

การทดสอบ

การทดสอบวัดค่า Inhibition Constant (K_i) เพื่อตรวจสอบความสามารถในการทำนายอนุพันธ์ที่ออกแบบได้ด้วยวิธี Structure-based drug design ผลการศึกษาพบว่า มีเพียงอนุพันธ์เดียวซึ่งมีหมู่แทนที่ Phenylpropyl (อนุพันธ์หมายเลข 4) ที่สามารถสังเคราะห์ได้ ซึ่งผลการทดสอบ K_i กับ Wildtype PfDHFR พบว่าค่า K_i ของอนุพันธ์ 4 มีค่าใกล้เคียงกับ K_i ของ PYR

อย่างไรก็ตาม ได้นำอนุพันธ์อื่นๆ แต่มีโครงสร้างใกล้เคียงและทดสอบค่า K_i แล้วมาเปรียบเทียบกับอนุพันธ์ที่ทำนายได้ ซึ่งสรุปไว้ในตารางที่ 3

ตารางที่ 3 ตรวจสอบความสามารถในการทำนายอนุพันธ์ที่ออกแบบกับค่า K_i ของอนุพันธ์ที่ได้จากการทดลอง

Compd	Fragment		K_i (nM)	Rel to PYR
	Prediction	Synthesis		
PYR	-	-	0.6 ± 0.2	1.0
1.	$-(CH_2)_2COMe$	$-(CH_2)_3COOMe$	0.3 ± 0.2	0.5
2.	$-(CH_2)_2Ph$	$-(CH_2)_3Ph$	0.7 ± 0.1	1.2
3.	$-CH_2CH(CH_3)_2$	-	-	-
4.	$-(CH_2)_3Ph$	$-(CH_2)_3Ph$	0.7 ± 0.1	1.2
5	$m-CH_2CH(CH_3)_2$	-	-	-
6	$p-CH_2NHMe$	$p-OMe$	0.9 ± 0.1	1.5
7	$p-CH_2CH_3$	$p-Me$	0.4 ± 0.0	0.7

บทวิจารณ์

โครงสร้างสามมิติของโมเลกุลของเอนไซม์ไดไฮโดรโฟเลตรีดักเทส จาก *Plasmodium falciparum* ที่สร้างโดยเทคนิค Comparative Structural Modeling ทำให้เข้าใจกลไกบางอย่างในระดับโมเลกุลได้ดียิ่งขึ้น โดยเฉพาะในเรื่องของอันตรกิริยาระหว่างเอนไซม์กับยารักษาโรคมาลาเรียในกลุ่มสารยับยั้งโฟเลตในระดับโมเลกุล สามารถคาดคะเนอิทธิพลของ Mutation ที่จะมีผลกระทบต่อการ

จับยึดของโมเลกุลของยา ได้อย่างจำเพาะเจาะจง และได้ทดลองออกแบบของอนุพันธ์ของ CYC และ PYR จากโครงสร้างของเอนไซม์

ในการสร้าง Homology Model ของ PfDHFR ข้อมูลจาก Sequence Alignment ทำให้ทราบว่า การที่ Primary Sequence ของ PfDHFR มี 2 ช่วงที่ไม่สามารถเทียบเคียงกับ Primary Sequence ของ DHFR อื่นๆได้ เป็นไปได้ว่า ส่วนพิเศษที่มีลักษณะเป็น Insertion ของ PfDHFR นั้นอาจเกี่ยวข้องกับหน้าที่ที่เป็น Bifunction ของเอนไซม์ โดยช่วยในการจับยึดกับ Thymidylate Synthase ซึ่งเป็น Co-partner ของเอนไซม์ ทั้งนี้เมื่อสังเกต DHFR อื่นๆ ที่นำมาเปรียบเทียบจะเป็น Monofunctional enzyme ด้วยกันทั้งสิ้น อย่างไรก็ตาม ประเด็นดังกล่าวยังเป็นเพียงข้อสันนิษฐาน ปัจจุบันยังไม่มีหลักฐานที่ชัดเจนที่จะอธิบายหน้าที่ของ Insertion ทั้งสองได้

สำหรับในส่วนที่เป็น Conserved Sequence นั้นพบว่า Tertiary Structure ของ PfDHFR จะคล้ายกับ DHFR ที่นำมาเปรียบเทียบ Homology Model สามารถระบุการจับยึดที่สำคัญระหว่างกรดอะมิโนในโปรตีนและสารยับยั้ง โดยเฉพาะลักษณะอันตรกิริยาชนิด Hydrogen Bonding ของ Sidechain Asp54 กับ HN ของวง Pyrimidine และ Triazine ของ PYR และ CYC ตามลำดับ

