X-ray Crystallographic Studies of Candida albicans Dihydrofolate Reductase

HIGH RESOLUTION STRUCTURES OF THE HOLOENZYME AND AN INHIBITED TERNARY COMPLEX*

Marc Whitlow‡, Andrew J. Howard§, David Stewart, and Karl D. Hardman¶
From the Genex Corporation, Protein Engineering Department, Gaithersburg, Maryland 20877

Lee F. Kuyper†, David P. Baccanari, Mary E. Fling, and Robert L. Tansik
From Glaxo Wellcome Inc., Research Triangle Park, North Carolina 27709

The recent rise in systemic fungal infections has created a need for the development of new antifungal agents. As part of an effort to provide therapeutically effective inhibitors of fungal dihydrofolate reductase (DHFR), we have cloned, expressed, purified, crystallized, and determined the three-dimensional structure of Candida albicans DHFR. The 192-residue enzyme, which was expressed in Escherichia coli and purified by methotrexate affinity and cation exchange chromatography, was 27% identical to human DHFR. Crystals of C. albicans DHFR were grown as the holoenzyme complex and as a ternary complex containing a pyrroloquinazoline inhibitor. Both complexes crystallized with two molecules in the asymmetric unit in space group P21. The final structures had R-factors of 0.199 at 1.85-A resolution and 0.155 at 1.60-A resolution, respectively. The enzyme fold was similar to that of bacterial and vertebrate DHFR, and the binding of a nonsynergistic diamino(pyrazolokinazoline) inhibitor and the interactions of NADPH with protein were typical of ligand binding to other DHFRs. However, the width of the active site cleft of C. albicans DHFR was significantly larger than that of the human enzyme, providing a basis for the design of potentially selective inhibitors.

Inhibition of the enzyme dihydrofolate reductase (DHFR) is a well-established mechanism for drug action (1–3). Commonly cited clinical agents whose therapeutic effects stem from DHFR inhibition include methotrexate (4), pyrimethamine (5, 6), and trimethoprim (7). Methotrexate is a nonselective inhibitor employed in cancer chemotherapy, and the latter two drugs are used for antimicrobial therapy: pyrimethamine is an antimalarial agent that selectively inhibits the DHFR of the protozoan Plasmodium falciparum, and trimethoprim is a broad spectrum antibacterial inhibitor that shows a striking selectivity for DHFR from many bacteria versus the enzyme from the human host.

DHFR has been studied extensively (8, 9). High resolution x-ray crystal structures of the enzyme from bacterial and vertebrate sources have been reported, and NMR studies of the enzyme in solution have provided additional structural information. This wealth of structural data has been utilized in efforts to design new inhibitors and to understand the molecular details of enzyme-inhibitor interactions that determine affinity and selectivity (10–13).

This paper focuses on DHFR from the fungus Candida albicans. That organism and the opportunistic fungi Cryptococcus neoformans and Aspergillus fumigatus are of increasing clinical importance, especially in immunocompromised patients (14–17). The incidence of systemic fungal infections has risen considerably in recent years. C. albicans is now the fifth leading cause of microbial infection in the hospital setting (18), and opportunistic fungal infections present a major problem to AIDS patients (19, 20). Although several antifungal agents, such as amphotericin B and theazole class of drugs, are currently available (14, 16), there is clearly a critical need for the development of new antifungal agents and novel approaches to antifungal therapy. In our approach to develop new therapeutically useful inhibitors of fungal DHFR, we have cloned, expressed, purified, crystallized, and determined the x-ray crystal structure of C. albicans DHFR. Details of those studies are described herein.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification of C. albicans DHFR—A 4.9-kilobase pair HindIII fragment of genomic C. albicans DNA was cloned and detected by its ability to mediate resistance to trimethoprim in Escherichia coli (21). A 2.3-kilobase pair HindIII-StuI region was identified by deletion analysis as the carrier of DHFR gene (DFR1) by retention of trimethoprim resistance. Sequence determination of this region (GenBank™ accession number U84887) revealed an open reading frame whose deduced amino acid sequence was indicative of a DHFR and whose amino-terminal sequence corresponded to the determined amino-terminal sequence of the enzyme purified from C. albicans (21).

C. albicans DHFR, expressed in E. coli BL21(DE3), p1889, was purified to homogeneity using procedures similar to those described previously (21). Briefly, 28-g cells were lysed by the method of Godson and Sinheimer (22) using lysosome and Brij 58. The crude lysate (133 ml containing 62,000 units of enzyme and 3.2 g of protein) was subjected to ammonium sulfate fractionation (40–90%). The precipitate (53,000 units and 2.0 g of protein) was redissolved in 50 mM sodium
phosphate buffer, 1 mM diithiothreitol, 0.1 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride, pH 7.0 (buffer A) and applied to a 2.5 × 12-cm methotrexate affinity column. The column was washed with buffer A and then with the same buffer containing 1 mM KCl (buffer B). The addition of 2 mM folic acid to buffer B eluted DHFR (33,000 units of enzyme) from the column. Buffer B was exchanged by dialysis against 20 mM KMES, 20% glycerol, pH 7.5, and the sample was applied to a 2.5 × 10-cm S-Sepharose (Pharmacia Biotech Inc.) column to remove folic acid. A KCl gradient eluted the DHFR (17,000 units and 36 mg of protein) at approximately 80 mM. The enzyme was concentrated on an Amicon YM10 membrane to 1.1 mg/ml. The purified enzyme migrated as a single band on SDS-polyacrylamide gel electrophoresis with an apparent molecular weight of 25,000. The UV spectrum of the pure DHFR had a sharp peak at 280 nm and a shoulder at 290 nm, but lacked any absorbance above 310 nm, suggesting the absence of folic acid.

