8)	1.197(9)	1.096(2)	1.254(11)
no. of protein atoms	544			
no. of water molecules	36			
rmsd bonds (Å)	0.01			
rmsd angles (deg)	2.49			
model quality (Ramachandran plot)				
most favored (%)	96.4			
additionally allowed (%)	3.6			

mm horizontal × 1.0 mm vertical), photon flux on sample ∼3 × 10^{12} photons s^{-1}. Appreciable changes were observed in the data after several patterns had been collected. Precisely, significant changes in the lattice parameters accompanied by a gradual increase of peak broadening and significant loss of intensity are common characteristics of radiation damage effects. Figure S1 shows the gradual evolution of six diffraction profiles of sample B collected at room temperature and 0.8012034(76) Å wavelength. Accurate lattice parameters were extracted from these data sets via the LeBail method[21] using the TOPAS Academic software suite.[22] Figure S2 illustrates an anisotropic variation of the three orthorhombic lattice dimensions with increasing irradiation time ($\Delta a/a_f$ = 0.421%, $\Delta b/b_f$ = 0.042%, $\Delta c/c_f$ = 0.171%, $\Delta V/V_f$ = 0.633%). The life time of sample B in the beam was identical at both wavelengths chosen for data collection (∼2 min). In order to increase the counting statistics without compromising the data quality, the capillary was translated to give a fresh region of sample as soon as radiation damage effects were evident in the peak positions and widths. Identical scans were summed together. Our earlier work has shown that cryo-cooling of protein powders leads to microstrains and considerable broadening of the diffraction peaks,[23] analogous to the increased mosaicity observed in single-crystal studies.[24] Thus, experiments were performed at room temperature.

Additional Single-Crystal X-ray Diffraction Measurements and Refinement. Data were collected on the bending magnet beamline X11 (EMBL/DESY, Hamburg). The crystal was immersed into the mother liquid, containing additionally 20% [v/v] MPD, and flash cooled to 100 K using a nitrogen cryostream. The data set was integrated, scaled, and merged using the HKL suite.[25] The final refined model from the powder diffraction data was used directly as an input for the "automated model building" protocol in ARP/wARP[26] producing a model with R = 20.9% and R_{free} = 23.8%, implying the isomorphous content in both powder and single-crystal materials. The model was manually inspected in 'O'[27], and using in parallel iterative cycles of refinement in Refmac5[28] double side-chain conformations and solvent molecules were added where necessary. The final cycles of refinement were carried out implementing anisotropic temperature factor refinement for the protein atoms. Data collection and refinement statistics are summarized into Table 1.

Results and Discussion

Enhanced Intensity Extraction from Powder Diffraction Data. The SH3.2 structure belongs to the $P2_12_12_1$ space group

(21) Lebail, A.; Duroy, H.; Fourquet, J. L. *Mater. Res. Bull.* **1988**, *23*, 447−452.
(22) Coelho, A. TOPAS-Academic **2004**, http://members.optusnet.com.au/∼alancoelho.
(23) Jenner, M. J.; Wright, J. P.; Margiolaki, I.; Fitch, A. N. *J. Appl. Crystallogr.* **2007**, *40*, 121−124.
(24) Garman, E. *Acta Crystallogr.* **1999**, *D55*, 1641−1653.
(25) Otwinowski, Z.; Minor, W. *Macromol. Crystallogr.* **1997**, *A276*, 307−326.
(26) Perrakis, A.; Morris, R.; Lamzin, V. S. *Nat. Struct. Biol.* **1999**, *6*, 458−463.
(27) Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr.* **1991**, *A47*, 110−119.
(28) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr.* **1997**, *D53*, 240−255.