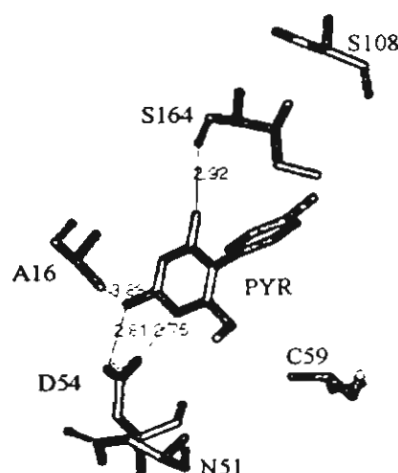
แม้ว่าการจับยึดของ CYC ที่ทำนายจาก Homology Model จะระบุว่ามีความคล้ายคลึงกับ PYR แต่พบว่า 2,2'-dimethyl substituent จะอยู่ใกล้ Sidechain ของ Ala16 มากกว่า Ethyl group ของ PYR ซึ่ง Mutation ชนิด A16V จะมี Sidechain เปลี่ยนจาก Alanine เป็น Valine มีขนาดใหญ่ขึ้น ส่งผลให้เกิด Steric Constraint เมื่อ CYC จับกับ Mutant ชนิด A16V แต่ผลกระทบบดังกล่าวจะไม่เกิดขึ้นกับกรณี PYR จับกับ A16V Mutant ข้อสรุปดังกล่าวได้รับการยืนยันจากการศึกษาด้วยวิธี Molecular Dynamics Calculation ของระบบที่ประกอบด้วย Wildtype, A16V และ A16V+S108T PfDHFR ที่เกิดคอมเพล็กซ์กับ CYC และ PYR และการทดสอบค่า K_i จากอนุพันธ์ของ Cycloguanil ที่สังเคราะห์ให้มี Methyl group แทนที่เพียงหมู่เดียว (Rastelli 2000) นอกจากนี้ ผลการศึกษานี้สอดคล้องได้เป็นอย่างดีกับผลการทดลองที่เคยพบว่า A16V PfDHFR มีระดับการดื้อ Cycloguanil มากถึงประมาณ 217 เท่าเมื่อเทียบกับ Wildtype PfDHFR เท่านั้น ในขณะที่ค่า K_i ของ Pyrimethamine กับ A16V สูงกว่า Wildtype เพียง 4 เท่า ซึ่งอยู่เกณฑ์ที่ถือว่าสามารถยับยั้งเอนไซม์ได้ (Sirawaraporn 1997)

ตำแหน่ง Mutation ที่สร้างปัญหาการดื้อยาอีกตำแหน่งหนึ่งคือ Ser108 Mutant ในกลุ่มนี้จะมี S108N เป็นหลัก เช่น S108N N51I+S108N C58R+S108N N51I+C59R+S108N N51I+C59R+S108N+I164L ระดับการดื้อยาของ Mutant ในกลุ่มนี้จะสูงทั้ง CYC และ PYR โดยเฉพาะ Quadruple Mutation ซึ่งค่า K_i ของยาทั้งสองจะสูงกว่า Wildtype หลายร้อยเท่า (Sirawaraporn 1997) อิทธิพลของ Mutation จาก Ser108 เป็น Asn108 น่าจะเกี่ยวข้องกับ Steric Conflict เป็นส่วนสำคัญ กล่าวคือ Sidechain ของ Asn108 มีขนาดใหญ่กว่า Ser108 ทำให้บริเวณ

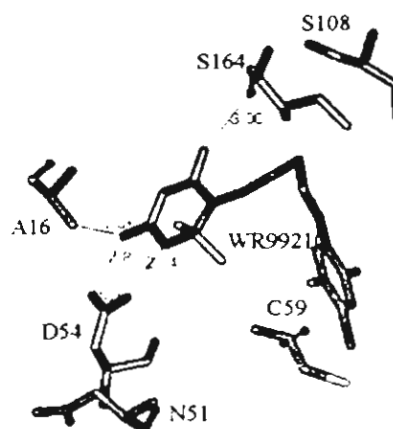
Binding Pocket แคบลงและมีรูปร่างเปลี่ยนไป ซึ่งทั้ง CYC และ PYR เป็นโมเลกุล Biphenyl ring ที่ค่อนข้าง Rigid ดังนั้นการปรับโมเลกุลเพื่อให้จับยึดกับ Mutant ได้ดีเหมือนกับ Wildtype จึงเป็นเรื่องยาก จากรายงานการทดสอบค่า K_i ของอนุพันธ์ของสารยับยั้งโฟเลตที่มีความยืดหยุ่นกว่า เช่น กลุ่ม Trimethoprim และ WR99210 พบว่าสามารถเพิ่มระดับความสามารถในการยับยั้ง Mutant ในกลุ่ม S108N ได้ดีกว่า CYC และ PYR (Rastelli 2000; Tamchompoo 2002; Yuvaniyama 2003)