DHFR Crystallization—The C. albicans DHFR protein was stored at 4 °C in its original buffer of 20 mM KMES, 1 mM diithiothreitol, 0.1 mM EDTA, 20% glycerol, pH 7.5. Prior to use, aliquots of DHFR were buffer exchanged with 20-fold molar excess of enzymatically reduced NADPH (Type III; Sigma) in the crystallization buffer (50 mM KMES, 1 mM NADPH, 1 mM diithiothreitol, pH 6.5) using Centricron-10 microconcentrator units (Amicon, Beverly, MA). The DHFR-NADPH solutions were concentrated to 17–20 mg/ml DHFR. Crystals were prepared using the hanging drop vapor diffusion technique at 4 °C (25) in Ligation cultures (Table I). The crystallization buffer contained 44% PEG-3350 in the crystallization buffer. The hanging drops contained 5 μl of the DHFR-NADPH solution and 5 μl of the reservoir solution.

The ternary complex between C. albicans DHFR, NADPH, and 1,3-diamino-7-(1-ethylpropyl)-7H-pyrrrolo[3,2-f]quinazoline (GW345) was crystallized in a manner similar to that used for the binary complex as described above. A 250 μM solution of the GW345 was prepared in 50 mM HCl, which was in turn used to prepare the crystallization buffer (50 μM GW345, 50 mM NADPH, 50 mM KMES, 1 mM diithiothreitol, pH 6.5). A 3-fold molar excess of GW345 and a 3-fold molar excess of NADPH were added to the original DHFR solution. The DHFR-NADPH-GW345 solution was stored at 4 °C overnight, buffer exchanged with the crystallization buffer, and concentrated to 17–20 mg/ml DHFR using a Centricron-10 microconcentrator unit. The hanging drops contained 5 μl of the DHFR-NADPH-GW345 solution and 5 μl of the reservoir solution. One or more successive cycles of macroseeding were required to produce crystals large enough for diffraction data acquisition (25).

X-Ray Data Collection and Reduction—X-Ray diffraction data were collected using a Siemens electronic area detector mounted on a Super V5 rotating anode X-ray source used to generate CuKα X-rays. Monochromatization was provided by a rotating anode, operating at 70 mA and 40 kV with a 0.3-mm Al filter. Data were collected using a Siemens electronic area detector mounted on a Super V5 rotating anode, operating at 70 mA and 40 kV with a 0.3-mm Al filter. Data were collected by the area detector and were recorded as a series of discrete frames, each comprising a 0.20 or 0.25° oscillation counted for between 30 and 110 s, depending upon the carriage angle. Usually 400–480 data frames, corresponding to 100° of crystal rotation about a vertical axis, were accumulated. Subsequently, the data were transferred to a Digital Equipment Corp. VAX 11/780 computer for processing. During the data collection, the area detector chamber was mounted 10 cm from the crystal. The carriage angle (2θ) varied from 0 to 30°, enabling the detector to intercept data from θ ≈ 1.58 Å, depending upon its position. Diffraction data collected by the area detector were recorded as a series of discrete frames, each comprising a 0.20 or 0.25° oscillation counted for between 30 and 110 s, depending upon the carriage angle. Usually 400–480 data frames, corresponding to 100° of crystal rotation about a vertical axis, were accumulated. Subsequently, the data were transferred to a Digital Equipment Corp. VAX 11/780 computer for processing. During the course of data collection, several data sets from different crystal orientations were recorded. The crystals were repositioned in the x-ray beam by adjustment of a goniostat φ motor set 45° from vertical. The crystals were very stable in the x-ray beam allowing collection of an extensively replicated data set at 1.60 Å from a single crystal. The XENGEN software package was used for data reduction (26).

Crystal Density Determination—The density of the C. albicans DHFR-NADPH crystals was determined using a density gradient column. A gradient maker was used to prepare a carbon tetrachloride (ρ = 1.59 g/ml) xylene (ρ = 0.88 g/ml) gradient in a 50-ml buret. The crystals were removed from the hanging drop, dried, and added to the top of the density gradient, and their final position was recorded. Density standards of 1.05, 1.02, and 1.00 g/ml were prepared and water. A small drop of each density standard was added to the top of the column and its final position recorded. A plot of the density of standards versus their position in the buret was prepared, and the crystal density was determined by interpolation.

Molecular Replacement—The starting molecular replacement model was based on the E. coli and mouse DHFR crystal structures (27, 28) and a sequence comparison with the C. albicans enzyme, see Fig. 1. The coordinates of E. coli DHFR were superimposed on those of the mouse DHFR using SUPPFIT. The molecular replacement model was based primarily on the mouse DHFR structure due to the higher sequence homology of the mouse and C. albicans proteins.

