Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 6587–6591, September 1989
Biochemistry

Crystal structure analysis of auromomycin apoprotein (macromomycin) shows importance of protein side chains to chromophore binding selectivity

PATRICK VAN ROEY*† AND TERRY A. BEERMAN‡

*Medical Foundation of Buffalo, Inc., Buffalo, NY 14203; and ‡Grace Cancer Drug Center, Roswell Park Memorial Institute, Buffalo, NY 14263

Communicated by Herbert A. Hauptman, June 12, 1989 (received for review April 28, 1989)

ABSTRACT **The crystal structure of macromomycin, the apoprotein of the antitumor antibiotic auromomycin, has been determined and refined at 1.6-Å resolution. The overall structure is composed of a flattened seven-stranded antiparallel β-barrel and two antiparallel β-sheet ribbons. The barrel and the ribbons define a deep cleft that is the chromophore binding site. The cleft is very accessible and in this structure is occupied by two 2-methyl-2,4-pentanediol and two water molecules. The overall shape of the binding site is similar to that of the analogue actinoxanthin. Highly specific side chains that are not conserved between different analogues extend into the binding site and may be important to the chromophore binding specificity.**

Macromomycin (MCR) is the apoprotein (M_r = 11,000) (1) of the antitumor antibiotic auromomycin (AUR) (2), which is isolated from the culture filtrate of *Streptomyces macromomyceticus* (3). AUR is one of a large group of protein antibiotics isolated from different strains of *Streptomyces* that also includes neocarzinostatin (NCS), the mechanism of action of which has been studied most extensively (4, 5), and actinoxanthin (AXN), for which a partially refined crystal structure has been reported (6).

MCR carries and protects a nonprotein chromophore (Chr, termed AUR–Chr, M_r = 667) (7), which is the cytotoxic and mutagenic component of AUR. AUR–Chr is noncovalently bound to MCR and can be extracted with organic solvents such as methanol. Isolated AUR-Chr is very labile to heat and UV-irradiation, especially in aqueous solution. AUR is capable of inhibiting cell growth by causing strand breaks in DNA (8). AUR–Chr binds to DNA as a weak intercalator and causes single- and double-strand breaks (9, 10). The mechanism of action involves a free-radical intermediate (11), the production of which is stimulated by reducing agents, mainly sulfhydryls. AUR–Chr has not yet been identified but the highly unusual molecular structure of NCS–Chr ($C_{35}H_{33}NO_{12}$; M_r = 659) (12) has been determined from spectroscopic studies.

AUR–Chr and NCS–Chr act in similar fashion but differ substantially in specific details of the mechanism (13, 14). AUR–Chr is more active than NCS–Chr in damaging DNA (15). AUR has extremely high cytotoxicity, especially in whole-cell systems where its cytotoxicity levels are three orders of magnitude greater than those of NCS [D_{10} values (concentration of drug required to kill 90% of the total population of cells) are 25 pM and 15 nM, respectively] (15, 16). AUR is active in the absence of reducing agents and is only stimulated 5-fold by sulfhydryls (17). NCS requires reducing agents and is stimulated 1000-fold by them in cell-free DNA-nicking experiments. NCS scissions occur primarily at thymidine and adenosine with the preferred sequences 5′-ATG-3′ and 5′-TTT-3′, whereas AUR nicks preferentially at guanosine in the sequences 5′-TGT-3′ and 5′-AGG-3′ (18). It is therefore likely that the Chrs have related core structures but also that they have significantly different functional groups. Although detailed studies of the mechanism of action of AXN have not been reported, available evidence indicates that it belongs to the same class of compounds as AUR and NCS.

The apoprotein MCR is not involved in the interaction with DNA and is not known to enter the target cell. How the Chr is released, transported across the cell membrane, and stabilized within the cell is not yet known. The apoproteins of AUR, NCS, and AXN are similar in size, 108–113 amino acids, and have about 45% sequence homology (19). However, AUR–Chr is a poor ligand for NCS apoprotein (17) and MCR does not bind NCS–Chr well (20). Therefore, the overall folding of the apoprotein is expected to be similar but the geometry or composition of the binding site should show differences that are related to the differences in the structures and activities of the Chrs.