การออกแบบโมเลกุลของยาด้วยเทคนิค Structure-Based Drug Design สามารถแสดงผลได้อย่างรวดเร็วและการทำนายให้ผลเป็นที่น่าพอใจ แม้ว่า โครงสร้างของโมเลกุลเป้าหมายซึ่งเป็นข้อมูลหลักในการออกแบบจะเป็น Homology Model ซึ่งปัจจุบันเราทราบโครงสร้างสามมิติของเอนไซม์ที่เป็น Wildtype PfDHFR-WR99210, C58R+S108N-PYR complex และ N51I+C59R+S108N+I164L-WR99210 complex (Yuvaniyama 2003) แล้ว (ในขณะที่ศึกษาอยู่นั้น ยังไม่สามารถหาโครงสร้างเอ็กซ์เรย์ของเอนไซม์ได้) จากเปรียบเทียบบริเวณ Binding Site พบว่าใกล้เคียงกัน (รูปที่ 8) ทำให้อนุพันธ์ที่ออกแบบมามีนัยสำคัญในระดับหนึ่ง

A. Homology model



B X-ray structure



รูปที่ 8 โครงสร้างที่ได้จาก Homology Technique (A) กับ X-ray Crystallography (B) เปรียบเทียบตำแหน่งของกรดอะมิโน Ala16, Asn51, Asp54, Cys59 , Ser108 และ Ile164 กับตำแหน่งของสารยับยั้ง Pyrimethamine (PYR) และ WR99210

แม้ว่าการออกแบบอนุพันธ์ของยาได้มาจากโครงสร้างของ Wildtype PfDHFR เท่านั้น ยังไม่ได้ครอบคลุมในส่วนที่เป็น Mutants ซึ่งการออกแบบโดยใช้โครงสร้าง X-ray ของ Mutant ที่มีระดับการดื้อยาสูงๆ เช่น Quadruple Mutation เป็นสิ่งที่น่าสนใจ นอกจากนี้ ความรุนแรงของการดื้อยาใน Mutant ชนิดนี้ ทำให้มีแนวโน้มว่าการออกแบบอนุพันธ์ในกลุ่มของ Flexible Antifolates เช่น

Trimethoprim และ WR99210 จะมีความน่าสนใจมากกว่าในกลุ่ม Rigid Antifolates เช่น CYC และ PYR อย่างไรก็ดี ยังมีอุปสรรคที่สำคัญของเทคนิคนี้ คือ Fragment ที่มีอยู่ในฐานข้อมูลมีจำนวนน้อย ทำให้การสืบค้นขาดความหลากหลาย และอนุพันธ์ที่ออกแบบมาส่วนใหญ่ไม่สามารถสังเคราะห์ได้ การประเมินความน่าเชื่อถือของเทคนิคจึงไม่สามารถกระทำได้ ทำให้ผลงานวิจัยขาดความสมบูรณ์

อย่างไรก็ตาม การออกแบบด้วยเทคนิค Structure-Based Drug Design อาจใช้ประโยชน์ในแง่ของการเสนอแนวโน้มนั้น เช่น ปัญหาที่อนุพันธ์ที่ออกแบบมาไม่สามารถสังเคราะห์ได้ อาจแก้ไขโดยสังเคราะห์สารให้มีโครงสร้างทางเคมีใกล้เคียงกับอนุพันธ์นั้น ซึ่งจากการเปรียบเทียบในตารางที่ 3 ค่า K_i ของอนุพันธ์ที่มีหมู่แทนที่ใกล้เคียงกับที่ออกแบบไว้อยู่ในเกณฑ์ที่ดี

สรุป

งานวิจัยนี้ใช้เทคนิคการออกแบบโมเลกุลของยารักษาโรคมาลาเรียโดยอาศัยข้อมูลโครงสร้างสามมิติของโมเลกุลเป้าหมาย โดยเลือก เอนไซม์ไดไฮโดรโฟเลตรีดักเทส จาก *Plasmodium falciparum* อาศัยเทคนิค Comparative Structural Modeling เพื่อสร้างโครงสร้างของเอนไซม์ แล้วจึงออกแบบโมเลกุลโดยวิธีการสืบค้น Fragment ที่มีอันตรกิริยาที่เหมาะสม โครงสร้างของเอนไซม์ทำให้เข้าใจอย่างลึกซึ้งถึงลักษณะการจับยึดที่สำคัญของยากับเอนไซม์ สามารถอธิบายสาเหตุของการดื้อยาของเอนไซม์กลายพันธุ์บางชนิดได้ และผลจากการสืบค้นและคัดกรองทำให้ได้อนุพันธ์ของ Pyrimethamine จำนวน 7 ชนิด และในจำนวนนี้ 1 ชนิดได้ทำการทดสอบซึ่งให้ผลใกล้เคียงกับ Pyrimethamine ที่ทดสอบกับ PfDHFR ชนิดไวต์ไทป์



Interaction of Pyrimethamine, Cycloguanil, WR99210 and their Analogues with *Plasmodium falciparum* Dihydrofolate Reductase: Structural Basis of Antifolate Resistance