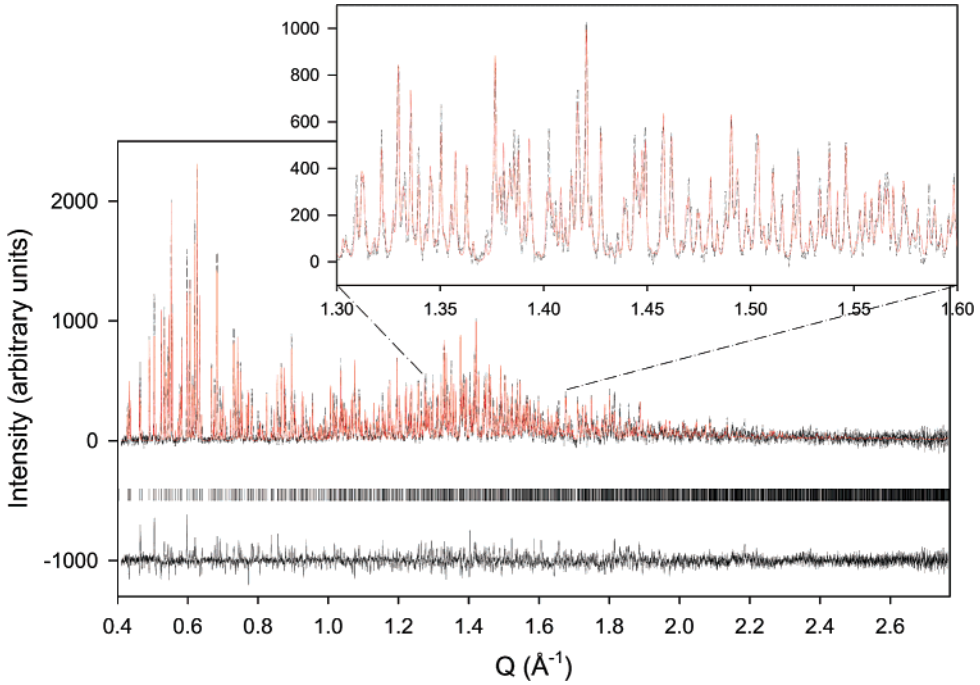


Figure 1. Final fit of one out of four data sets employed for stereochemically restrained Rietveld analysis. The data were collected on sample A at 295 K (ID31, $\lambda = 1.252481(32)$ Å). The dashed black, red, and lower black lines represent the experimental data, calculated pattern, and difference between experimental and calculated profiles, respectively. The vertical bars correspond to Bragg reflections compatible with the refined orthorhombic structural model. The inset corresponds to magnification of the observed and calculated profiles in the Q range between 1.3 and 1.6 Å^{-1}. The background intensity has been subtracted for clarity.

comprising one molecule per asymmetric unit. The same orthorhombic phase was identified for both microcrystalline samples employed for this study with small differences between their cell dimensions (Table 2). In addition, radiation-induced processes were evident from the gradual alteration of the diffraction peak positions associated with an anisotropic expansion of the three lattice parameters (Figures S1 and S2). Advantage has been taken of the anisotropic peak shifts in order to improve separation of adjacent peaks and improve the accuracy of the extracted peak intensities. Therefore, we selected four profiles with enhanced statistics collected on sample A for intensity extraction. These profiles were measured sequentially, and therefore, they correspond to different levels of radiation damage and slightly different lattice parameters. Intensities were extracted via a multipattern Pawley refinement where each diffraction profile is calculated as a sum of overlapping reflections, the intensities of which are variables in a least-squares procedure. The four data sets were fitted using the same integrated intensities for each pattern but different unit cell parameters (first data set: $a = 24.70420(9)$ Å, $b = 36.42638(14)$ Å, $c = 72.09804(26)$ Å). Reflections with a reasonable signal-to-noise ratio could be observed up to a d-spacing resolution of about 2.4 Å. This procedure allowed us to extract a set of 2764 intensities with improved effective completeness (Figure S3). The intensity values are in excellent agreement with those obtained subsequently from single-crystal diffraction experiments (~90% and 68% correlation at 5.0 and 3.7 Å resolution, respectively) (Figure S4).

Structure Solution via the Molecular Replacement Method. For structure determination, we randomly flagged the structure factor amplitudes for cross-validation using the CCP4 suite programs.[29] Molecular replacement was performed using both Molrep[30] (Figure S5) and Phaser[31] (data not shown). On the basis of the sequence similarity we selected two search models (PDB entries 1W70 and 1OOT) corresponding to 38% and 40% sequence homology, respectively, and a data resolution limit of 3.7 Å. Although identical results were evident for both cases, the correct solution was more clearly distinguished from the noise peaks when the 1W70 was employed (Figure S5).

Structure Refinement and Model Building. We chose the solution obtained from Molrep (model 1W70) as the initial model for maximum likelihood refinement in REFMAC5[28] using tight restraints and a resolution limit at 3.5 Å. The initial model provided a moderate fit to the powder data ($R = 43.5\%$ and $R_{free} = 48.1\%$) and electron density maps of sufficient quality to identify the correct protein sequence of the SH3.2 domain via "O"[27] (Figure S6). Further refinement cycles increasing the resolution limit to 2.9 Å enhanced the quality of the model and led to improved agreement factors ($R = 28.9\%$ and $R_{free} = 32.6\%$).