The model was constructed from the mouse DHFR structure using the following logic. Residues 1–3 of the mouse enzyme were omitted, because the C. albicans protein contained an insertion of one residue (Asn) between residues corresponding to 3 and 4 of the mouse enzyme. C. albicans DHFR had a Pro inserted between residues 12 and 13 of the mouse sequence; thus, residues 12 and 13 were removed from the model. Mouse DHFR residues Pro25 and Pro26 were removed from the model because Pro25 corresponded to an insertion relative to the fungal enzyme. Residues 43–47 of mouse DHFR form part of a loop structure and were not homologous with the corresponding residues of C. albicans DHFR and were thus not included in the model. Residues 76–111 in mouse DHFR (residues 65–97 in E. coli) compose a loop-helix-loop structure. Residues 88–97 of the mouse enzyme are structurally homologous to the corresponding residues of E. coli DHFR (residues 73–82), and residues 105–110 in mouse DHFR have sequence homology with residues 102–107 in C. albicans DHFR. These observations led us to include residues 78–87 and 98–104 in the model. Residues 139–147 in mouse DHFR showed little sequence homology with the E. coli or C. albicans enzymes and therefore were removed. Residues 175–184 in mouse DHFR are structurally homologous to residues 148–159 in E. coli DHFR. Although sequence homology was low between C. albicans and mouse DHFR for mouse residues 175–184, structural homology of this region for the E. coli and mouse proteins was relatively high, and the residues were included in the model. The following side chains were kept because they are either identical or nearly identical in C. albicans and mouse DHFR: Ile7, Val8, Ala9, Ile16, Pro23, Trp24, Leu27, Arg28, Glu29, Tyr33, Phe34, Thr38, Thr39, Asn48, Val50, Ile51, Met52, Arg54, Lys55, Thr56, Ser59, Thr60, Trp61, Glu62, Lys63, Arg65, Pro66, Leu67, Asp69, Thr70, Ser74, Asp75, Glu76, Ser77, Asn78, Val79, Tyr80, Ile81, Thr83, Asn84, Trp85, Val86, Pro87, Leu88, Arg90, Glu91, Tyr97, Leu98, Ser99, Arg100, Lys101, Val102, Lys103, Gln104, Val105, Val106, Ser107, Glu109, Arg110, Ser114, Val115, Ile116, Val117, Val118, Tyr120, Trp123, Phe124, His125, Asn126, Leu127, Val128, Thr129, Ile130, Tyr133, and Phe139. The NADPH molecule was also included in the model. In total, the model consisted of 845 atoms of the 1567 atoms in the mouse DHFR-NADPH crystal structure.

A self-rotation function was run on the C. albicans DHFR-NADPH complex data at 5–10, 4–10, and 3–10 Å using MERLOT (29). The initial molecular replacement solution was performed using the 845-atom search model, described above, and MERLOT (29). The rotational searches were performed using the Crowther fast rotation function (30) with a 5° grid size. Patterson cutoffs of 12, 14, and 16 Å were used with 5–10, 4–10 or 3–10 Å data. The three top rotation solutions were refined using the Lattman rotation function with a 1 Å grid at 3–10 Å, and each solution was subjected to translation searches using 3–10 Å data (see Table I).

The initial molecular replacement solution was refined using RT (a rigid-body rotation and translation program) at 4–10 Å and then at 3–10 Å. RT is similar to the program of Fujinaga and Read (31) except that it is based on the refinement programs of Hardman et al. (32) and uses the Winograd Fourier transform for P21.

Each molecule was refined independently using the restrained least squares procedures (33) modified by Finzel (34) to incorporate the fast Fourier algorithms of Ten Eyck and Agarwal (35, 36) in the program PROFOFT. A noncrystalllographic averaging of the 2Fo − Fe map was used in the early stages of refinement, using a set of programs written

2 K. D. Hardman, unpublished program.

3 M. Whitlow and K. D. Hardman, unpublished results.

4 M. Fujinaga and R. Read, unpublished program.
DHFR is 33 residues longer than protein and the two isozymes share 34% sequence identity.

Crystallization of C. albicans DHFR—

Crystal structures of Candida dihydrofolate reductase

| Resolution (Å) | Rotation peak a | Peak height (e) |
|----------------|-----------------|-----------------|
| A             | α   | β   | γ   | U | V | W |
| 3.0–10.0      | 96.2 | 94.2 | 189.6 | 0.00 | 0.50 | 0.08 | 3.4 |
|               | 0.78 | 0.50 | 0.34 | 3.3 |
| 4.55          | 71.0 | 336.0 | 0.34 | 0.50 | 0.92 | 5.4 |
|               | 45.9 | 124.0 | 43.0 | 0.96 | 0.50 | 0.40 | 3.3 |
| Cross         | 96.2 | 94.2 | 189.6 | 0.10 | 0.05 | 0.71 | 6.0 |
|               | 0.22 | 0.53 | 0.70 | 5.3 |
| Cross         | 96.2 | 94.2 | 189.6 | 0.17 | 0.04 | 0.99 | 5.0 |
|               | 0.03 | 0.10 | 0.78 | 5.0 |
|               | 0.11 | 0.89 | 0.78 | 5.0 |

a The refined rotation solutions from the Lattman rotation function at 3–10 Å.

by one of us (M. Whittow) for this structure solution. Model building was performed on an Evans and Sutherland PS330 using PSFRRODO version 6.4.

Structural Comparisons—Structural comparison of the two molecules of the asymmetric unit of C. albicans DHFR and comparison of the fungal enzyme structure with that of human DHFR employed the program O and a least squares fitting routine written by N. Srinivasan.α α-Carbons from five of the central β-strands (strands A, B, E, F, and H) and the B-helix were used to superimpose the human and C. albicans DHFR structures. The asymmetric unit of both crystal structures contained two molecules. The two molecules in the human DHFR structure are essentially identical, and molecule 1 was arbitrarily chosen for comparison to the C. albicans DHFR structure. However, molecule 1 of the fungal enzyme structure displays conformational distortions that may be due to its interactions with molecule 2. We therefore selected molecule 2 of the C. albicans DHFR structure for a detailed comparison with the human form of the enzyme.

RESULTS AND DISCUSSION

Comparison of C. albicans DHFR Amino Acid Sequence with Those of Other DHFRs—A 2.3-kilobase pair HindIII-SceI region of C. albicans genomic DNA was cloned and identified by its retention of trimethoprim resistance. Sequencing revealed an open reading frame without introns with a deduced amino acid sequence, which was indicative of DHFR and an amino-terminal sequence of the enzyme purified from C. albicans, see Fig. 1 (21). The deduced amino acid sequence differed at residue 84 from a reported sequence (37) in that the codon reported here encodes lysine, whereas the corresponding codon has a single base difference and encodes arginine. A silent nucleotide difference was observed in the third position of codon 71. Only trivial differences were seen when sequences outside of the genes were compared; these included 12 single nucleotide differences and one short insertion at ~261 base pair from the start codon in the sequence reported here relative to that of Daly et al. (37).