Giulio Rastelli,^a Worachart Sirawaraporn,^{b,*} Pornthep Sompornpisut,^c
Tirayut Vilaivan,^c Sumalee Kamchonwongpaisan,^d Rachel Quarrell,^c Gordon Lowe,^c
Yodhathai Thebtaranonth^f and Yongyuth Yuthavong^d

^aDipartimento di Scienze Farmaceutiche, Università di Modena e Reggio Emilia, Via Campi, 183-41100 Modena, Italy

^bDepartment of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

^cDepartment of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

^dNational Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Bangkok 10400, Thailand

^eDyson Perrins Laboratory, South Parks Road, Oxford University, Oxford OX1 3QY, UK

^fDepartment of Chemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

Received 29 October 1999; accepted 15 January 2000

Abstract—The nature of the interactions between *Plasmodium falciparum* dihydrofolate reductase (pDHFR) and antimalarial antifolates, i.e., pyrimethamine (Pyr), cycloguanil (Cyc) and WR99210 including some of their analogues, was investigated by molecular modeling in conjunction with the determination of the inhibition constants (K_i). A three-dimensional structural model of pDHFR was constructed using multiple sequence alignment and homology modeling procedures, followed by extensive molecular dynamics calculations. Mutations at amino acid residues 16 and 108 known to be associated with antifolate resistance were introduced into the structure, and the interactions of the inhibitors with the enzymes were assessed by docking and molecular dynamics for both wild-type and mutant DHFRs. The K_i values of a number of analogues tested support the validity of the model. A 'steric constraint' hypothesis is proposed to explain the structural basis of the antifolate resistance. © 2000 Elsevier Science Ltd. All rights reserved.

Introduction

The dihydrofolate reductase (DHFR) domain of *Plasmodium falciparum* bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-TS) is one of the few well-defined targets in malarial chemotherapy. The enzyme catalyzes the nicotinamide adenine dinucleotide phosphate (NADPH) dependent reduction of dihydrofolate to tetrahydrofolate. DHFR has received considerable attention as it is the target of pyrimethamine

(Pyr) and cycloguanil (Cyc) and other antifolates used for prophylaxis and treatment of *P. falciparum* infection. The rapid emergence of antifolate resistant *P. falciparum* has unfortunately compromised the clinical utilities of the drugs, and thus highlights the urgent need to search for new effective antifolate antimalarials.

Analysis of the gene encoding DHFR from *P. falciparum* parasites obtained from geographically distant sources and with different degrees of resistance suggested the association of antifolate resistance with point mutations of the gene coding for *P. falciparum* DHFR.^{1–9} Mutations of one or more residues at amino acid positions 16, 51, 59, 108, and 164 of the *P. falciparum* DHFR were identified to be involved in antifolate resistance. Evidence available recently also supports the hypothesis that antifolate resistance in malaria evolved in a step-wise selection of mutants, with multiple mutations subsequent to a single S108N mutation, and that

Abbreviations: Cyc, cycloguanil; Pyr, pyrimethamine; pDHFR, *Plasmodium falciparum* dihydrofolate reductase; RMS, root mean square; TES, *N*-tris (hydroxymethyl)-methyl-2-aminoethane-sulfonic acid; H₂folate, 7,8-dihydrofolate; fs, femto-second; RESP, restrained electrostatic potential.

*Corresponding author. Tel.: +662-245-5195; fax: +662-248-0375; e-mail: scwsr@mahidol.ac.th

mutants which do not survive in nature were either poorly resistant to the drugs or had insufficient catalytic activity of DHFR to support DNA synthesis.¹⁰

While the parasites with A16V+S108T double mutation in the *dhfr* gene are resistant to Cyc but susceptible to Pyr, those with cross resistance to both drugs have either single or multiple mutations at amino acid residues 51, 59, 108 and 164.^{5–7,11} Since the structures of Pyr and Cyc are very closely similar, the questions arise as to why Cyc-resistant but Pyr-sensitive evolved while the reverse has never been reported in the natural isolates, and whether the binding of Pyr and Cyc to the wild-type and the mutant DHFRs are different. Similar questions can also be raised with WR99210, a dihydro-triazine whose structure is closely related to Cyc. The fact that WR99210 is highly effective against both Pyr-sensitive and -resistant parasites in vivo¹² and the DHFRs in vitro¹³ has made WR99210 an attractive compound for further studies.