The model was further improved using the GSAS Rietveld[32] refinement suite.[33] In this case the diffraction profiles were included based on the observed signal-to-noise ratio, data resolution, and full width at half-maximum (fwhm) of the observed reflections. Two data sets for sample A (containing different levels of radiation damage) and two more for sample B (measured at two different wavelengths) were employed in a combined stereochemically restrained Rietveld refinement. This kind of restrained least-squares procedure has been previously applied successfully on reference proteins[5−9] and is described in detail elsewhere.[9] Effectively, use of several data sets was

(29) Bailey, S. *Acta Crystallogr.* **1994**, *D50*, 760−763.
(30) Vagin, A. A.; Teplyakov, A. *J. Appl. Crystallogr.* **1997**, *30*, 1022−1025.
(31) Read, R. J. *Acta Crystallogr.* **2001**, *D57*, 1373−1382.
(32) Rietveld, H. M. *J. Appl. Crystallogr.* **1969**, *2*, 65−71.
(33) Larson, A. C.; Von Dreele, R. B. General Structure Analysis System (GSAS); Los Alamos National Laboratory Report LAUR 86-748; 2004.

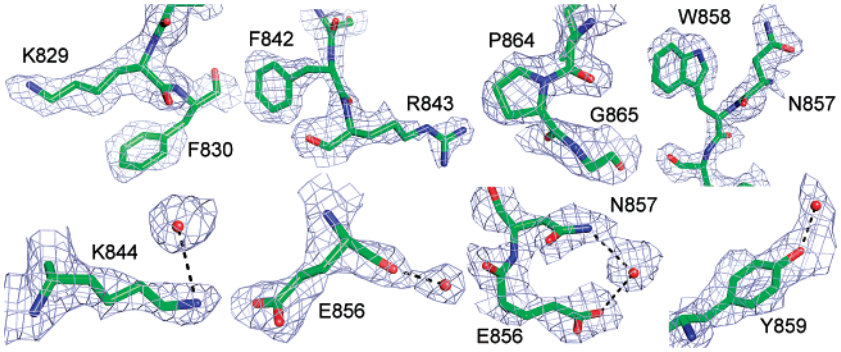


Figure 2. Selected regions of the final refined structural model in stick representation, and the corresponding total omit map contoured at 1σ. This figure was generated using PYMOL (http://pymol.sourceforge.net/).

essential for augmenting the robustness of the structure refinement and the quality of the electron density maps. The different data sets have slightly different lattice parameters due to sample preparation or induced by exposure to radiation. These cause relative shifts in the positions of neighboring peaks, thus reducing correlations between overlapping reflections. In total, 1980 stereochemical restraints were imposed in order to refine the positions of 544 protein atoms in the asymmetric unit using experimental data in the resolution range of 15.3−2.27 Å for the data sets collected for sample A and 15.3−2.8 Å for the data sets collected for sample B. Two coefficients (A's and B's) were refined to account for the solvent scattering. These coefficients were varied separately for each of the different patterns, and they can therefore account for some of the small differences in peak intensities at low angle between the different profiles.

Periodical evaluations of the protein stereochemistry were crucial for monitoring the progress of the refinement, and different validation programs were employed (Supporting Information). In summary, after merging all four data sets and extracting the calculated structure factors during the Rietveld refinement the program SFCHECK[34] was employed to evaluate them against the observed ones. In the same procedure in SFCHECK a total OMIT map based on the Bhat procedure[35] was generated. Visual inspection of the refined model against the OMIT map was performed in Coot.[36] Where necessary, side-chain rotamers and displacement of small groups of main-chain atoms were adjusted and the refinement repeated.

Detection of Water Molecules. Water molecules were detected and added by inspecting the total OMIT map from SFCHECK in Coot. Peaks in the OMIT electron-density map over 1 sigma level that fulfill the geometry criteria were assigned and along with the protein atoms then subjected to further cycles of stereochemically restrained Rietveld refinement. Repetition of this process finally yielded 36 water molecules. The refinement proceeded smoothly for the four profiles leading to good quality of the final fit (total agreement factors: $R_{wp} = 3.82\%$, $R_p = 2.86\%$). Details of the refinement statistics are listed in Table 2, and the final fits to the four profiles are shown in Figures 1 and S7. Selected regions of the refined coordinates and total OMIT map computed at the final steps of analysis are presented in Figure 2.

