

**CRYSTALLIZATION AND PRELIMINARY X-RAY DIFFRACTION
STUDIES OF THE ENGRAILED HOMEODOMAIN AND OF AN
ENGRAILED HOMEODOMAIN/DNA COMPLEX**

Beishan Liu, Charles R. Kissinger and Carl O. Pabo

**Howard Hughes Medical Institute
Department of Molecular Biology and Genetics
The Johns Hopkins University School of Medicine
Baltimore, MD 21205**

Enrique Martin-Blanco and Thomas B. Kornberg

**Department of Biochemistry and Biophysics
University of California
San Francisco, CA 94743**

Received July 19, 1990

The homeodomain from the engrailed protein of *Drosophila* has been crystallized from ammonium phosphate at pH 6.8. The crystals form in space group P6₁22 (or P6₅22), with cell dimensions $a = b = 44.8 \text{ \AA}$ and $c = 118.2 \text{ \AA}$. These crystals diffract to 1.8 \AA resolution. A complex containing the engrailed homeodomain and a duplex DNA site also has been crystallized. The cocrystals form in space group C2 with $a = 131.2 \text{ \AA}$, $b = 45.5 \text{ \AA}$, $c = 72.9 \text{ \AA}$ and $\beta = 119.0^\circ$. These crystals diffract to 2.6 \AA resolution.

© 1990 Academic Press, Inc.

The homeodomain is a 60-residue DNA-binding motif found in many eukaryotic proteins that regulate gene expression (1). Amino acid sequences of homeodomains show some similarities with the helix-turn-helix motif of prokaryotic DNA-binding proteins (2,3), and NMR studies of the Antp homeodomain have confirmed that homeodomains contain a helix-turn-helix structure (4). Although genetic studies have pinpointed some of the critical amino acids involved in recognition (5,6,7), relatively little is known about the homeodomain/DNA interactions.

In order to better understand the structural basis of homeodomain/DNA interactions, we have begun X-ray crystallographic studies of the homeodomain from the engrailed protein of *Drosophila*. The engrailed protein is required for proper development of segments during embryogenesis (8). Closely related homeodomains have been found in several other organisms, including man (9), and sequence comparisons suggest that the engrailed protein is representative of a distinct subfamily of homeodomain proteins (1). We report here the crystallization of the engrailed homeodomain and crystallization of a complex of the engrailed homeodomain with duplex DNA.

METHODS

A polypeptide containing the 60 residues of the engrailed homeodomain (10), along with an additional methionine residue on the N-terminus, was expressed in *E. coli* (11). Gel shift and DNase footprinting experiments confirmed that this polypeptide specifically recognizes engrailed binding sites *in vitro*. To purify the polypeptide, overproducing *E. coli* cells were sonicated and cellular debris was removed by centrifugation. Adding 0.5% polyethyleneimine and centrifuging again removed many other proteins. The engrailed homeodomain was further purified by chromatography on an S-Sepharose FPLC column and on a C₄ HPLC column. The protein was concentrated to about 9 mg/ml for crystallization experiments.

Crystals of the engrailed homeodomain were grown using the hanging-drop vapor diffusion technique (12), and the best crystals were grown from 74% saturated ammonium phosphate (pH 6.8) at room temperature. Crystals of an engrailed homeodomain-DNA complex were obtained by systematically varying the DNA length when surveying crystallization conditions (13). Crystallization conditions were tested with several different binding sites, since it is not yet known which site(s) are important for development. The largest crystals were obtained using a DNA duplex that has 21 bases on each strand and has an overhanging 5' end:

ATTAGGTAATTACATGGCAA
AATCCATTAATGTACCGTTT .

(This had initially been synthesized for cocrystallization with the yeast $\alpha 2$ protein, but gel shift experiments show that it binds to the engrailed homeodomain with a $K_D = 1-2 \times 10^{-9}$ M.) These crystals were also grown using the hanging-drop vapor diffusion method. The drops contained 4.5 mg/ml of protein and an equimolar amount of duplex DNA in a buffer containing 30 mM bis-Tris-HCl, pH 6.7. When the drops were set up, the pH was raised to 8.0-9.0 by the addition of ammonium hydroxide; crystals grew as the ammonium hydroxide diffused out and the pH returned to 6.7.