In order to address the problems of binding modes of Pyr, Cyc, and WR99210, structural data of *P. falciparum* DHFR are needed. In the absence of pDHFR structure from X-ray diffraction analysis, it is necessary to construct and predict the structure of the enzyme by a homology modeling approach.¹⁴ Such modeling approach has been exploited for the prediction of the structure of the new drug leads capable of inhibiting *P. falciparum* DHFR at micromolar and submicromolar levels.¹⁵ Using the model built from the 3D-structure of *Leishmania major* DHFR-TS, McKie et al.¹⁶ predicted that the S108N mutation of the *P. falciparum* DHFR would lead to a steric clash with the *p*-Cl atom of Pyr. The hypothesis was readily tested by modifying the Pyr structure to avoid the steric constraint, hence resulting in two Pyr analogues which are highly effective against Pyr-resistant DHFR and malarial parasites.¹⁶ Indeed, through analysis of a number of physicochemical characteristics of residues at position 108 of *P. falciparum* DHFR and compared with their activities against a series of drugs including Pyr, Cyc, TMP (trimethoprim), and WR99210, Warhurst has recently pointed out that mutant DHFRs with bulky groups at residue 108 would cause steric constraint and hence result in poor binding with the inhibitor.¹⁷

In this paper, we describe the construction of a homology model of DHFR domain of the bifunctional *P. falciparum* DHFR-TS, and the structures of the complexes between Pyr, Cyc and WR99210 with the wild-type, A16V, S108T, and A16V+S108T mutant enzymes. We predict from the model that A16V mutation of *P. falciparum* DHFR imposes a steric conflict for binding to Cyc as opposed to Pyr and WR99210, and that the steric conflict becomes more pronounced when combined with mutation at residue 108 in the A16V+S108T double mutants. The hypothesis was tested against experimental data on the binding strengths of analogues of these drugs to both wild-type and mutant enzymes. This model was used to design and synthesize new compounds which were tested against wild-type and resistant *P. falciparum* in vitro.

Materials and Methods

Homology modeling of *Plasmodium falciparum* DHFR

The structure of the DHFR domain of *P. falciparum* bifunctional DHFR-TS was obtained from homology modeling simulations. The model was constructed by alignment of *P. falciparum* DHFR primary sequence¹⁸ with the DHFR sequences of which the crystal structures are already known; these include DHFRs from *Escherichia coli*,¹⁹ *Lactobacillus casei*,²⁰ human,^{21,22} chicken liver²³ and *Pneumocystis carinii*.²⁴ Unfortunately, the *Leishmania major* bifunctional DHFR-TS was not included into the sequence alignment as the crystal structure was not publicly available at the time the model was built. The alignment of *P. falciparum* DHFR sequence¹⁸ was carefully checked by using the available sequence alignment evaluation functions within MODELLER 3.0,²⁵ and by superimposing the structurally conserved regions of the available DHFR crystal structures.

The model of pDHFR was constructed using MODELLER 3.0.²⁵ A primary homology model was constructed based on the conserved regions of most DHFRs by removing the sequences of the two insertions, i.e., residues 22–38 and 64–98. Next, several models were constructed on the whole sequence by adding one at a time the local alignments found to fill the two insertions. Local alignments were predicted using BLAST (Basic Local Alignment Search Tool)²⁶ at the National Center for Biotechnology Information, NCBI (web site hyperlink <http://www.ncbi.nlm.nih.gov/blast>). A stringent significant threshold of 3 was adopted for the BLAST search. Then, combinations of these, and finally all the local alignments were introduced. The model showing the best score as judged by the value of the Modeller objective function, and with the least root mean square (RMS) deviation with respect to the conserved regions of the DHFR templates, was saved for further refinement with molecular dynamics. The quality of the stereochemistry and of the structures was validated using the PROCHECK²⁷ and WHATIF tools available at the Biotech Validation Suite for Protein Structures (web site <http://biotech.embl-heidelberg.de:8400/>).

Molecular dynamics refinement of the structure

The pDHFR structure resulting from MODELLER was refined by means of molecular mechanics and molecular dynamics calculations using AMBER 4.1²⁸ and the all-atom force field by Cornell et al.²⁹ Bond lengths involving hydrogens were constrained using the SHAKE algorithm³⁰ and a time step of 2.0 femtosecond (fs). Molecular dynamics were performed at 300 K using a residue-based cut-off of 10 Å. Pairlists were updated every 25 time steps. Hydrogens were added to the structure predicted by MODELLER using the stored internal coordinates of the AMBER all-atom data base and then minimized, the heavy atoms being kept fixed in their original positions. All Lys and Arg residues were positively charged and Glu and Asp residues were negatively charged.

The reduced form of the cofactor (NADPH) was used for the calculations, according to the tighter binding of NADPH over NADP⁺ to DHFR.^{31–33} Force field parameters for NADPH were adapted from previous simulations that employed the oxidized form (NADP⁺) of the cofactor.³⁴ The geometry of the reduced nicotinamide ring was fully optimized at the 6-31G* level using Gaussian 94, and partial charges on atoms were calculated from an electrostatic potential fit to a 6-31G* ab-initio wave function, followed by a standard RESP (restrained electrostatic potential) fit.^{35,36} Bond, angle, dihedral and van der Waals parameters of the nicotinamide ring were assigned consistently with the Cornell et al. force field, and then tested for their ability to reproduce the 6-31G* optimized geometry. The remaining adenosine and phosphate parameters of NADPH were taken from NADP⁺.³⁴

Molecular mechanics and molecular dynamics calculations on the two long insertions in the pDHFR sequence were performed first to allow more relaxation of the non-conserved and less validated regions of the structure without affecting the geometry of the most conserved regions. The two loops were energy-minimized with 3000 steps of conjugate gradient minimization, followed by 10 ps molecular dynamics at 300 K. This protocol was repeated 10 times, with energy-minimization after each molecular dynamics. To perform these first molecular mechanics and molecular dynamics calculations, a distance-dependent dielectric constant with a dielectric multiplication constant of 4 was used to simulate the presence of water.