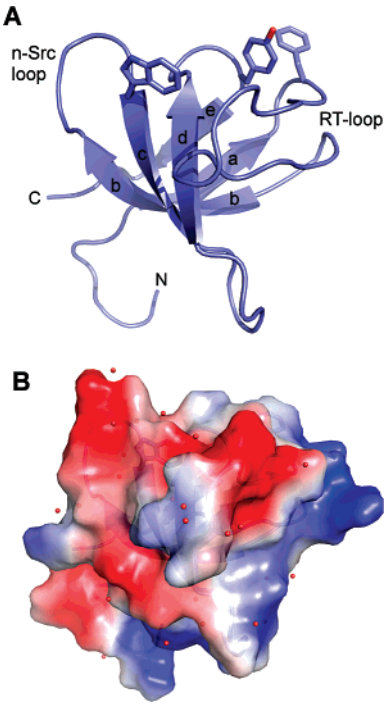


Figure 3. Powder-diffraction structure of the ponsin SH3.2 domain. (A) Ribbon representation of the SH3.2 indicating the secondary structure elements of the domain. The main hydrophobic residues of the binding interface as well as the positions of the n-Src and RT loops are indicated. (B) Electrostatic potential representation of the domain identifying additionally the water molecules as red spheres. This figure was generated using PYMOL (http://pymol.sourceforge.net/).

Structure Validation via Single-Crystal Methods. The powder and single-crystal structural models were superimposed and compared using the LSQKAB software.[37] The structure of the ponsin SH3.2 comprises the one common for SH3 domains five β-strand-fold arranged as two orthogonal β-sheets, forming an antiparallel β-barrel (Figure 3). The powder diffraction model is virtually identical to the one obtained from single-crystal measurements with a main-chain root-mean-square displacement (rmsd) of 0.532 Å and minor differences only at the N-terminus (Figure 4). The water molecules were compared using the program LSQMAN.[38] Despite the different data collection temperatures and resolution limits for the powder (room temperature) and single-crystal (100 K) measurements, 20 out of 36 water molecules are in good agreement (rmsd $\approx$ 1.368 Å).

(34) Vaguine, A. A.; Richelle, J.; Wodak, S. J. *Acta Crystallogr.* **1999**, *D55*, 191−205.
(35) Bhat, T. N. *J. Appl. Crystallogr.* **1988**, *21*, 279−281.
(36) Emsley, P.; Cowtan, K. *Acta Crystallogr.* **2004**, *D60*, 2126−2132.
(37) Kabsch, W. *Acta Crystallogr.* **1976**, *A32*, 922−923.
(38) Kleywegt, G. J. *Acta Crystallogr.* **1999**, *D55*, 1878−1884.