The high cytotoxicity, the mechanism of action, and the sequence specificity of AUR and its analogues make these compounds of special interest for the development of antitumor agents and for the study of DNA-damaging processes in general. The crystal structure analysis of MCR and AUR is under investigation to obtain detailed information about the structural and chemical properties of AUR, including the identification of AUR–Chr and its interaction with MCR and DNA. In this paper we report the crystal structure of the apoprotein MCR.

MATERIALS AND METHODS

MCR was purified from lyophilized culture filtrate as described (8). Purification of AUR yields two partially separated fractions, one active (AUR) and one inactive (MCR). Activity measurement of the AUR fraction gave D_{10} values of about 25 pM. Only the MCR fraction was used for this crystallographic study. Crystallization experiments for MCR have produced several crystal forms (21). Crystals of the best-diffracting form were obtained by hanging-drop vapor-diffusion techniques from a solution at 8 mg/ml in 25 mM Tris·HCl, 5 mM $CaCl_2$, 35% (wt/vol) 2-methyl-2,4-pentanediol (MPD) (pH 8.0), equilibrated against 68% MPD in 25 mM Tris·HCl (pH 8.0). The crystals belong to the monoclinic space group $P2_1$, with cell dimensions a = 36.29(1) Å, b = 35.54(1) Å, c = 38.04(1) Å, β = 99.59(2)°. One molecule per asymmetric unit (V_M = 2.20 Å^3) corresponds to a solvent content of approximately 44% (22). A full report on the structure determination and refinement will be published elsewhere. The structure was determined by multiple isomorphous replacement methods. Intensity data for the native

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: AUR, auromomycin; AXN, actinoxanthin; Chr, chromophore; MCR, macromomycin; MPD, 2-methyl-2,4-pentanediol; NCS, neocarzinostatin.

†To whom reprint requests should be addressed.

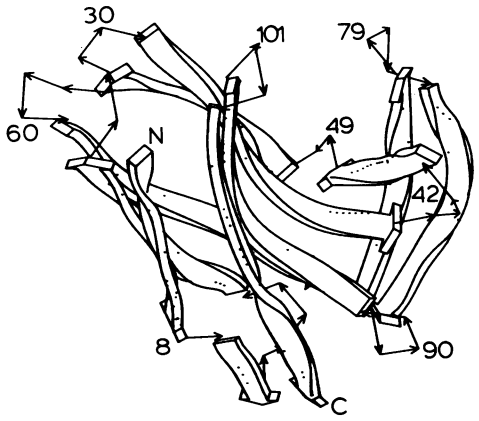

FIG. 1. Folding of macromomycin. Ribbons represent β-sheet structural components. The seven-stranded β-barrel is on the left and the two β-sheet ribbons are on the right. N and C indicate the amino and carboxyl termini, respectively. This figure was prepared using program ARPLOT (29).