C. albicans DHFR gene encodes a 192-amino acid protein, which is 27% identical to the corresponding human enzyme (see Fig. 1). DHFR from the opportunistic organism Pneumocystis carinii contains 14 more residues than the C. albicans protein and the two isoforms share 34% sequence identity. C. albicans DHFR is 33 residues longer than E. coli DHFR, and 30% of the bacterial enzyme's sequence is identical to that of the fungal protein.

Crystallization of C. albicans DHFR—Crystals of the binary complex of C. albicans DHFR and NADPH grew as somewhat irregular shaped crystals with well defined, slightly curved faces. The largest of these grew to 0.6 × 0.3 × 0.1 mm in 30% PEG-3350. A 1.83 Å data set was collected from one of the DHFR-NADPH binary complex crystals. The space group of this crystal was P21, with unit cell dimensions of a = 77.20 Å, b = 67.57 Å, c = 38.66 Å, α = γ = 90.0°, and β = 93.06°. The data set was complete and the mean I/σ drop to two (D_s) at 1.86 Å (see Table II).

The density of the C. albicans DHFR-NADPH complex crystals was 1.18 gm/cm³ as determined by a carbon tetrachloride/xylene density gradient column. With one or two molecules in the asymmetric unit cell, the calculated solvent contents would be 72 or 43%, respectively. With two molecules in the asymmetric unit, the Matthews coefficient (38) was 2.27 Å³/dalton. A 4.1 σ self-rotation function peak was found at 158°, β = 142°, and γ = 22°, using data between 3–10 Å. Both the density of the crystals and the self-rotation peak suggested that there were two molecules in the asymmetric unit.

Molecular Replacement Solution and Refinement—Three possible rotation solutions were identified using the Crowther rotation function in MERLOT at three resolution ranges (3–10, 4–10, and 5–10 Å) and three Patterson cutoffs (12, 14, and 16 Å). The highest peaks were found using 3–10 and 4–10 Å data. The highest rotation peak was consistently found to be around α = 95°, β = 95°, and γ = 190°, with a peak height of 4.2–4.8 σ. The two other solutions were α = 45°, β = 70°, and γ = 335° and α = 45°, β = 125°, and γ = 45°, with peak heights of 3.5–4.4 and 3.5–4.2 σ, respectively. The three rotation solutions were refined using Lattman's rotation function at 3–10 Å resolution. Translation searches were run between each of the rotation solutions and their 2-fold symmetry related orientations, between the first and second rotation solutions, and between the first and third rotation solutions (see Table I). A consistent set of translation peaks were found for the first and second rotation solutions. The final molecular replacement solution was α = 96.2°, β = 94.2°, γ = 189.6° at x = 0.39, y = 0.00, z = 0.17 (first rotation solution) and α = 46.5°, β = 71.0°,
Crystal Structures of Candida Dihydrofolate Reductase

The molecular replacement solution was initially refined using rigid body refinement at 4–10 Å resolution and then at 3–10 Å resolution. After rigid body refinement the R-factor was 0.49 and the correlation coefficient was 0.39. Data between 1.85 and 10 Å was used in the restrained least squares refinement (see Fig. 3). Only residues 148–154 in molecule 2 had higher temperature factors in molecule 1 than in molecule 2. Regions had higher temperature factors in molecule 1 than in molecule 2. Therefore a total of 119 cycles of refinement and eight model building sessions, the final R-factor was 0.33%, the noncrystallographic averaging was discontinued. After a total of 119 cycles of refinement and eight model building sessions, the final R-factor for the C. albicans DHFR-NADPH complex was 0.199 (see Table III). A representative 2Fo – Fc electron density is shown in Fig. 2. Only residue 22 in both molecules and residue 141 in molecule 1 had ϕ-ψ values significantly outside the acceptable regions of a Ramachandran plot (39).

The final model contained two molecules of C. albicans DHFR, two NADPH and 162 water molecules. The following residues were assigned alternative side chain positions: Val5, Arg32, Ile33, Lys37, Thr41, Arg42, Asn47, Ile53, Ser55, and Val109 in molecule 1 and Val6, Arg7, Arg8, Ghu12, Asp105, Ile110, Ser125, Gln141, Ser142, Lys150, and Glu174 in molecule 2. The mean temperature factors for the two C. albicans DHFR molecules were 17.85 and 15.09 Å², respectively; for the two NADPH molecules were 21.15 and 14.10 Å², respectively; and for the 162 water molecules was 17.84 Å². The following four regions had higher temperature factors in molecule 1 than in molecule 2: residues 44–50, 90–107, 138–143, and 174–183 (see Fig. 3). Only residues 148–154 in molecule 2 had higher temperature factors than in molecule 1. Asn17 of molecule 1 was in close contact to Arg29 of symmetry-related molecule 1. In molecule 2 Asn17 was in close contact with Arg79 and Ghu120 of a symmetry-related molecule 2. There was a large contact surface between molecules 1 and 2 involving residues Met1, Ghu2, Asn3, Ile6, Ghu6, Ser6, Asn101, Ser104, and Ser125 in molecule 1 and Leu2, Lys2, Asn2, Asn123, Asn124, Ser125, Arg184, and Lys185 in molecule 2. The two loops containing Pro138, Ser139, Lys142, and Asp181 of molecule 2 were in close proximity to residues Asn3, Arg106, Ser128, His129, and Lys192 in a symmetry-related molecule. Molecule 2 was involved in far more intermolecular contacts than molecule 1, which may explain its lower mean temperature factor.

