Acta Cryst. (1996). D**52**, 942-945

Conformational Change of the Adenovirus DNA-Binding Protein Induced by Soaking Crystals with $K_3UO_2F_5$ Solutions

PANAGIOTIS N. KANELLOPOULOS,[a] DEMETRIUS TSERNOGLOU,[b] PETER C. VAN DER VLIET[c] AND PAUL A. TUCKER[a]*

[a]*Structural Biology Programme, European Molecular Biology Laboratory, D 69012 Heidelberg, Germany,* [b]*Biochemistry Department, Universita 'La Sapienza', Piazzale Aldo Moro 5, I 00185 Rome, Italy, and* [c]*Laboratory for Physiological Chemistry, Utrecht University, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands. E-mail: tucker@embl-heidelberg.de*

(*Received 27 November 1995; accepted 29 April 1996*)

Abstract

Soaking crystals of the C-terminal DNA-binding domain of the adenovirus single-stranded DNA-binding protein with a buffer containing $K_3UO_2F_5$ results in a 9% change of the crystallographic *c* axis without destruction of the crystals or appreciable loss of resolution. The crystals belong to space group $P2_12_12_1$ with $a = 79.7$, $b = 75.6$ and $c = 60.6$ Å. The three-dimensional structure has been refined to 2.7 Å with a crystallographic *R* factor of 0.206. Antiparallel chains of protein molecules running through the entire crystal are linked by uranyl ions. The relative orientation of protein monomers is flexible, even in the crystalline state, and allows changes in the packing of the protein chains.

1. Introduction

The soaking of heavy-atom compounds into protein crystals has been a technique of key importance in obtaining phase information during the X-ray analysis of protein crystals. It is the usual procedure for the method of isomorphous replacement (Green, Ingram & Perutz, 1954) and has also been used in the multi-wavelength anomalous dispersion method (Hendrickson, 1991). Isomorphous replacement requires that the heavy-atom compound enters the crystal without disturbing either the crystal packing or the structure of the protein. Amongst the classes of heavy-atom compounds (for overview see Blundell & Johnson, 1976) are uranyl ion containing salts. $K_3UO_2F_5$, a member of this class, has been successfully used a number of times.

Soaking trials with $K_3UO_2F_5$ were performed during derivative searches with crystals of the adenovirus single-stranded DNA-binding protein (DBP). The derivative was non-isomorphous showing a 9% change in the *c*-axis length. DBP is a well known member of the class of polypeptides called single-stranded DNA-binding proteins (for a review see Chase & Williams, 1986). The three-dimensional structure of the C-terminal domain has been solved in two crystal forms (Tucker *et al.*, 1994; Kanellopoulos, Tsernoglou, Van der Vliet & Tucker, 1996). The structures show chains of protein molecules in the crystal which form as a result of a hydrophobic interaction between the 17 C-terminal residues of one molecule and a hydrophobic cleft in the adjacent molecule. Deletion of this 17 amino-acid C-terminal extension results in reduction of cooperativity (Tucker *et al.*, 1994). Here, we describe the changes caused to the DBP crystal structure upon soaking DBP crystals in a solution of $K_3UO_2F_5$.

2. Experimental

The C-terminal domain of DBP was obtained by controlled chymotryptic digestion of the full-length protein and was crystallized as described previously (Tsernoglou, Tucker & Van der Vliet, 1984). The crystals were initially transferred to a 50 m*M* acetate buffer pH 5.8 containing 2.0 *M* $NaNO_3$ in four steps and subsequently soaked in the same buffer containing 0.4 m*M* $K_3UO_2F_5$ for 24 h at 293 K. X-ray intensities for a single crystal were collected to 2.7 Å resolution using a 18 cm MAR image-plate detector and pyrolytic graphite-monochromatized Cu $K\alpha$ radiation ($\lambda = 1.5418$ Å) produced by a rotating-anode generator running at 40 kV, 90 mA. Data collection was by the rotation method using 1° rotation frames. The X-ray data were processed and scaled using the program *XDS* (Kabsch, 1988). The crystals were found to be orthorhombic, belonging to the space group $P2_12_12_1$, with unit-cell dimensions $a = 79.7$, $b = 75.6$ and $c = 60.6$ Å. The data-collection statistics are summarized in Table 1. The cell dimensions of the non-derivatized crystals are $a = 79.5$, $b = 76.0$ and $c = 67.4$ Å. The difference in cell dimensions arises from the use of the uranyl fluoride since transfer to the acetate/$NaNO_3$ buffer alone has no significant effect on the cell dimensions.