and two heavy-atom derivatives were measured using one crystal for each data set. The native and the $HgCl_2$-derivative data were measured on an Enraf–Nonius model CAD4 diffractometer. The $K_2Pt(NO_2)_4$-derivative data were measured on a Rigaku model AFC5 diffractometer. Anomalous data were measured for both derivatives and featured prominently in the phasing because both are single-site derivatives. Heavy-atom phasing at 3.0-Å resolution led to the following parameters (for definition of terms, see ref. 23): figure-of-merit, 0.84. For Pt derivative: $x = 0.2051$; $y = 0.2500$; $z = 0.5166$; $B = 37.85$ Å^2; occupancy = 81.47 electrons; phasing power = 2.56; R_{cullis}(centric) = 0.43. For Hg derivative: $x = 0.5789$; $y = 0.0680$; $z = 0.8702$; $B = 22.83$ Å^2; occupancy = 28.75 electrons; phasing power = 1.16; R_{cullis}(centric) = 0.68. The resulting Fourier map allowed unambiguous location of 106 of the 112 residues, including most side chains. The structure was refined using restrained least-squares methods (program PROFFT) (24, 25) in combination with model building (program FRODO) (26, 27). Over 10 rounds of refinement, the resolution limit was gradually increased from 3.0 to 1.6 Å. The structure refined to a conventional R value of 0.152 for the 10,831 data [$F > 3\sigma(F)$, where F is the observed structure factor amplitude and σ is its estimated standard deviation] that were used in the refinement and 0.179 for all 12,674 data between 8.0- and 1.6-Å resolution. Root-mean-square deviations from ideality during the last cycle are as follows: bond lengths, 0.024 Å; angle distances, 0.054 Å; bond angle, 2.6°; peptide planarity, 0.023 Å; and hydrogen bonds, 0.193 Å. Four residues were refined with disordered side chains. The solvent structure that could be refined includes at this time two MPD molecules and 76 water molecules. Atomic coordinates have been deposited with the Brookhaven Protein Data Bank (28) (Chemistry Department, Brookhaven National Laboratory).

RESULTS AND DISCUSSION

The structure of MCR, shown schematically in Fig. 1, consists of one ordered domain and two external loops, formed by residues 36–46 and 72–86. The ordered domain is a flattened seven-stranded antiparallel β-barrel. Its folding is similar to that of the C domain of immunoglobulin (30). The loops form twisted antiparallel β-ribbons that partially overlap at the base of the barrel but run perpendicular to one another. The β-barrel and the two ribbons define a deep U-shaped cleft that, based on analogy of the structure with those of AXN and NCS, is thought to be the Chr-binding site. Preliminary reports on the study of NCS (31) have indicated the presence of residual electron density in its cleft and several dyes,§ including the intercalator thiolatoterpyridine platinum, have been shown to bind to apo-NCS in the cleft. The end turns of the ribbons have higher thermal parameters than the rest of the molecule. The main chain atoms of residues Glu-40 and Pro-41 have B values of about 31 Å^2 and those of Asp-79 have B values of about 40 Å^2. Except for the terminal residues Ala-1 and Ala-112, the B values of the main chain atoms of all other residues are less than 18 Å^2. This indicates that the two ribbons are the most flexible part of the molecule and that they could have different positions and conformations when the Chr is bound. The cleft is open on three sides and has approximate dimensions of 12 × 9 × 6 Å within the boundary of the protein. Using a high estimate of 1.5 g/cm^3 for its density, AUR-Chr requires about 700 Å^3 of space. Changes in the positions of the ribbons to close the binding site would reduce further the approximately 650 Å^3 space available in the open form observed in this structure. It is therefore impossible to completely enclose a compound the size of AUR–Chr. This may explain the fact that the Chr is easily extracted with methanol although it has a relatively high binding affinity for the apoprotein ($K \approx 10^{-10}$ M). This binding-site geometry and accessibility of bound Chr is in contrast with most binding proteins, such as the periplasmic sugar- and amino acid-binding proteins (32, 33), which are able to completely enclose their ligands and have clearly defined hinge regions or long flexible loops for this purpose.

Fig. 2 shows the complete structure of MCR and a detailed view of the binding site is shown in Fig. 3. [Figs. 2–5 and 7 have been prepared using a modified version of the program PLUTO by W. D. S. Motherwell and W. Clegg (1978).] The cleft is formed by the residues 32–48, 51–53, 73–81, and 94–108. The base of the cleft is very flat and consists of four strands of antiparallel β-sheet. All functional side chains are

§Sieker, L. C., American Crystallographic Association Meeting, June 22–27, 1986, Hamilton, Ontario, Canada (abstr. PB67).