Comparison of the Two Molecules in the Asymmetric Unit—Superposition of the two molecules of the asymmetric unit of the fungal enzyme crystal structure using the program O (40) showed that the two molecules differed in conformation in the region of their interface. The root mean square deviation resulting from the superposition of the two molecules using all α-carbons was 1.22 Å. Omission of residues 1–5, 45–47, and 94–105 (via the lsq improve option in O with a cutoff distance of 1.5 Å) lowered the root mean square deviation to 0.55 Å. As shown in Fig. 4A, residues 1–5, 45–47, and 94–105 represent much of contact surface between the two molecules, suggesting that the differences are probably due to crystal packing forces. The overlaid molecules are shown in Fig. 4B. The active site regions and the binding geometry of inhibitors and cofactor in the two molecules were essentially identical. Comparison of each of the two molecules to human and chicken DHFR structures showed that molecule 2 of the fungal enzyme structure was more similar to the vertebrate protein structures, and it was therefore chosen for a more detailed comparison with the structure of human DHFR.

Structural Features of C. albicans DHFR and Comparison to Human DHFR—Structures of DHFR from E. coli (27), Lactobacillus casei (27), chicken (41), mouse (28), human (42), and P. carinii (43) are known, and all exhibit the same general fold that is dominated by an extended central β-sheet with flanking α-helices. The structure of C. albicans DHFR displays a similar fold (see Figs. 4 and 5).

Dihydrofolic acid binding site being identical and the remaining 3 compounds. Such differences might be used for the design of inhibitors that would selectively bind to the fungal protein and provide a margin of safety for the human host.

The fungal enzyme contained 192 amino acid residues compared with 187 for the human protein, and sequence identity was 27% (see Fig. 1). Sequence identity in the active site was significantly higher, with 9 of the 12 residues that compose the dihydrofolate binding site being identical and the remaining 3 being homologous (C. albicans/human: Ile8/Ile7, Met20/Leu22, Ghu22/Glu24, Ile133/Phe131, Phe36/Phe34, Met65/Met72, Thr58/Thr56, Ser81/Ser77, Ile69/Ile66, Leu69/Leu74, Arg8/Arg70, Ile112/Val115).

Despite the low overall sequence homology, the three-dimensional structures of the two proteins were strikingly similar. Seven strands of the central β-sheet and much of the helical portions of the enzymes closely superimposed, as depicted in Fig. 5 (A and B). Differences were observed in various exterior features of the structures. β-Strand G shows a significant disruption in the human enzyme, similar to that first reported for chicken DHFR (44). An even greater disruption was observed in the corresponding strand of the C. albicans enzyme; the disrupted region of β-strand G (residues 160–176) in C. albicans DHFR contained two turns of α-helical conformation, similar to that found in the recently reported structure of DHFR.

---

**Table II**

| Data set | Dmin (Å) | Dmax (Å) | R(Fo – Fc) | R(Fo – Fc) | θ of reflections | Completeness |
|----------|----------|----------|-------------|-------------|----------------|--------------|
| Native   | 1.83     | 1.86     | 6.64        | 7.04        | 140,620        | 31,520       | 0.903        |
| GW345    | 1.60     | 1.60     | 4.58        | 4.22        | 139,612        | 45,668       | 0.917        |

- a Dmin = smallest D spacing for which reflections were measured.
- b Dmax = spacing at which (I/s(I)) = 2.
- c R(Fo – Fc) = 2(I(I) – I/s(2I/I/s)).
- d R(Fo – Fc) = 2(I(I) – I/s(2I/I/s)).
- e Completeness = number of unique reflections measured/number of unique reflections possible to the stated resolution limit.

\[ γ = 336.0° \text{ at } x = 0.17, \ y = 0.467, \ z = 0.46 (\text{second rotation solution}). \]

---

**Table III**

| Data set | Native | GW345 |
|----------|--------|-------|
| Resolution (Å) | 1.85     | 1.60     |
| # of reflections | 31,520 | 45,668 |
| R-factor | 0.199 | 0.155 |
| Distances (Å) | 0.03 | 0.20 |
| Bonds (1–2) | 0.04 | 0.038 |
| Angles (1–3) | 0.05 | 0.040 |
| Intraplanar (1–4) | 0.06 | 0.040 |
| Planar groups (Å) | 0.03 | 0.022 |
| Chiral centers (Å) | 0.30 | 0.294 |
| Nonbonding contacts (Å) | 0.20 | 0.172 |
| Single torsion | 0.20 | 0.161 |
| Multiple torsion | 0.20 | 0.143 |
| Possible hydrogen bonds | 0.20 | 0.127 |
| Torsion angles (degree) | 6.0 | 4.6 |
| Planar (ω) | 1.0 | 17.6 |
| Staggered (10.0) | 3.8 | 14.9 |
| Orthonormal (10.0) | 32.4 | 31.0 |

- a R-factor = \( \frac{\Sigma|F_o|-F_c|}{\Sigma F_c} \).
from *P. carinii* (43).

The most dramatic difference between the *C. albicans* and human DHFR structures appeared in an extended loop region formed by residues 79–90 of the *C. albicans* protein and residues 77–87 of human DHFR. In the human enzyme structure, as well as all other structures of DHFR reported to date, this loop is packed against the C-helix in a preceding section of the sequence (residues 54–60 in the human enzyme). In contrast, that loop in the *C. albicans* DHFR structure is folded away from the C-helix and forms a short $\beta$-structure, increasing the core $\beta$-sheet of the structure to nine strands. The effects of this conformational difference appear to extend to the active site regions of the two proteins and involve the position of the C-helix. This helix forms a portion of one side of the active site, and its position relative to the active site could have an impact on substrate and inhibitor affinity. Because of this difference in the C-helix position, the width of the active site cleft in the fungal enzyme is significantly greater than that of the human protein. The distance between $\alpha$-carbons of key active site residues Glu$^{32}$ and Ile$^{62}$ that are on opposite sides of the cleft in the *C. albicans* enzyme was 16.4 Å, 1.4 Å greater than the corresponding (Glu$^{30}$ and Ile$^{60}$) distance in human DHFR. The difference in cleft width increases to 2.0 Å when measured between Glu$^{32}$ and Gln$^{64}$ of the fungal protein and the related residues of the human enzyme. Thus, despite the highly conserved residue composition of the active site, this observed difference in cleft geometry might provide a basis for the design of selective inhibitors. A similar difference was reported in a comparison of chicken and *E. coli* DHFR and was invoked as a source for the enzyme selectivity of the antibacterial drug trimethoprim (41, 45).

