Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Anisotropic behaviour of the C-terminal Kunitz-type domain of the α3 chain of human type VI collagen at atomic resolution (0.9 Å)

Bernadette Arnoux,* Arnaud Ducruix and Thierry Prangé

Laboratoire de Cristallographie et RMN Biologiques (UMR 8015 CNRS), Faculté de Pharmacie 4, Avenue de l'Observatoire, 75006 Paris, France

Correspondence e-mail: arnoux@pharmacie.univ-paris5.fr

Received 5 November 2001
Accepted 22 April 2002

PDB Reference: collagen Kunitz-type domain, 1kth.

The C-terminal Kunitz-type domain from the α3 chain of human type VI collagen (C5), a single amino-acid residue chain with three disulfide bridges, was refined at 0.9 Å resolution in a monoclinic form, space group $P2_1$ with one molecule per asymmetric unit, using data collected at cryogenic temperature (110 K). The average protein $\langle B \rangle$ factor decreases from 21 Å^2 at room temperature (RT) to 12 Å^2 at cryotemperature (100 K, CT). The spatially close N- and C-termini remain highly disordered. The different structural motifs of C5 were analyzed in terms of rigid-body displacement (TLS analyses) and show dominant libration motion for the secondary structure.

1. Introduction

The availability of low-temperature data-recording techniques is a way to experimentally analyze the dynamic motion of residues in proteins, compared with normal temperature experiments (Parkin *et al.*, 1997; Harata *et al.*, 1999; Wilson & Brunger, 2000; Teixeira *et al.*, 2001). The C-terminal domain of human type VI collagen (α3 chain), consisting of 58 amino-acid residues and named C5, has a BPTI-like structure and provides a good model. C5 diffracts to atomic resolution and allows anisotropic analyses of both local and overall motions of atoms or groups of atoms in the crystal matrix (Merigeau *et al.*, 1998). The structure of a longer fragment of the α3 chain (76 residues) including the C5 domain has been solved in solution by the NMR technique (Zweckstetter *et al.*, 1996) and is the origin of controversial results regarding the thermal behaviour of the central core (Sørensen *et al.*, 1997). The aim of the present study is to clarify some points about the thermal motion of C5 in the crystalline state, a study difficult to perform at room temperature (Arnoux *et al.*, 1995).

2. Experimental

2.1. Crystallization, data collection and data processing

Purified C5 protein was a gift from K. Norris (Novo Nordisk). Crystals were grown at 291 K as previously described (Merigeau *et al.*, 1998), and soaked in 20% glycerol for cryocooling prior to X-ray exposure. A high-resolution (0.9 Å) and a fast low-resolution (1.5 Å) data set were collected at 110 K from a single crystal with 1° oscillation per frame at the BW7A synchrotron beamline, DESY, Hamburg. A MAR345 imaging plate was used at a wavelength of 0.900 Å. Data were processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Crystals belong to the monoclinic space group $P2_1$, with unit-cell parameters a = 25.24, b = 37.74, c = 28.49 Å, β = 109.29° and Z = 2. The unit-cell parameters are slightly smaller than those recorded at 293 K. Crystal mosaicity was refined to 0.4°. A total of 179 075 observations in the resolution range 8–0.9 Å were reduced to 33 572 unique reflections. The merged data set contains 95.2% of the expected reflections in this resolution range (92.1% in the last shell). The overall $R_{\rm merge}$ was 4.1% (34% in the highest resolution shell).

2.2. Refinement

An initial model was calculated using the 1.2 Å structure of C5 without water molecules (Arnoux *et al.*, 1995; PDB code 2knt). The refinement process started with a rigid-body refinement cycle with resolution limited to 2 Å. The second step was composed of a series of slow-cooling steps refining positional parameters using *X-PLOR*/*CNS* (Brünger *et al.*, 1987, 1998) at a resolution of 1.8 Å. In a third step, *ARP*/*wARP* (Perrakis *et al.*, 1999) was used to localize 115 water molecules. Water peaks were selected according to their distances to hydrogen-bonding partners and were removed when the peak height in the $2F_o - F_c$ electron density was less than 0.7σ. All water molecules were carefully inspected manually using *TURBO-FRODO* (Roussel & Cambillau, 2001). An isotropic B-factor cutoff of 35 Å^2 for solvent molecules was introduced to limit overfitting. Refinement converged to R = 21.5% ($R_{\rm free}$ = 26.1%). Electron densities which could include either sulfate or phosphate ions were localized in two places. Isotropic

Table 1
Final refinement statistics.