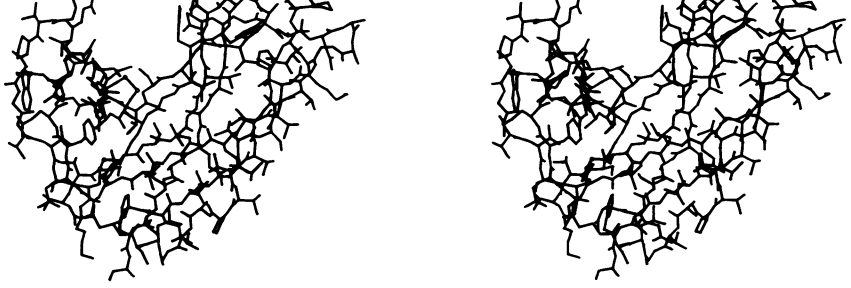

FIG. 2. Stereo diagram showing the structure of MCR illustrating the overall size and shape of the binding site relative to the size of the protein.

FIG. 3. Stereo diagram showing the residues (bold) that border the binding site superimposed on the α-carbon tracing (fine line). Only the side chains that extend into the binding site have been included.

located on the side walls of the cleft. They include His-32, Val-38, Ile-44, Ser-51, Asp-53, Ser-94, Asp-100, and the disulfide bridge Cys-36-Cys-46. These residues are distributed in a very organized manner along the walls of the binding site. The polar neutral residues (residues 36, 46, 51, and 94) are close to the base of the binding site, the charged residues (residues 32, 53, and 100) are higher up at the barrel side of the cleft, and the hydrophobic residues (residues 38 and 44) are exposed on the ribbon side. His-32 is held parallel to the β-sheet of the barrel by salt bridges with the two aspartate side chains. In this structure, the binding site contains two MPD molecules and two water molecules. As illustrated in Fig. 4, the MPD molecules form hydrogen bonds with the main-chain carbonyl group of His-32 (MPD-114) and $O\gamma$-hydroxyl group of Ser-51 (MPD-115), with O–O distances of 3.04 Å and 2.78 Å, respectively. These MPD molecules, which have additional hydrogen bonds with water molecules, partially substitute for the missing Chr. The two water molecules (O-120 and O-166) are located between the protein and MPD molecules.

MCR, AXN, and NCS have similar overall structures, including a well-defined binding-site cleft. Coordinates of the structure of AXN, partially refined at 2.5-Å resolution, are available (28), allowing comparison of the overall folding and the geometry of the binding sites of AXN and MCR. Fig. 5 shows the comparison of the α-carbon tracings of MCR and AXN after least-squares fitting of the main-chain atoms. The average distance between corresponding main-chain atoms in the two structures is 1.62 Å, with a maximum of 6.54 Å for the α-carbon of residue 102. This overlap shows that the MCR binding site is somewhat deeper and somewhat more closed at the top than that of AXN. However, differences are not greater than about 1.0 Å except at the turn between residues 98 and 103. This loop is more over the binding site in MCR, where it does not follow the β-barrel geometry as closely as in AXN.

If the binding sites have similar shapes, then the reported binding selectivity must be a result of differences in the functional groups that are involved in the protein–Chr interactions. Fig. 6 lists the sequences of the sections of MCR that