**Fig. 2.** Stereoview of a representative $2F_o - F_c$ electron density map of the *C. albicans* DHFR-NADPH structure contoured at 1 $\sigma$. Residues Ile$^8$ to Val$^{10}$ in molecule 1 are shown. This figure was prepared using XtalView (50), with minor modification to the postscript file.

**Fig. 3.** Variation in temperature factor (B-value) along the polypeptide chain in the *C. albicans* DHFR-NADPH structure. Molecule 1 is shown in solid lines and molecule 2 is shown in dashed lines. This figure was prepared using XtalView (50), with minor modification to the postscript file.

The ternary complex of the *C. albicans* DHFR with Compound GW345—A 1.60-Å data set was collected from one of the DHFR-NADPH-GW345 ternary complex crystals. The data set was 92% complete with the mean I over $\sigma$ drop to two (D $\sigma^2$) at the edge of the data set (see Table II). After collecting data on a number of *C. albicans* DHFR ternary complex crystals, an average $P2_1$ unit cell was determined to be $a = 76.91$ Å, $b = 67.28$ Å, $c = 38.49$ Å, $\alpha = \gamma = 90.0^\circ$, and $\beta = 93.07^\circ$.

The starting model for the refinement of the DHFR-NADPH-GW345 ternary complex was derived from another ternary complex, containing a different inhibitor, that had been refined from the holoenzyme *C. albicans* DHFR structure at 1.7 Å to an R-factor of 0.158. Prior to fitting GW345, 11 cycles of restrained least squares refinement were performed, electron density maps were calculated and then GW345 was fit to the difference
map (see Fig. 6). After six more rounds of refinement, for a total of 54 cycles of refinement, the final R-factor was 0.155 at 1.6 Å resolution (see Table III). The Ramachandran plot (39) of the final model had only a few residues that lie just outside of the acceptable regions (see Fig. 7). The residues of the binary complex that displayed abnormal ϕ-ψ values (residues 22 and 141) have normal conformations in the more highly refined structure of the ternary complex.

The final structure contained 333 water molecules with a mean temperature factor of 22.67 Å², 182 more than were found in the holoenzyme structure. The following residues in the C. albicans DHFR ternary structure were assigned alternative side chain positions: Arg³⁰, Ile³³, Thr⁴¹, Arg⁴², Ile⁵³, Arg⁷⁹, Ser⁸⁰, Glu⁸⁴, Val¹⁰⁹, Leu¹²¹, Leu¹³¹ and Glu¹³⁶ in molecule 1 and Lys¹, Arg³⁴, Ly¹⁶⁵, Arg⁷⁹, Ile⁶⁶, Ser¹⁰⁴, Asp¹⁰⁵, Ile¹¹¹, Gln¹⁴¹, Ser¹⁴², Leu¹⁷³, and Glu¹⁷⁴ in molecule 2. Twelve of these residues had alternative side chain positions in the holoenzyme structure. The mean temperature factors for components of the two complexes in the asymmetric unit were: protein, 19.22 and 17.22 Å²; NADPH, 14.89 and 12.88 Å²; and GW345, 14.63 and 13.97 Å², respectively. Temperature factor variations between the two complexes in the asymmetric unit were similar to those observed in the holoenzyme structure. The protein structures of the ternary and holo complexes were essentially identical (α-carbon root mean square deviation 0.31 Å).

Ligand Binding—Inhibitor GW345 is a member of a series of high affinity inhibitors of C. albicans DHFR (46). The series of compounds inhibit human DHFR more strongly than the fungal enzyme, with a Ki ratio in the range of 10–100. The Ki values of GW345 are 0.22 nM for C. albicans DHFR and 4.5 pM for human DHFR (46).

The hydrogen bond interactions between GW345 and the fungal enzyme were similar to those observed for other diaminopyrimidine-containing inhibitors in various DHFR complexes (2). The pyrimidine ring of GW345 (pKᵦ 8.2) is presumably protonated in the fungal DHFR complex, in analogy to trimethoprim (pKᵦ 7.1) and methotrexate (pKᵦ 5.2) complexes with bacterial DHFR (47, 48). As shown in Fig. 8, four hydrogen bonds linked GW345 to the C. albicans DHFR active site. The protonated ring nitrogen and the 2-amino group donated hydrogen bonds to the carboxylate group of Glu³². The 2-amino...
group also donates a hydrogen bond to a buried water molecule (not shown). A corresponding water molecule has been observed in all other refined structures of DHFR. The 4-amino moiety served as the donor in hydrogen bonds to the backbone carbonyl groups of Ile9 and Ile112.

The 3-pentyl substituent of GW345 was designed to take advantage of the hydrophobic region of the protein near Ile62. The pentyl moiety adopted a conformation that oriented its chain of carbons perpendicular to the pyrroloquinazoline ring system and positioned its methyl groups in regions of the active site that were predicted by the computer program GRID to interact favorably with methyl groups (46). The addition of the pentyl group to the pyrroloquinazoline ring system does not impart any significant selectivity between the fungal and human enzymes (46), although the geometry of the hydrophobic region of the active site in which the pentyl group resides is different in the two proteins, as described above. The conformational flexibility of the pentyl chain may permit optimal interactions with both enzyme binding sites.