The orientation of the DBP molecule in the new unit cell was determined by molecular replacement using the program *AMoRe* (Navaza, 1994). The DBP native structure was used as a search model, omitting the 17 C-terminal amino acids. By application of the fast rotation function, using reflections in the resolution range 15–3.5 Å and an integration radius of 22 Å, one major peak appeared at 10σ. This solution was used to perform a translational search, using reflections from 15

Table 1. *Summary of data collection*

Resolution range (Å)	15–2.7
No. of observed reflections	34497
No. of unique reflections	10312
Completeness (%)	98.7
R_{sym}*	14.7

* $R_{sym} = \sum_{\mathbf{h}} \sum_i |F_i^2(\mathbf{h}) - \langle F^2(\mathbf{h})\rangle| / \sum_{\mathbf{h}} \sum_i F_i^2(\mathbf{h})$.

to 3.0 Å. A single solution was obtained, and was refined by the rigid-body refinement procedure, implemented in *AMoRe*, to an initial crystallographic *R* factor of 0.39 in the resolution range 15–3.0 Å. The solution corresponds approximately to 6° rotations about the *b* and *c* axes and an approximate 7 Å translation along the *c* axis.

The model was further refined using the program *X-PLOR* (Brünger, 1992*a*). Initially, the molecular replacement solution was subjected to an additional four cycles of rigid-body refinement, and during each cycle both the number of reflections and the rigid groups were increased. At this point, the missing 17 C-terminal amino acids were remodelled from $2F_o - F_c$ and $F_o - F_c$ Fourier syntheses. Additionally, four heavy-atom peaks were located in the electron-density maps and U atoms were placed in positions corresponding to the density. The model building was performed using the program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991). The model was then subjected to refinement using simulated-annealing methods. Specifically, each step of refinement included simulated-annealing refinement with a starting temperature of 3000 K, 20 cycles of restrained temperature-factor refinement, and finally, 160 cycles of conjugate-gradient minimization. All data in the resolution range 6.0–2.7 Å were used and no sigma cut-off was applied. In the next steps, additional cycles of model building and refinement were performed. To decide whether or not to include residues, the free *R* value was used (Brünger, 1992*b*), and simulated-annealing omit maps were calculated. The refinement with *X-PLOR* gave a resulting model with a crystallographic *R* factor of 0.213 and a free *R* factor of 0.314 (10% of the data) for the resolution range 6.0–2.7 Å. The model was finally refined using *TNT* (Tronrud, Ten Eyck & Matthews, 1987) by the conjugate-gradient direction method. 31 water molecules have been included in the final model. The crystallographic *R* factor for all data (10 312 reflections) in the resolution range 15–2.7 Å was 0.206. The regions 174–178, 299–326 and 454–462 of the polypeptide chain were either invisible or poorly defined in the electron density and are omitted from the final model.*

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1ANV, R1ANVSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0603). At the request of the authors, the atomic coordinates and structure factors will remain privileged until 1 January 1997.

3. Results and discussion

The refined model shows a good fit to the $2F_o - F_c$ electron-density map, calculated for reflections in the resolution range 15–2.7 Å. The real-space correlation coefficient per complete residues has been calculated using the program *O* (Jones *et al.*, 1991) and is good in general, apart from loop regions (data not shown). The pattern is similar for the temperature factors, and the mean temperature factor for the structure is 30.2 Å^2. The r.m.s. deviations from ideality are 0.01 Å for bond lengths and 2.04° for bond angles. No residues lie in the disallowed regions of the Ramachandran plot, as calculated with the program *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993). Based on the Luzzati plot (Luzzati, 1952), the maximum coordinate error was found to be approximately 0.35 Å at 2.7 Å resolution.

The three-dimensional structure is similar, but not identical, to that of the unsoaked crystals (Tucker *et al.*, 1994) (Fig. 1). Superposition of the common Cα