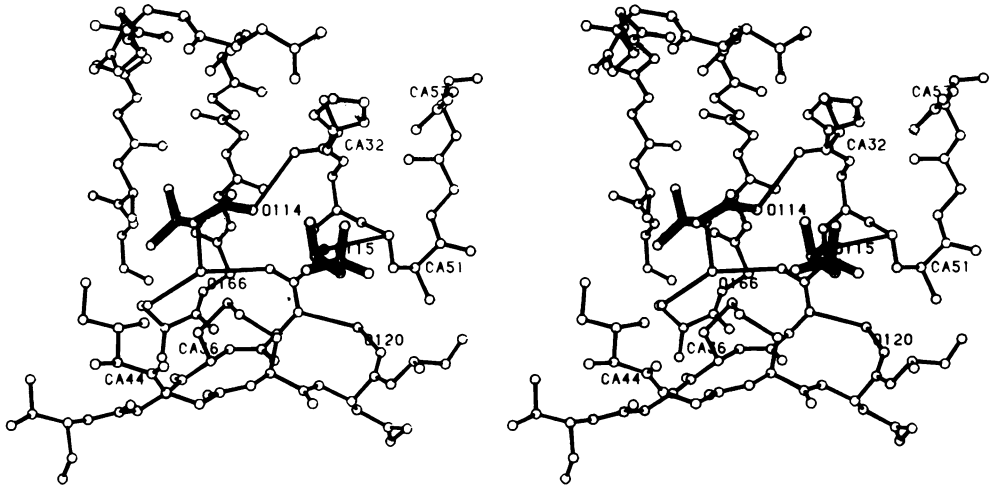

FIG. 4. Stereo diagram showing the location of the two MPD molecules (numbered 114 and 115) and the two water molecules (120 and 166) in the binding site and their hydrogen-bonding contacts.

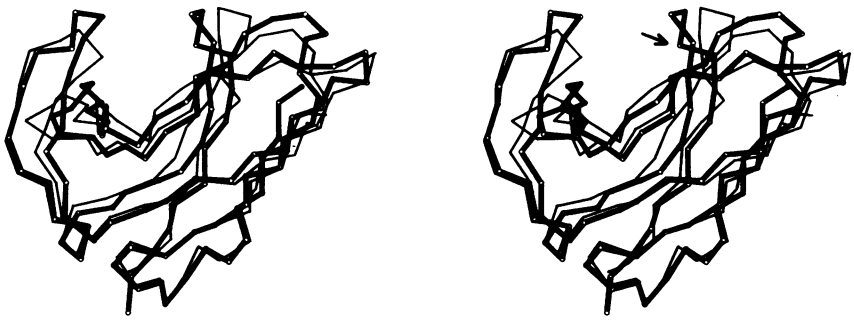

FIG. 5. Stereo diagram showing the comparison of the α-carbon tracings of MCR (bold) and AXN (fine line). MCR residue 102, which is the furthest from the corresponding residue in AXN, is indicated with an arrow in the right-hand view.

```
MCR  -His32 Val Gly Gln Cys Ala Val Val Glu Pro Gly Val Ile Gly Cys Asp47-
AXN  -Tyr30 Ile Ala Gln Cys Ala Pro Val Gly  -  Gly Gln Asn Ala Cys Asn44-
NCS  -Asp33 Val Gly Gln Cys Ala Trp Val Asp Thr Gly Val Leu Ala Cys Asn48-

MCR  -Ser51 Thr Asp Val Thr55-    -Val75 Val Gly Ala Asp Gly80-
AXN  -Ala48 Thr Ser Phe Thr52-    -Gly71 Gln Thr Pro Ser Gly76-
NCS  -Phe52 Ser Ser Val Thr56-    -Gly75 Phe Leu Phe Asp Gly80-

MCR  -Cys93 Ser Ala Gly Leu Gly Ser Asp Ser Gly Glu Gly  -  Ala Ala106-
AXN  -Cys89 Asn Leu Gly Ala Gly  -  Asn Ser Gly Leu Asn Leu Gly His102-
NCS  -Cys93 Gln Val Gly Leu  -  Ser Asp Ala Ala Gly Asn Gly Pro Glu106-
```

FIG. 6. Binding-site residues in MCR and the corresponding residues in the sequences of AXN and NCS. Boxes indicate residues that have side chains penetrating into the binding sites of MCR and AXN and thus provide contact points with the Chr.