The binding of NADPH was similar in the holo and ternary complexes and showed typical DHFR-cofactor interactions. The cofactor was draped over the edge of the central β-sheet with its nicotinamide moiety positioned in the active site via two hydrogen bonds between its amide group and the backbone of Ala11 (see Fig. 8). In the only atypical NADPH-protein interaction, Glu120 of molecule 2 appeared to form two hydrogen bonds with the adenine moiety; presumably the carboxyl group of Glu120 was neutral, donating a hydrogen bond to N7 and accepting a hydrogen bond from the amino group of the adenine portion of the cofactor. The remaining cofactor-protein interactions were analogous to those described in detail for previously reported structures of DHFR holo and ternary complexes (8, 9).

**Fig. 5.** A, stereoview of a superposition of *C. albicans* DHFR (red) and human DHFR (blue). NADPH and GW345 of the *C. albicans* DHFR structure are shown as ball-and-stick models. B, closeup stereoview of the active site region of the superposition of *C. albicans* DHFR (red) and human DHFR (blue) shown in A. Side chains of selected amino acid residues of both enzymes are shown in ball-and-stick representation, and side chains of *C. albicans* DHFR are labeled. This figure was prepared using MOLSCRIPT (51).
A volume of negative electron density was observed adjacent to Arg72 in the initial difference maps of the ternary complex and was tentatively modeled as a phosphate ion (see Fig. 6), although phosphate was not present in the crystallization solution. Based on subsequently determined structures of ternary complexes containing other inhibitors in which a molecule of MES was observed bound to Arg72, we suggest that a disordered molecule of MES is similarly bound in the GW345 complex.

Summary—To summarize, we have cloned, expressed, purified, crystallized, and determined the three-dimensional structure of *C. albicans* DHFR. The structure of the holoenzyme was solved at 1.85 Å by molecular replacement using a model derived from mouse DHFR. The overall fold of *C. albicans* DHFR is similar to that of other DHFRs, having an extended β-sheet with flanking α-helices. A ternary complex of *C. albicans* DHFR with NADPH and the nonselective inhibitor GW345 (Ki = 0.22 nM) diffracted to 1.60 Å. The structure of the protein and the binding of NADPH in the ternary complex were essentially identical to that found for the holoenzyme. GW345 was bound to *C. albicans* DHFR through four hydrogen bonds, as expected, and the pendant isopentyl group was oriented perpendicular to the pyrroloquinazoline ring system and positioned in a hydrophobic region of the protein active site. The binding cleft of *C. albicans* DHFR was 1.3–2.0 Å larger than that of human DHFR. This difference in the cleft geometry might provide a basis for selective inhibition. Indeed, a series of DHFR inhibitors that do show significant selectivity for *C. albicans* DHFR versus the human enzyme was recently reported (49). Some of those inhibitors, represented by Structure 1, were found to bind to the fungal enzyme in a manner that forces the arylthio substituent against the side of the binding cleft (near Ile62) where the human and fungal DHFR protein structures differ in backbone geometry. The larger cleft of the fungal enzyme structure appears to accommodate the sterically demanding shape of this class of inhibitor better than the smaller binding site of human DHFR. X-ray crystal structures of *C. albicans* DHFR in complex with several members of that series of selective inhibitors were consistent with the observed selectivity and will be reported in due course. Finally, the coordinates of the two
structures reported here have been deposited with the Protein Data Bank under the entries 1AI9 and 1AOE, for the holoenzyme and the inhibited enzyme structure, respectively.

Acknowledgments—We thank Jan Kopf and Cynthia A. Richards for their excellent technical contributions to the molecular biology and Drs. Reetta Raag and Joseph Chan for helpful discussions.

