The structure of DNA binding protein II at 6 Å resolution

Isao Tanaka, Stephen W. White, Krzysztof Appelt, Keith S. Wilson[+] and Jan Dijk

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin-Dahlem 33, Germany

Received 24 October 1983

The structure of DNA binding protein II from *Bacillus stearothermophilus* is described at 6 Å resolution. The molecule exists as the dimer in our crystals, and there are 3 independent dimers related by non-crystallographic symmetry in the unit cell. The dimer is compact and globular with dimensions of about $32 \times 35 \times 39$ Å^3.

DNA binding protein *Thermophile* Eubacterium *X-ray analysis*

1. INTRODUCTION

A number of DNA binding proteins in prokaryotic cells have been described (review [1]). One protein in particular, DNA binding protein II, has been studied extensively. This protein, which reportedly binds to double stranded DNA, is abundant and ubiquitous in both eubacteria and archebacteria. Although its precise function in the cell is still uncertain it is assumed to exert a histone-like role in the stabilisation of the DNA structure. The protein has been found associated with ribosomes and their subunits [2] where its function, if any, is even more obscure.

The protein has a monomeric relative molecular mass of approximately 9500 and is assumed to occur as a tetramer in solution. The amino acid sequences of the proteins from a number of prokaryotes, both eubacteria and archebacteria, are now available (review [3]).

We have reported preliminary crystallisation conditions for the DNA binding protein II from the thermophilic bacterium *Bacillus stearothermophilus* [2]. In this paper the results of the structural analysis of these crystals at low resolution are presented. The protein occurs in the form of dimers, 3 of which are present in the unit cell. The molecular boundary is very well defined and it is possible to see helical regions probably linked by pleated sheet structure.

2. MATERIALS AND METHODS

2.1. *Protein isolation*

DNA binding protein II was extracted and purified as in [2].

2.2. *Protein crystallization*

The protein was crystallised by the hanging-drop vapour diffusion method [4] as in [2]. Our studies here have used the monoclinic crystal form which grows optimally from 35% 2-methylpentane-2,4-diol, at pH 8.0, with 80 mM phosphate buffer, and at a protein concentration of about 100 mg/ml. The heavy atom derivative was prepared by first transferring a native crystal to a solution containing 40% 2-methylpentane-2,4-diol with 100 mM sodium acetate/ammonium sulphate mixture at pH 7.5, followed by transfer to a similar solution containing in addition 0.2 mM $K_3UO_2F_5$. The crystal was soaked in this solution for 3 days.

2.3. *Data collection*

Data were recorded photographically using an Enraf-Nonius Arndt-Wonacott oscillation camera. CuK_α X-radiation was provided by a Seifert sta-

[+] Present address: Department of Physics, University of York, York, England

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

tionary anode operating with a fine focus tube at 40 kV and 30 mA. Native and derivative data were initially evaluated to a nominal resolution of 6 Å to permit rapid screening of potential heavy atom derivatives and calculation of a low resolution image of DNA binding protein II. The data recorded for the native protein and the uranyl derivative are summarised in table 1. Both data sets were recorded from a single crystal of dimensions of about 0.2 × 0.2 × 1.5 mm^3. The complete data collection took 4 days per crystal, with the crystal being translated every 24 h on the long axis to expose a fresh region to the beam and reduce the effects of radiation damage.

Table 1

Summary of the 6 Å X-ray data for DNA binding protein II

Crystal	Concentration of heavy atom reagent	Soak time	R_{sym} (%)	Δ_{iso} (%)
1. Native protein		–	3.64	–
2. $K_3UO_2F_5$	0.2 mM	3 days	3.84	20.3

R_{sym} is defined as:

$$\sum_h \sum_i |I_{ih} - \bar{I}| \; / \; \sum_h \sum_i I_{ih}$$

where $\bar{I}$ is the mean intensity of i equivalent reflections with indices given by h. Δ_{iso} is the mean fractional isomorphous difference summed over all reflections

3. RESULTS AND DISCUSSION

DNA binding protein II crystallises in a number of different forms, two of which we have described previously [2]. Of these two the monoclinic form is most suitable for X-ray analysis and it is this form which we have studied. The crystals grow optimally under the conditions given above and diffract to a resolution better than 2.5 Å. The space group is P2 with cell dimensions a = 65.5 Å, b = 37.3 Å, c = 65.5 Å and β = 114.5°. Based on the results of previous solution studies [2] we initially assumed that the molecule would exist as a tetramer and this led to possible values of V_M (daltons of protein per $Å^3$ packed into the crystal) of either 1.85 for two tetramers or 3.70 for one tetramer in the unit cell. These values lie at the two extremes of the range observed for protein crystals [5]. The true value of V_M is given below.

We have to date produced a single heavy-atom derivative of the protein using $K_3UO_2F_5$ as described above. There is a single binding site for the heavy atom, the position being easily calculated from the difference Patterson synthesis. The least-squares refinement of the derivative is described in table 2.

Using the isomorphous and anomalous scattering of this derivative we have calculated single isomorphous phases for DNA binding protein II with a mean figure of merit of 0.72 at 6 Å resolution. The derivative appears to be highly isomorphous and the crystals diffract to beyond 3 Å. Thus, it will be useful for the high resolution analysis of the structure.