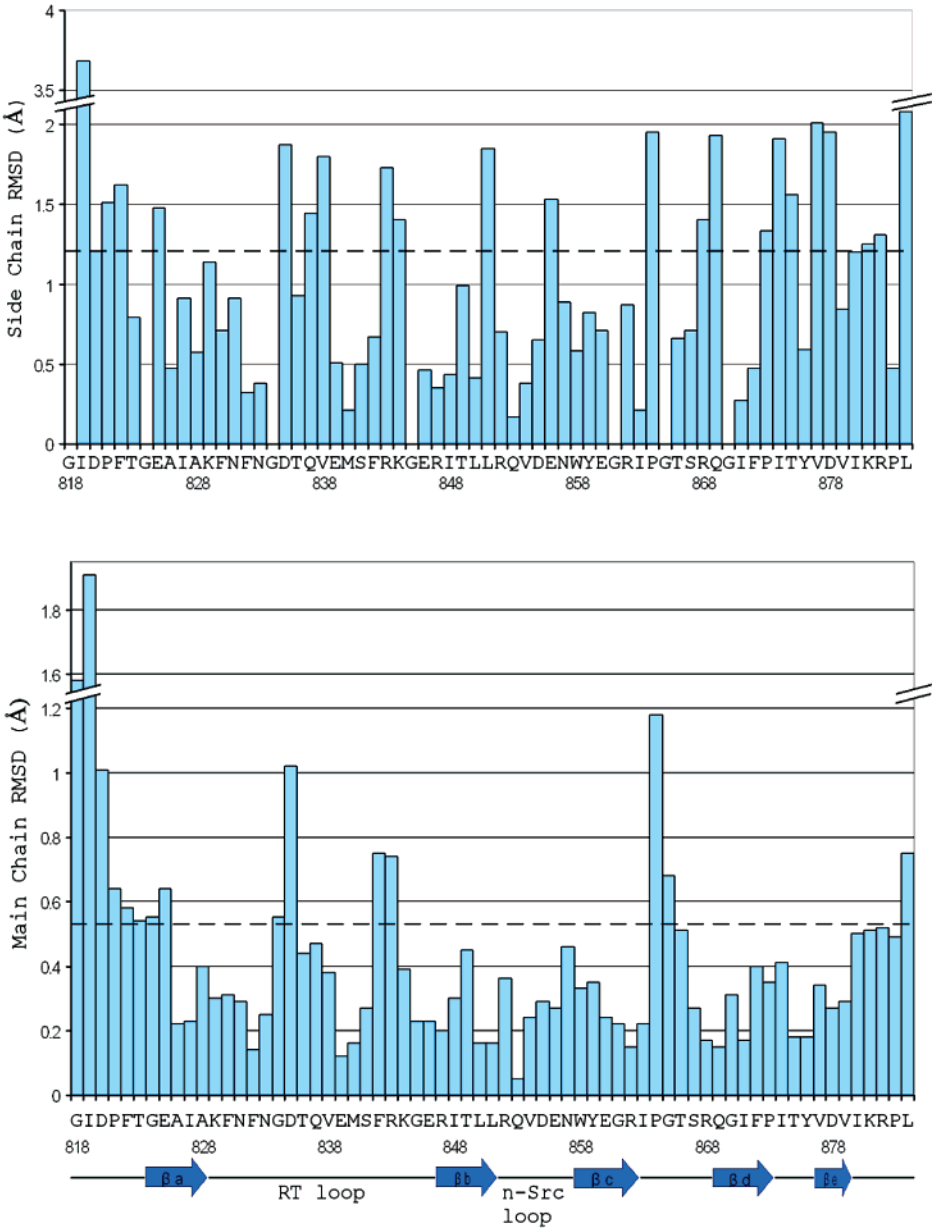


Figure 4. rms displacements of the main-chain (lower panel) and side-chain (upper panel) atoms between the refined models from powder and single-crystal diffraction data versus residue number. The dashed black lines correspond to the mean rmsd values, equal to 0.532 and 1.205 Å for the main- and side-chain atoms, respectively. The secondary structure assignment is shown at the lower panel.

Conclusions

Determination of an unknown protein structure from powder data is unique: 544 protein atoms and 36 water molecules were located in the asymmetric unit using a model of moderate similarity in molecular replacement and model building based on electron density maps. The studied domain of ponsin, SH3.2, is of significant biological interest owing to its role in adhesion-mediated signaling events and muscle differentiation processes.[15,16] Furthermore, the SH3 motif is one of the most abundant found in a large variety of signaling processes thus linked to several diseases and is therefore an important therapeutic target for drug design.[13,39] In addition to traditional structure determination of such domains via NMR or single-crystal X-ray diffraction, we demonstrated that microcrystalline samples of good quality powders can also provide adequate information content. Up to a resolution limit where overlapping intensities dominate, we show that single-crystal methods can be successfully applied following extraction of a set of extracted intensities from the measured powder diffraction profiles. There are various strategies for resolving overlapping reflections in powder diffraction data sets and improving the extracted intensity values. In the present study we exploited radiation-induced anisotropic lattice strains in a specially modified multipattern Pawley refinement taking into consideration likelihood criteria.[12] The outcome of this analysis is a set of improved extracted intensities, in good agreement with single-crystal intensities later obtained independently. The powder extracted data were sufficient for the structure solution of the domain via the molecular replacement method. Maximum likelihood refinement and inspection of the electron density maps allowed the

(39) Saksela, K. *Curr. Drug Targets Immune Endocr. Metabol. Disord.* **2004**, *4*, 315−319.

gradual building of the structural model and unraveling of the molecular conformation: a substantial result from powder diffraction data. The protein conformation is refined in a multiple data set via stereochemically restrained Rietveld analysis taking advantage of sample-induced anisotropic lattice strains. Finally, the model is further optimized, and details of the structure in addition to 36 water molecules are detected in total OMIT maps and verified by single-crystal diffraction experiments. Further advantages gained from the use of microcrystalline samples and powder diffraction measurements include verification of the homogeneity and phase purity of the protein precipitant and accurate lattice parameters allowing for direct observation of slight structure modifications and exploitation of radiation- and sample-induced effects. The domain structure is characterized at both room temperature, by powder diffraction, and 100 K, by measurements on single crystals, with only minor differences observed between them. This work provides new insights on how the molecular basis of materials such as small proteins can be adequately unraveled even when suitable single crystals are unavailable.

Coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 2O31 and 2O2W for the single crystal and powder diffraction model, respectively.

Acknowledgment. We thank ESRF and EMBL/DESY for provision of beam time. We also thank Bob Von Dreele for provision of software employed for the analysis of the powder diffraction data presented herein.

Supporting Information Available: Assessment of powder diffraction data and Rietveld analysis, (2) tables S1 and S2, and Figures S1−S9. This material is available free of charge via the Internet at http://pubs.acs.org.

JA073846C