In order to include the effects of explicit water molecules at the active site of DHFR, the final structure obtained was solvated with a spherical cap of 2113 TIP3P³⁷ water molecules centered on the center of mass of pDHFR. Calculations were continued using a dielectric constant of 1 and the explicit water molecules. In order to allow the equilibration of the solvent around the protein residues, 100 ps molecular dynamics at 300 K were performed on the water molecules. Then, the whole system (enzyme, water, and NADPH) was energy-minimized with 3000 steps, with 5 Kcal/mol restraint applied on the backbone atoms of the enzyme with the exception that the backbone atoms of loops 1 and 2 were left free. The restraint was imposed in order to avoid initial large drifts from the homology model structure, which is based on experimentally-validated structures of DHFRs. Then, 2000 steps of minimization without any constraints were performed, and the resulting structure was compared with the homology model and the DHFR templates to ensure that the conserved regions maintained similar 3D-structure. A 200 ps molecular dynamics simulation at 300 K was then performed on the whole system with a 1 Kcal/mol restraint on the backbone atoms. Whilst performing molecular dynamics, coordinates were collected every 0.2 ps and averaged every 10 ps for visual inspection. Finally, 2000 steps of energy-minimization without constraints were performed. The values of RMS deviations between the final pDHFR model and the crystal structures of DHFRs from different sources used for homology modeling

were calculated. The considerably low RMS deviations, 1.40 with *E. coli*, 1.47 with *L. casei*, 1.21 with human, 1.24 with chicken liver, and 1.26 with *P. carinii*, revealed that the pDHFR model was close to these structures.

Docking of pyrimethamine, cycloguanil and WR99210, and molecular dynamics of the complexes

The inhibitors were docked into the active site of pDHFR, using the final structure obtained from the explicit solvent molecular dynamics calculations. The inhibitors were modeled as protonated state at N1, as evidenced by the NMR studies of both the bound and free states of Pyr, methotrexate (MTX) and trimethoprim (TMP).^{38,39} The protonated amino portions of the inhibitors were initially positioned to interact with the anionic side chain of D54. This choice was strongly suggested by the binding modes of other antifolates with the conserved aspartic or glutamic acid residue corresponding to D54 of the pDHFR in the available crystal structures of DHFRs from different sources. While Pyr and Cyc have conformationally constrained structures, the more conformationally flexible 2',4',5'-trichlorophenoxy propyloxy group of WR99210 was docked into the active site using the orientations of folate and MTX in the available crystal structures as guidelines.

To obtain a set of atomic charges for the inhibitors, their structures were completely optimized at the STO-3G basis set level, using Gaussian 94. Charges were calculated from an electrostatic potential fit to a 6-31G* ab-initio wave function, followed by a standard RESP fit. Atom types were assigned consistently with the Cornell et al. force field,²⁹ and some parameters were adjusted to reproduce the STO-3G optimized geometry. The parameters for the chlorine atoms of the inhibitors were as previously described.⁴⁰

The molecular mechanics and molecular dynamics calculations of the complexes with the inhibitors were performed in explicit water molecule environment. The same solvation sphere of the pDHFR–NADPH binary complex was used, but the water molecules that are replaced by the inhibitors were removed from the structure. Two thousand steps of minimization on the inhibitor, NADPH, the water molecules and all the protein residues within 10 Å from the inhibitor were performed, followed by 300 ps molecular dynamics at 300 K. A 1 Kcal/mol restraint on the backbone atoms was applied to avoid undesirable drifts from the initial structure.

Molecular dynamics of the pDHFR mutant complexes

The mutant pDHFRs complexed with Pyr, Cyc and WR99210 were constructed by changing alanine to valine (A16V), serine to threonine (S108T), and both (A16V + S108T) in the corresponding structures of the wild-type pDHFR–inhibitor complexes. Five hundred steps of energy-minimization were performed on the inhibitor and the mutated residue only in order to position the side chain of valine and/or threonine with respect to the inhibitor and to relieve steric conflicts due

to the growing of the bulkier amino acid side chains. This was followed by 1000 steps of minimization on the protein residues within 10 Å from the inhibitor, the inhibitor, and all the water molecules. The molecular dynamics at 300 K was then started from these minimized complexes and run for over 300 ps. The 1 Kcal/mol restraint on the backbone atoms was applied as usual.