No. of atoms	
Protein	$C_{288}H_{425}N_{77}O_{84}S_6$
Solvent (waters)	91 H_2O
Phosphate ions	2 PO_4^{3-}
Refinement	
Resolution range (Å)	8–0.95
No. of parameters/restraints	5109/6012
R factor (%), for 4σ observed data (No. of data)	12.1 (22448)
R factor (%), for all data (No. of data)	13.2 (28107)
Free *R* factor (%) for 6% of observed data (No. of data)	16.2 (1147)
R.m.s. deviations (number/value)	
1–2 bond length (Å)	494/0.016
1–3 bond-distance angles (Å)	665/0.032
Planes (Å)	159/0.041
Non-zero chiral volumes (Å^3)	62/0.11
Zero chiral volumes (Å^3)	77/0.1
Isotropic $\langle B \rangle$ values (Å^2)	
Main chain	12.2
Side chain	17.1
Water molecules	22.2
Phosphate ions	27.5
Residuals in final Fourier map (min/max) (e)	−0.32/+0.52

refinement continued with *SHELX*97. At the end of the first cycle, R = 21.8% and R_{free} = 24.5% (the same 6% of the data were used to calculate R_{free} as were used in *X-PLOR*).

In the next step, isotropic riding H atoms were introduced into the model, followed by individual anisotropic B factors. The structure was further refined by conjugate-gradient least-squares minimization using *SHELX*97 (Sheldrick & Schneider, 1997). Anisotropic B factors of protein atoms were restrained in the atom bond direction (DELU/SIMU commands). The anisotropic B factors of water molecules were not restrained. A few amino-acid residues showed clear alternate positions and were refined with the sum of their occupation factors constrained to unity. In the last steps, restraints on anisotropic factors were gradually relaxed so as not to bias the TLS analyses. The final model includes four residues with alternate positions (Arg1 and Thr13, Ile18 in loop L1) unambiguously assigned, two phosphates and 91 water molecules. The final model led to R = 12.1%, R_{free} = 16.2% (Table 1).

3. Results and discussion

Data recorded at 110 K (CT experiments) confirmed the overall conformation of the C5 fragment at room temperature (RT). They also reveal the last residue to have a different orientation, showing a clear signature for its terminal carboxyl group.

Both phosphate and sulfate ions were present in the crystallization medium. Two tetrahedral anions were located and phosphate was confirmed based on the refinement of a free parameter set to the X—O distance (X = P for phosphate or S for sulfate), where the expected value is 1.55 Å for P—O or 1.42 Å for S—O. This parameter converged to an average of 1.52 (2) Å, corresponding to two phosphates, not sulfates.

In solution, the RT structure of C5 solved by NMR (Zweckstetter *et al.*, 1996) showed a highly mobile structure with at least 24 residues in multiple conformations. A second investigation made by Sørensen *et al.* (1997) on the present shorter fragment (58 residues) concluded that the global fold was nearly identical to the crystal structure (Arnoux *et al.*, 1995) recorded at room temperature.

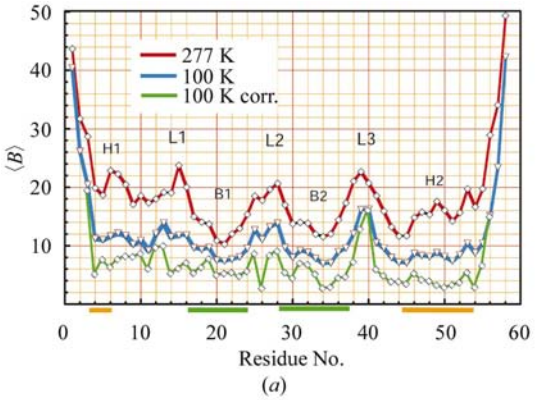
In the crystalline state at room temperature, C5 shows important dynamic motion, but only in the two N- and C-terminal extremities. At 110 K, the N- and C-terminal regions still remain highly flexible as is the case to a lesser extent for the three loops L1–L3. Fig. 1(*a*) shows the evolution of the average $\langle B \rangle$ factors of RT and CT refinements. There is a systematic decrease in the $\langle B \rangle$ factors of about 10 Å^2 between the RT and CT refinements. This overall decrease is also evident from initial Wilson plots ($\langle B \rangle_{RT}$ = 21.5 and $\langle B \rangle_{CT}$ = 11.9 Å^2).

3.1. TLS analyses

The inner core of C5 is rather rigid but the side chains at the surface presents a high degree of collective motion that was analyzed in terms of a many-rigid-body model. The TLS parameters (Schomaker & Trueblood, 1968; Winn *et al.*, 2001) were derived using programs *ANISOANL* and *TLSANL* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994).

A first analysis was performed using the whole main chain as a single rigid group, removing from the calculations the three residues at both N- and C-termini. It was clear that the structure underwent a rigid-body motion as shown from the radial and tangential distributions of the thermal U_{ij} projections along a radial vector about the centre of mass (U_{rad} are roughly constant, while U_{tang} linearly increase). This is also evident from the *RASMOL* representation (Merritt & Bacon, 1997) shown in Fig. 1(*b*). The anisotropic U_{ij} components were corrected from this rigid-body motion to derive residuals corresponding to the static disorder in the crystal. The corrected B factors are shown as a green line in Fig. 1(*a*). Obviously, the loops cannot be described as rigid blocks compared with the other structured parts of the molecule.