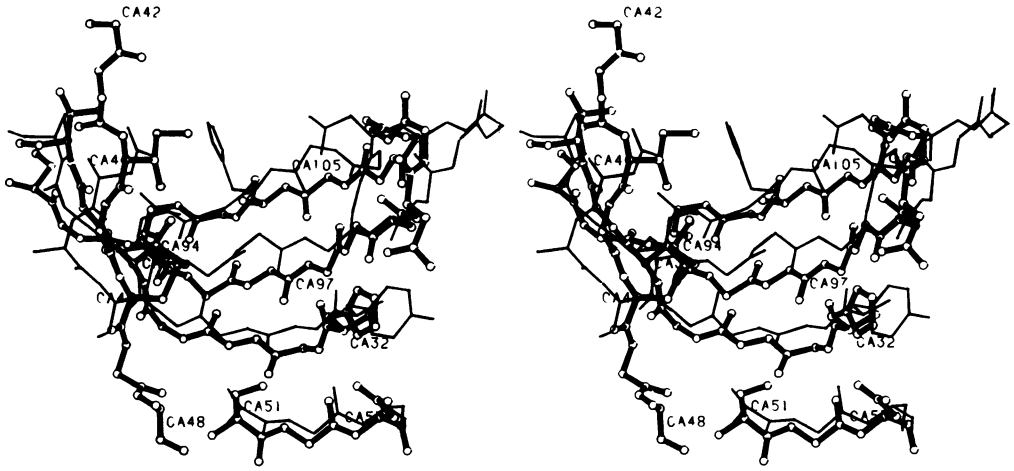

FIG. 7. Stereo diagram showing the comparison of the residues that border the binding sites of MCR (bold) and AXN (fine line). Note the substitutions of His-32 by Tyr-30 (right), Ile-44 by Asn-41 (left), and Ala-106 by His-102 (central). Loop 74 to 83, which is part of the binding site but would partially obstruct this view, is omitted for the sake of clarity.

border the binding site and the corresponding residues in AXN and NCS, aligned according to the highest sequence homology (18). Fig. 7 shows the comparison of the binding site residues of MCR and AXN after least-squares fitting of the main-chain atoms. Although the overall shapes of the binding sites are similar, the locations of functional side chains along the binding sites are very different. Most significant are the substitutions of the residues His-32, Ile-44, Asp-53, Ser-94, Gly-104, and Ala-106 in MCR by Tyr-30, Asn-41, Ser-50, Asn-90, and His-102 in AXN, respectively. These are residues of very different size and polarity. AXN does not have the separation of the charged and hydrophobic residues at the opposite sides of the binding site that is observed for MCR. It also has a large charged residue (His-102) in a location where MCR only has a small hydrophobic residue. These differences must reflect corresponding differences in the Chr structures. On the basis of the sequence comparisons in Fig. 6, it appears that analogous differences will be found for the NCS-binding site.

The structure of AUR needs to be determined to allow the

complete analysis of the interaction between AUR–Chr and MCR. But the information obtained from the analysis of the structure of MCR and its comparison with that of the apoprotein of AXN already shows that the binding sites of these proteins are much more open and accessible than those found for most binding proteins and that the nature of the residues involved in the binding of the Chr rather than the overall shape of the binding site is the cause for the selectivity of the Chr binding of these protein carriers.

The structure determination of MCR was completed by P.V.R. during a leave-of-absence at the Howard Hughes Medical Institute (Houston, TX). The generous hospitality, encouragement, and advice provided by Dr. F. A. Quiocho of the Structural Biology Laboratory at the Howard Hughes Medical Institute is gratefully acknowledged. This research was supported by Public Health Service Grants CA-34769 (P.V.R.) and CA-28495 (T.A.B.) from the National Cancer Institute, Department of Health and Human Services.