Synthesis of derivatives of cycloguanil and WR99210

Compounds I, II, III, XIV and XV were prepared by a three-component condensation between 4-chloroaniline, dicyanodiamide and acetone in the presence of concentrated HCl (I) or two-component condensation between the arylbiguanides and the aldehydes (II, III, XIV and XV) both in the presence of concentrated HCl according to the literature.^{41–45} The new compound IV was synthesized by a three-component condensation between 4-chloroaniline, dicyanodiamide and methylal. Compounds V–XII were synthesized by a modified two-component condensation⁴⁴ between the corresponding biguanides and ketones in the presence of concentrated HCl as catalyst and triethyl orthoacetate as water scavenger, the details of which will be published elsewhere (Vilaivan et al., in preparation). All cycloguanil derivatives were isolated as monohydrochloride salts and were purified by crystallization from water or water–ethanol. New compounds gave clean ¹H nmr and mass spectra (APCI or MALDI) and provide satisfactory elemental analysis (C,H,N).

Preparation of pfDHFRs and inhibitor binding studies

The recombinant plasmids pET-pfDHFR (wild-type) and pET-pfDHFR (A16V+S108T) were prepared as previously described¹⁰ and were employed for the preparation of enzymes used for testing the compounds synthesized in the present study. Purification of both wild-type and mutant pfDHFRs was achieved by affinity chromatography on MTX-Sepharose CL-6B column.⁴⁶ The activity of pfDHFR was determined spectrophotometrically by monitoring the decrease in absorbance at 340 nm at 25°C according to the standard assay method described previously.⁴⁶ For determination of the *K_i* values of a number of synthesized analogues, the assay reactions were initiated with affinity-purified enzyme (~0.005–0.01 U ml⁻¹), and the initial velocities in the presence of varying concentrations of inhibitors were analyzed by non-linear least square fit using KaleidagraphTM software. Other analyses of the protein were as earlier described.¹⁰

Results

The alignment of the sequence of pfDHFR with the sequences of DHFRs from selected sources of which the X-ray structures are known is shown in Figure 1. There is significant homology among the DHFRs compared, with the exception that the pfDHFR sequence contained two extra insertions between residues 22–38 and residues 64–98. Our initial pfDHFR model was therefore

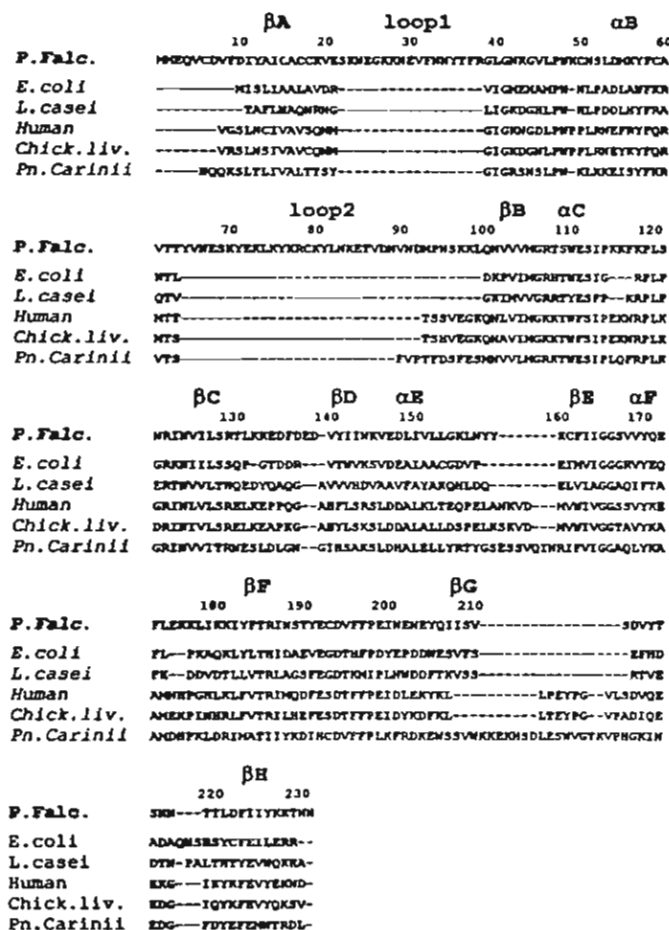


Figure 1. Alignment of the amino acid sequences of dihydrofolate reductase of *Plasmodium falciparum*,¹⁸ *Escherichia coli*,⁴⁹ *Lactobacillus casei*,⁵⁰ human,^{51,52} chicken liver⁵³ and *Pneumocystis carinii*.⁵⁴ The percentages of the identical amino acids between *P. falciparum* DHFRs and DHFRs from *E. coli*, *L. casei*, human, chicken liver, and *P. carinii* are 29, 24, 32, 32, and 24%, respectively. The elements of secondary structure are indicated as α , β , and loop.

constructed without the sequences of the two insertions. The resulting model was then used to check for the folding of the conserved regions, to make superimpositions and comparisons with other known templates.