A superposition of several sections of the electron density synthesis calculated with these phases is shown in fig.1. There is excellent contrast bet-

Table 2

The parameters for the least-squares refinement of the heavy atom derivative at 6 Å resolution

Derivative	Occupancy	x	y	z	B	f_H	E	R_c
$K_3UO_2F_5$	1.380	0.229	0.000	0.424	15.0	167	77	42.1%

The occupancy is on an arbitrary scale and the temperature (B) value has not been refined. f_H is the mean calculated heavy atom contribution, E is the r.m.s. lack of closure error. R_c is the R factor for the centric terms:

$$R = \sum_h ||F_{PH} - F_P| - f_H| \; / \; \sum_h |F_{PH} - F_P|$$

where F_P is the structure factor amplitude of the native protein and F_{PH} that of the heavy atom derivative

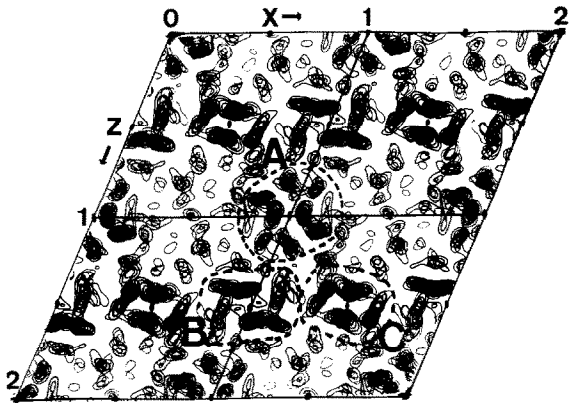

Fig.1. A superposition of 12 sections of the 6 Å electron density map of the DNA binding protein II in the *xz* projection. The slab of density is 22.4 Å thick and extends from $y = 0.85$ to 1.40. Four unit cells of the crystal are shown. There is excellent contrast between regions of protein and solvent and the molecular boundaries between the 3 molecules A–C can be clearly seen.

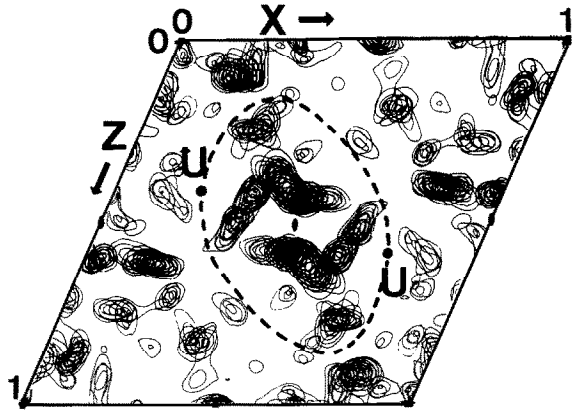

Fig.2. A superposition of 8 sections of one unit cell of the map in the *xz* projection, at 6 Å resolution. The slab of density is 14.9 Å thick and extends from $y = 0.95$ to 1.35. Molecule C is at the centre of the cell and strong rod-like features characteristic of α-helices at this resolution are apparent. 'U' indicates the uranium binding sites disposed about the 2-fold rotation axis in the centre of the unit cell.

ween regions of solvent and of protein, and the molecular boundary can be easily defined. The good contrast is readily explained by the low density of the solvent, which contains only a very low concentration of inorganic salt ions.

The explanation of the unexpected values of V_M now becomes clear. The unit cell contains in fact 3 dimers of the DNA binding protein, whose boundaries are indicated in fig.1. Each of these dimers is located on a different, parallel but independent 2-fold rotation axis of the space group, the monomers within a dimer being related to one another by these rotation axes. The 3 dimers are related to one another by non-crystallographic translations and rotations. The 3 dimers are labelled A–C in fig.1. Molecule B can be generated from molecule A by a translation to a different 2-fold axis by (0.0, 0.0, 0.5) fractions of the cell edge, followed by a rotation of 103° about this 2-fold. Molecule A requires a translation of (0.5, 0.325, 0.5), i.e., to a third 2-fold axis and by 0.325 along the *y* direction, followed by a rotation of 80° about this 2-fold, to generate molecule C.

The 3 molecules appear to be extremely similar in their structure, though we cannot exclude the possibility of small differences at high resolution due to different packing environments in the crystal. Superposition of molecule A onto B gives an overlap correlation of 76%, and onto C of 77% (100% indicating exact agreement of the densities).

The monomer appears to consist of two helices seen as the strong rod-like features in fig.2, supporting what could well be a region of β-pleated sheet. The helices are involved in the close contact between monomers within the dimers.

The presence of the non-crystallographic symmetry will greatly assist in the analysis of the structure, as it will enable us to average the electron density for the 3 dimers to improve the estimates of the phases for the data. We are presently extending the analysis to 3 Å resolution for the native protein and for the uranyl derivative, and also evaluating a promising $K_2Pt(NO_2)_4$ derivative. With the symmetry averaging procedure we hope that this will lead to a high resolution structure of the protein in the near future. We are also planning the use of synchrotron radiation to extend the resolution towards 2 Å. We have already shown that DNA binding protein II interacts with oligodeoxyribonucleotides in some preliminary studies in solution [2] and hope to be able to carry out some complementary studies in the crystal.

ACKNOWLEDGEMENTS

We thank Volker Brachvogel for technical assistance. This work has been partly funded by the Deutsche Forschungsgemeinschaft. I.T. is a Humboldt fellow.

REFERENCES

[1] Geider, K. and Hoffmann-Berling, H. (1981) Annu. Rev. Biochem. 50, 233–260.
[2] Dijk, J., White, S.W., Wilson, K.S. and Appelt, K. (1983) J. Biol. Chem. 258, 4003–4006.
[3] Kimura, M. and Wilson, K.S. (1983) J. Biol. Chem. 258, 4007–4011.
[4] Davies, D.R. and Segal, D.M. (1971) Methods Enzymol. 22, 266–269.
[5] Matthews, B.W. (1968) J. Mol. Biol. 33, 491–497.