X-ray diffraction data was collected on a Xentronics multiwire area detector using a Rigaku rotating anode X-ray source.

RESULTS AND DISCUSSION

Crystals of the Engrailed Homeodomain

The largest engrailed homeodomain crystals measure roughly 0.8 mm in their longest dimension and 0.4 mm in the shortest dimension. Precession photographs revealed that the crystals form in one of the enantiomorphic space groups P6₁22 or P6₅22 and have unit cell dimensions $a = b = 44.8 \text{ \AA}$ and $c = 118.2 \text{ \AA}$. Assuming one protein monomer in the asymmetric unit, the calculated volume per dalton of protein (V_m) of the crystals is $2.30 \text{ \AA}^3/\text{Da}$, and this corresponds to a solvent content of 47% (14).

Native data has been collected to 1.8 \AA resolution on a multiwire area detector. A 4.0 \AA data set was also collected after soaking the crystals in 1 mM K_2PtCl_4 , and two Pt

sites were identified in the difference Patterson map. Additional derivative soaks are in progress.

Cocrystals of an Engrailed Homeodomain/DNA Complex

The largest cocrystals of the engrailed homeodomain/DNA complex measure about 0.8 mm in their longest dimension and about 0.4 mm in the shortest dimension. Gel electrophoresis of dissolved crystals proved that they contained both protein and DNA.

Precession photographs revealed that the cocrystals form in space group C2 with $a = 131.2 \text{ \AA}$, $b = 45.5 \text{ \AA}$, $c = 72.9 \text{ \AA}$ and $\beta = 119.0^\circ$. The volume of the unit cell is most consistent with the assumption that there are two 1:1 complexes in the asymmetric unit. This gives a calculated V_m of $2.26 \text{ \AA}^3/\text{Da}$ and a solvent content of 46%.

A complete native data set has been collected to 2.6 \AA . We also have collected data from several cocrystals that contain iodo-uracil substituted for thymine at specific positions in the duplex, and difference Patterson maps suggest that these will be excellent derivatives. A high-resolution structure should give us a much better understanding of homeodomain-DNA interactions.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health to C.O.P. (GM31471) and to T.B.K. (GM24635). We thank Anatoli Collector and Cynthia Wendling for synthesizing the DNA, and we thank Cynthia Wolberger for encouragement and help at several critical stages of this project. B.L., C.R.K., and C.O.P. are in the Howard Hughes Medical Institute.

REFERENCES

1. Scott, M.P., Tamkun, J.W. & Hartzell, G.W., III (1989) Biochim. Biophys. Acta 989, 25-48.
2. Laughon, A. and Scott, M. P. (1984) Nature 310, 25-31.
3. Shepard, J.C.W., McGinnis, W., Carrasco, A.E., De Robertis, E.M., and Gehring, W.J. (1984) Nature 310, 70-71.
4. Quian, Y.Q., Billeter, M., Otting, G., Muller, M., Gehring, W.J., and Wuthrich, K. (1989) Cell 59, 573-580.
5. Hoey, T., Warrior, R., Manak, and Levine, M. (1988) Mol. Cell Biol. 8, 4598-4607.
6. Hanes, S.D. and Brent, R. (1989) Cell 57, 1275-1283.
7. Treisman, J., Gonczy, P., Vashishtha, M., Harris, E. & Desplan, C. (1989) Cell 59, 553-562.
8. Kornberg, T. (1981) Proc. Natl. Acad. Sci. USA 78, 1095-1099.
9. Patel, N.H., Martin-Blanco, E., Coleman, K.G., Poole, S.J., Ellis, M.C., Kornberg, T.B. & Goodman, C.S. (1989) Cell 58, 955-968.
10. Poole, S.J., Kauvar, L.M., Drees, B., and Kornberg, T. (1985) Cell 40, 37-43.
11. Martin-Blanco, E., Kornberg, T., unpublished.
12. McPherson, A. (1982). Preparation and Analysis of Protein Crystals, Wiley, New York.
13. Jordan, S.R., Whitcombe, T.V., Berg, J. M., and Pabo, C.O. (1985) Science 230, 1383-1385.
14. Matthews, B. (1968) J. Mol. Biol. 33, 491-497.