REFERENCES
1. Roth, B. (1980) Drugs Future 14, 249–268
2. Kuyper, L. F. (1989) in Computer-aided Drug Design (Perun, T. J., and Propst, C. L., eds) pp. 327–369, Marcel Dekker, Inc., New York
3. Schweitzer, B. I., Dicker, A. P., and Bertino, J. R. (1990) FASEB J. 4, 2441–2452
4. Bertino, J. R. (1993) J. Clin. Oncol. 11(1), 5–14
5. Frenkel, J. K., and Hitchings, G. H. (1957) Antibiots. Chemother. (Wash, D.C.) 7, 630–638
6. Brooks, R. G., Remington, J. S., and Luft, B. J. (1987) Antimicrob. Agents Chemother. 2, 297–306
7. Kass, E. H., Finland, M., and Platt, R. (1982) Rev. Infect. Dis. 4, 196–618
8. Kraut, J., and Matthews, D. A. (1987) in Biological Macromolecules and Assemblies (Jurnak, F., and McPherson, A., eds) Vol. III, pp. 1–71, John Wiley & Sons, Inc., New York
9. Freisheim, J. H., and Matthews, D. A. (1984) in Folate Antagonists as Therapeutic Agents (Sirotnak, F. M., Burchall, J. J., Ensminger, W. B., and Montgomery, J. A., eds) Vol. 1, pp. 69–131, Academic Press, New York
10. Kuyper, L. F., Roth, B., Baccanari, D. P., Ferone, R., Beddell, C. R., Champaess, J. N., Stammers, D. K., Dann, J. G., Norrington, F. E. A., Baker, D. J., and Goddard, P. J. (1985) J. Med. Chem. 28, 353–311
11. Kuyper, L. F. (1990) in Protein Design and the Development of New Therapeutics and Vaccines (Hook, J. B., and Poste, G., eds) pp. 297–327, Plenum Press, New York
12. Hitchings, G. H., Kuyper, L. F., and Baccanari, D. P. (1989) in Design of Enzyme Inhibitors as Drugs (Sandler, M., and Smith, H. J., eds) pp. 345–362, Oxford University Press, New York
13. Cheung, H. T. A., Searle, M. S., Feeney, J., Birdsall, B., Roberts, G. C. K., and Attardi, G. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 249–268
14. Kraut, J., and Matthews, D. A. (1987) in Computer-aided Drug Design (Perun, T. J., and Propst, C. L., eds) pp. 327–369, Marcel Dekker, New York
15. Hubscher, U., and Kornberg, A. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 13650–13654
16. Kuyper, L. F. (1989) in Design of Enzyme Inhibitors as Drugs (Sandler, M., and Smith, H. J., eds) pp. 345–362, Oxford University Press, New York
17. Walsh, T. J., Jarosinski, P. F., and Fromling, R. A. (1990) Diagn. Microbiol. Infect. Dis. 13, 37–40
18. Edwards, J. R. (1991) N. Engl. J. Med. 324, 1060–1062
19. Mills, J., and Masur, H. (1990) Sci. Am. 263, 50–57
20. Leung, G., and Mills, J. (eds) (1989) Opportunistic Infections in Patients with the Acquired Immunodeficiency Syndrome, Marcel Dekker, New York
21. Baccanari, D. P., Tananok, R. L., Joyner, S. S., Fling, M. E., Smith, P. L., and Freisheim, J. H. (1988) J. Biol. Chem. 264, 1100–1107
22. Godson, G. N., and Sinsheimer, R. L. (1967) Biochim. Biophys. Acta 149, 476–488
23. Wldawer, A., and Hodgson, K. O. (1975) Proc. Natl. Acad. Sci. U. S. A. 72, 398–399
24. Gilliland, G. L., and Davies, D. R. (1984) Methods Enzymol. 104, 370–381
25. Stura, E. A., and Wilson, I. A. (1990) Methods (Dalhousie) 1, 38–49
26. Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., and Salemm, F. R. (1987) J. Appl. Crystallogr. 20, 383–387
27. Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C., and Kraut, J. (1982) J. Biol. Chem. 257, 13650–13662
28. Stammers, D. K., Champaess, J. N., Beddell, C. R., Dann, J. G., Eliepoulos, E., Geddes, A. J., Ogg, D., and North, A. C. T. (1987) FEBS Lett. 218, 178–184
29. Fitzgerald, P. M. (1988) J. Appl. Crystallogr. 21, 275–278
30. Crowther, P. M. (1972) in The Molecular Replacement Method (Rossman, M. G., ed) pp. 173–178, Gordon and Breach, New York
31. Masters, J. N., and Astari, G. (1983) Gene (Amst.) 21, 59–63
32. Hardman, K. D., Agarwal, R. C., and Freisheim, J. H. (1982) J. Mol. Biol. 157, 69–86
33. Hendrickson, W. A. (1985) Methods Enzymol. 115, 252–270
34. Finzel, B. C. (1987) J. Appl. Crystallogr. 20, 53–55
35. Agarwal, R. C. (1978) Acta Crystallogr. Sec. A. 34, 791–809
36. Agarwal, R. C. (1981) in Refinement of Protein Structures, Proceedings of the Daresbury Conference, November 15 and 16, 1980, pp. 24–28 The Science
and Engineering Research Council, Daresbury, Warrington, UK

37. Daly, S., Mastromei, G., Yacoub, A., and Lorenzetti, R. (1994) Gene (Amst.) 147, 115–118

38. Matthews, B. W. (1968) J. Mol. Biol. 33, 491–497

39. Ramachandran, G. N., Venkatchaliam, C. M., and Krimm, S. (1966) Biophys. J. 6, 849–872

40. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. A 47, 110–119

41. Matthews, D. A., Bolin, J. T., Burridge, J. M., Filman, D. J., Volz, K. W., Kaufman, B. T., Bedell, C. R., Champness, J. N., Stammers, D. K., and Kraut, J. (1985) J. Biol. Chem. 260, 381–391

42. Oefner, C., D’arcy, A., and Winkler, F. K. (1988) Eur. J. Biochem. 174, 377–385

43. Champness, J. N., Achari, A., Ballantine, S. P., Bryant, P. K., Delves, C. J., and Stammers, D. K. (1994) Structure (Lond.) 2, 915–924

44. Volz, K. W., Matthews, D. A., Alden, R. A., Freer, S. T., Hunsch, C., Kaufman, B. T., and Kraut, J. (1982) J. Biol. Chem. 257, 2528–2536

45. Matthews, D. A., Bolin, J. T., Burridge, J. M., Filman, D. J., Volz, K. W., and Kraut, J. (1985) J. Biol. Chem. 260, 392–399

46. Kuyper, L. F., Baccanari, D. P., Jones, M. L., Hunter, R. N., Tansik, R. L., Joyner, S. S., Boytos, C. M., Rudolph, S. K., Knick, V., Wilson, H. R., Caddell, J. M., Friedman, H. S., Comley, J. C. W., and Stables, J. N. (1996) J. Med. Chem. 39, 892–903

47. Cocito, L., Roth, B., Temple, C., Jr., Montgomery, J. A., London, R. E., and Blakley, R. L. (1983) Arch. Biochem. Biophys. 236, 567–577

48. Bevan, A. W., Roberts, G. C. K., Feeney, J., and Kuyper, L. F. (1985) Eur. Biophys. J. 11, 211–218

49. Chan, J. H., Hong, J. S., Kuyper, L. F., Baccanari, D. P., Joyner, S. S., Tansik, R. L., Boytos, C. M., and Rudolph, S. K. (1995) J. Med. Chem. 38, 3608–3616

50. McRee, D. (1992) J. Mol. Graphics 10, 44–46

51. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950

52. Delves, C. J., Ballantine, S. P., Tansik, R. L., Baccanari, D. P., and Stammers, D. K. (1993) Protein Expression Purif. 4, 16–23

53. Baccanari, D. P., Stone, D., and Kuyper, L. (1981) J. Biol. Chem. 256, 1738–1747

54. Stone, D., Patterson, S. J., Raper, J. H., and Phillips, A. W. (1979) J. Biol. Chem. 254, 480–488