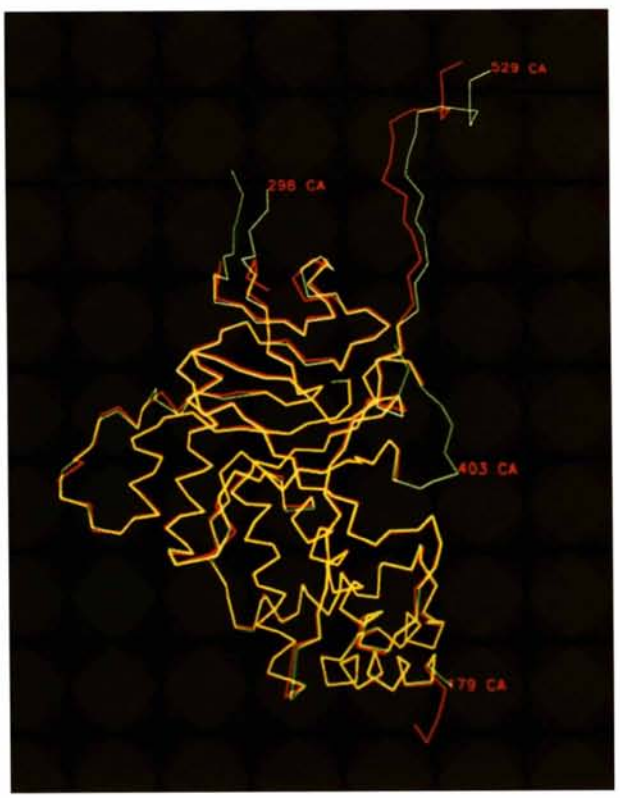

Fig. 1. Superposition of the Cα positions for the core of the present structure (green) and the non-derivatized DBP structure (red). Representative residues in regions of significant differences are marked. The 17 amino acids of the C-terminal extension were not used for the calculations of the least-squares fit. The transformation matrix relating the derivatized ($\mathbf{x}_g$) to the underivatized ($\mathbf{x}_r$) molecule is

$$\mathbf{x}_g = \begin{pmatrix} 0.9996 & 0.0054 & 0.0272 \\ -0.0083 & 0.9940 & 0.1095 \\ -0.0264 & -0.1097 & 0.9936 \end{pmatrix} \mathbf{x}_r + \begin{pmatrix} 0.3307 \\ 0.2269 \\ 7.4973 \end{pmatrix}$$

atom positions between residues 174 and 512 gives an r.m.s. deviation of 0.66 Å which is rather larger than the estimated coordinate error. The r.m.s. deviation for the C-terminal residues (513–529) after this superposition is much larger, being 5.4 Å. The presence of the uranyl ions has caused a number of differences. The first and most surprising difference is that soaking $K_3UO_2F_5$ into the native crystals causes a 9% shrinkage along the crystallographic *c* axis. Despite this fact, the crystals remain integral and their diffraction quality on a conventional X-ray source is comparable with that of the native crystals. Infinite chains of protein molecules run parallel and antiparallel to the crystallographic *y* axis. The difference in *c* axis length results from an additional attraction between these chains of protein molecules because of the presence of the uranyl ions. The way in which the protein chains are arranged in the crystal is shown in Fig. 2. The decrease in the length of the *c* axis results in approximately a 9% decrease in solvent content and, therefore, to a significant dehydration of the crystal.

The movement of the protein chain in essentially one dimension would be facilitated by flexibility of the protein chain. Evidence for such flexibility is provided by the structure of DBP in the second crystal form (Kanellopoulos *et al.*, 1996). The two reported crystal forms show that the difference in orientation of the C-terminal extension relative to the body of the protein results primarily from a change at a 'hinge' region around N512. In the two cases the main-chain torsion angle of this residue move between two favourable regions in the Ramachandran plot. In the derivative crystal yet another relative orientation (though more similar to the orientation in the first crystal form than in the second crystal form) of the C-terminal arm to the protein core is found (Fig. 2) supporting the idea that the protein chain is more flexible than suggested by a simple two-state model. The residues responsible for the flexibility again lie in the 'hinge' region of the tail (residues N512 and V513), while the conformation of the tail itself is almost identical to that in the non-derivatized structure. Fig. 2 also shows that part of a loop (I297–I331) not visible in the native structure is stabilized presumably by interaction with one of the heavy-atom sites (see below). These residues are in an extended conformation and the base of the loop has a significantly different conformation, which would support the hypothesis that it is flexible and could bend over the protein core during DNA binding (Tucker *et al.*, 1994). The loop region 401–405 is also stabilized, relative to the native structure, and shows good electron density. In contrast, the five first N-terminal residues (174–178) are not well defined and are omitted from the final model.

The structure at 2.7 Å resolution reveals two distinct metal-binding sites. The resolution of this structure does not allow the determination of the detailed coordination geometry of the U atoms, however the chemically stable uranyl (UO_2^{2+}) cation is a likely basic unit, which would be expected to be coordinated by electronegative ligands. The first metal-binding site is formed by a triangle of U atoms with distances between the U atoms of 3.4, 3.6 and 4.1 Å.* The three-membered cluster is surrounded almost exclusively by negatively charged residues. U533 is responsible for stabilization of part of the DBP large flexible loop (residues 294–335) *via* contacts with the side chains of E298 and of Q330. Two additional residues found in loop regions are responsible for the interactions with the remaining two members of the cluster. D429 approaches U534 and E225 approaches U535. Residue E225 comes from a symmetry-related molecule on the same chain of protein molecules as E298, whilst D429 comes from a symmetry-related DBP molecule on a different chain of protein molecules,

* We thank a referee for pointing out that uranium-containing clusters are known to be formed in multicomponent systems containing the uranyl ion, for example the $K^+/UO_2^{2+}/F^-/NO_3^-/H_2O$ system (Groeweghe, 1954).