1. Woynarowski, J. M. & Beerman, T. A. (1980) *Biochem. Biophys. Res. Commun.* **94,** 769–776.
2. Beerman, T. A. (1978) *Biochem. Biophys. Res. Commun.* **83,** 908–914.
3. Chimura, H., Ishizuka, M., Hamada, M., Hori, S., Kimura, K., Iwanaga, J., Takeuchi, T. & Umezawa, H. (1968) *J. Antibiot.* **21,** 44–49.
4. Goldberg, I. H. (1986) *Basic Life Sci.* **38,** 231–244.
5. Kappen, L. S., Chen, C.-Q. & Goldberg, I. H. (1988) *Biochemistry* **27,** 4331–4340.
6. Pletnev, V. Z., Kuzin, A. P., Trakhanov, S. D. & Kostetsky, P. V. (1982) *Biopolymers* **21,** 287–300.
7. Naoi, N., Miwa, T., Okazaki, T., Watanabe, K., Takeuchi, T. & Umezawa, H. (1982) *J. Antibiot.* **35,** 806–813.
8. Beerman, T. A. (1979) *Biochim. Biophys. Acta* **564,** 361–371.
9. Kalvin, D. M., Huang, C.-H., Lischwe, M. A., Peters, E. H., Prestayko, A. W. & Crooke, S. T. (1981) *Cancer Chemother. Pharmacol.* **7,** 41–50.
10. Suzuki, H. & Tanaka, N. (1983) *J. Antibiot.* **36,** 575–582.
11. Suzuki, H., Kirini, Y. & Tanaka, N. (1983) *J. Antibiot.* **36,** 583–587.
12. Edo, K., Mizugaki, M., Koide, Y., Seto, H., Furihata, K., Otake, N. & Ishida, N. (1985) *Tetrahedron Lett.* **26,** 331–334.
13. Kappen, L. S., Goldberg, I. H. & Samy, T. S. A. (1979) *Biochemistry* **18,** 5123–5127.
14. Suzuki, H., Miura, K., Kumada, Y., Takeuchi, T. & Tanaka, N. (1980) *Biochem. Biophys. Res. Commun.* **94,** 255–261.
15. Grimwade, J. E. & Beerman, T. A. (1986) *Mol. Pharmacol.* **30,** 358–363.
16. Rauscher, F. J., III (1987) Ph.D. Dissertation (State Univ. of New York at Buffalo, Buffalo, NY).
17. Kappen, L. S., Napier, M. A., Goldberg, I. H. & Samy, T. S. A. (1980) *Biochemistry* **19,** 4780–4785.
18. Takeshita, M., Kappen, L. S., Grollman, A. P., Eisenberg, M. & Goldberg, I. H. (1981) *Biochemistry* **20,** 7599–7606.
19. Samy, T. S. A., Hahm, K. S., Modest, E. J., Lampman, G. W., Keutman, H. T., Umezawa, H., Herlihy, W. C., Gibson, B. W., Carr, S. A. & Biemann, K. (1983) *J. Biol. Chem.* **258,** 183–191.
20. Kappen, L. S. & Goldberg, I. H. (1980) *Biochemistry* **19,** 4786–4795.
21. Van Roey, P., Weeks, C. M., DeJarnette, F. E., Beerman, T. A. & McHugh, M. (1981) *J. Mol. Biol.* **153,** 183–185.
22. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497.
23. Blundell, T. L. & Johnson, L. N. (1976) *Protein Crystallography* (Academic, London).
24. Hendrickson, W. A. (1985) *Methods Enzymol.* **115,** 252–270.
25. Finzel, B. C. (1987) *J. Appl. Crystallogr.* **20,** 53–55.
26. Jones, T. A. (1985) *Methods Enzymol.* **115,** 157–171.
27. Pflugarth, J. W., Saper, M. A. & Quiocho, F. A. (1984) in *Methods and Applications in Crystallographic Computing,* eds. Hall, S. & Ashiaka, T. (Clarendon, Oxford), pp. 404–407.
28. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542.
29. Lesk, A. M. & Hardman, K. D. (1985) *Methods Enzymol.* **115,** 381–390.
30. Richardson, J. S. (1981) *Adv. Protein Res.* **34,** 168–339.
31. Sieker, L. C. (1981) Ph.D. Dissertation (Univ. of Washington, Seattle).
32. Quiocho, F. A. & Vyas, N. K. (1984) *Nature (London)* **310,** 381–386.
33. Sack, J. S., Trakhanov, S. D., Tsigannik, I. H. & Quiocho, F. A. (1989) *J. Mol. Biol.* **206,** 193–207.