A more accurate description of the vibration modes is obtained by dividing C5 into several individual groups: the two helices (H1, H2), the two strands (B1, B2) and the three loops (L1–L3). For each group, TLS parameters were determined and analyzed. As expected, the four



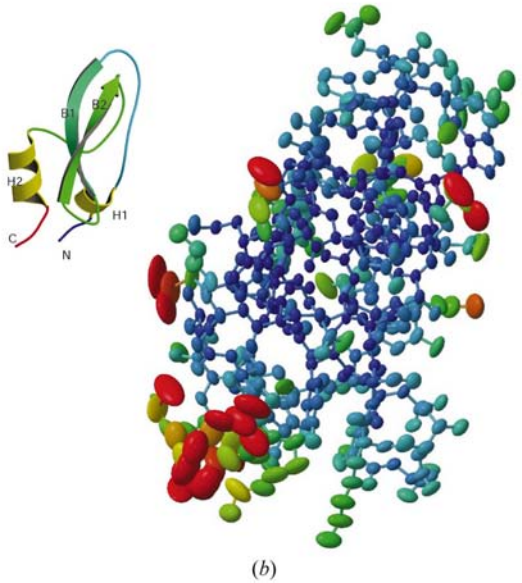


Figure 1
(*a*) Comparison between room-temperature (RT, red) and cryotemperature (CT, blue) average $\langle B \rangle$ factors in residues. The residual $\langle B \rangle$ factors after correction from a rigid-body motion are also displayed as a green curve. (*b*) *MOLSCRIPT*/*RASMOL* representation of the anisotropic structure (Kraulis, 1991; Merritt & Bacon, 1997). The ellipsoids are colour-coded from blue to red over the range 4–40 Å^2.

Table 2
Mean displacements (TLS components) given as the trace of the TLS tensors with respect to orthogonal axes.

Rigid group	No. of atoms	Trace of mean translation (Å)	Trace of mean libration (°)
α-Helix 1	30	0.13	9.70
β-Sheet 1	85	0.11	2.87
β-Sheet 2	65	0.09	4.83
α-Helix 2	70	0.09	7.37

secondary-structure motifs (helices and sheets) display mainly libration (L) motions, while the loops are not structured from a rigid-body point of view. Mean translational displacements (T component) about axes are all less than 0.2 Å^2 and screw (S) modes are lower than $\pm$0.3 Å°. Table 2 summaries the traces of TLS tensors (/3) in the four structural motifs (tensors themselves are not reported).

4. Conclusions

Compared with that at room temperature, the refinement of the C5 fragment at 100 K shows that the structure is segmented into discrete secondary-structure motifs (α-helices, β-sheets) with specifically correlated libration rigid motions. This observation is not valid for the loops connecting these motifs, as they undergo random non-concerted displacements. The present analysis of the C5 fragment compares with the room-temperature NMR and X-ray dynamic data and also corroborates normal mode calculations (Merigeau *et al.*, 1998). In addition, this study also shows that the N- and C-termini are mainly involved in local static disorder rather than in dynamic motion.

References

Arnoux, B., Mérigeau, K., Saludjian, P., Norris, F., Norris, K., Bjørn, S., Olsen, O., Petersen, L. & Ducruix, A. (1995). *J. Mol. Biol.* **246**, 609–617.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Harata, K., Abe, Y. & Muraki, M. (1999). *J. Mol. Biol.* **287**, 347–358.

Kraulis, P. E. (1991). *J. Appl. Cryst.* **24**, 946–950.

Merigeau, K., Arnoux, B., Norris, K., Norris, F. & Ducruix, A. (1998). *Acta Cryst.* D**54**, 306–312.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Parkin, S., Rupp, B. & Hope, H. (1997). *Acta Cryst.* D**52**, 18–29.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Roussel, A. & Cambillau, C. (2001). *TURBO-FRODO. A General-Purpose Molecular-Modelling Environment.* AFMB-CNRS, 13402 Marseille CEDEX 20, France.

Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* B**24**, 63–76.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Sørensen, M. D., Bjørn, S., Norris, K., Olsen, O., Petersen, L., James, T. & Led, J. J. (1997). *Biochemistry*, **36**, 10439–10450.

Teixeira, S., Lo Leggio, L., Pickersgill, R. & Cardin, C. (2001). *Acta Cryst.* D**57**, 385–392.

Wilson, M. A. & Brunger, A. T. (2000). *J. Mol. Biol.* **301**, 1237–1256.

Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.

Zweckstetter, M., Czisch, M., Mayer, U., Chu, M. L., Zinth, W., Timpl, R. & Holak, T. A. (1996). *Structure*, **4**, 195–209.