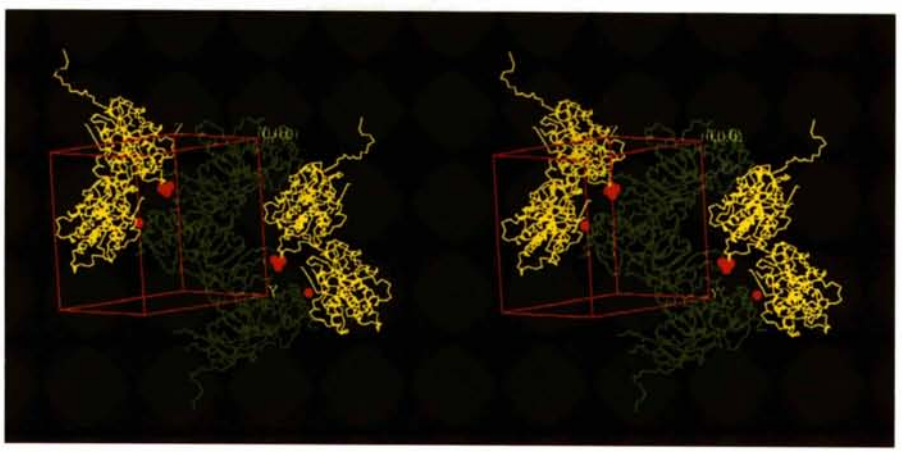

Fig. 2. A stereoscopic representation of the crystal packing, displaying how successive triple and single heavy-atom sites (red) link one DBP protein chain (green) with adjacent ones (yellow). The green chain is comprized of three molecules and the yellow chains of two molecules. Adjacent molecules in a protein chain are related by a crystallographic 2_1 screw axis along *b*.

serving as one of the means of attraction between protein chains. The other interaction between protein chains involves the single U atom U532 which interacts with the negatively charged side chains of residues D468 and D193 of a different protein chain (Fig. 2). N466 and a clear solvent site (HOH25) also interact with U532. It should be also noted, that the temperature factors of the U atoms in the clustered site, range between 50 and 55 $Å^2$ so they might not be fully occupied, whilst the temperature factor for the U atom of the single site is 38 $Å^2$.

From the crystal packing it is observed that successive triple-membered and single-membered uranyl sites bring the chains of protein molecules closer together along the crystallographic *c* axis (Fig. 2), which in turn is possible because different relative orientations of adjacent molecules in the chains are possible. In a similar manner $K_3UO_2F_5$ has been used in the past to obtain oriented specimens of tobacco mosaic virus (Holmes & Leberman, 1963).

References

Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*. London: Academic Press.

Brünger, A. T. (1992*a*). *X-PLOR Version 3.1, A System for X-ray Crystallography and NMR*. Yale University, New Haven, CT, USA.

Brünger, A. T. (1992*b*). *Nature (London)*, **355**, 472–475.

Chase, J. W. & Williams, K. R. (1986). *Annu. Rev. Biochem.* **55**, 103–136.

Green, D. W., Ingram, V. M. & Perutz, M. F. (1954). *Proc. R. Soc. London Ser. A*, **225**, 287.

Groeweghe, L. (1954). *Ind. Chim. Belge*, **19**, 1277–1282

Hendrickson, W. A. (1991). *Science*, **254**, 51–58.

Holmes, K. C. & Leberman, R. (1963). *J. Mol. Biol.* **6**, 439–441.

Jones, T. A, Zou, J. Y. Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Kanellopoulos, P. N., Tsernoglou, D., Van der Vliet, P. C. & Tucker, P. A. (1996). *J. Mol. Biol.* **257**, 1–8.

Laskowski, R. A., MacArthur, M. W., Moss, D. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291

Luzzati, P. V. (1952). *Acta Cryst.* **5**, 802–810.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163

Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.

Tsernoglou, D., Tucker, A. D. & Van der Vliet, P. C. (1984). *J. Mol. Biol.* **172**, 237–239.

Tucker, P. A., Tsernoglou, D., Tucker, A. D., Coenjaerts, F. E. J., Leenders, H. & Van der Vliet, P. C. (1994). *EMBO J.* **13**, 2994–3002.