The local alignments of the two extra sequences of pfDHFR were predicted from BLAST.²⁶ A stringent significance threshold of 3 was adopted, resulting in two significant homologous segments for loop 1 and five for loop 2. The predicted alignments were restored into the DHFR model. As for loop 1, the local alignments predicted to fill this insertion significantly reduced the quality of the model and also markedly affected the structure of the most conserved regions. On the contrary, one out of the five local alignments that BLAST identified to fill the second insertion gave an acceptable initial arrangement of loop 2 without affecting the quality and the structure of the conserved regions. Therefore, the initial homology model was constructed using the best local alignment only for loop 2. The loop 1, however, was subsequently constructed manually and then appended to the structure previously obtained

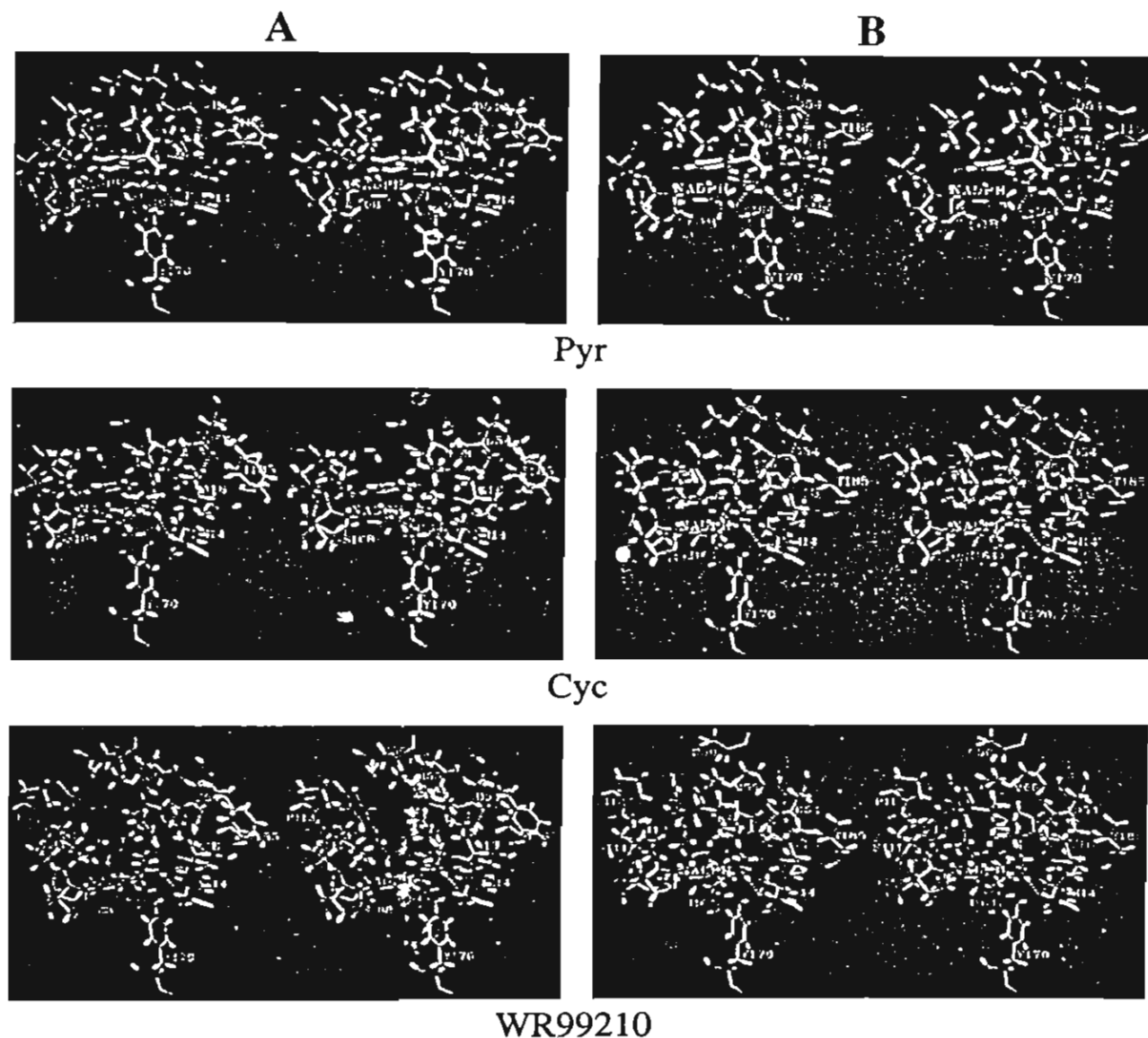


Figure 3. Stereoviews of wild-type (Panel A) and A16V + S108T (Panel B) pfDHFRs showing active site residues involved in the interaction with Pyr (yellow), Cyc (green), and WR99210 (red). The residues shown are I14, A16, C50, D54, M55, F58, S108, S111, I112, P113, I164, Y170 and NADPH molecule.

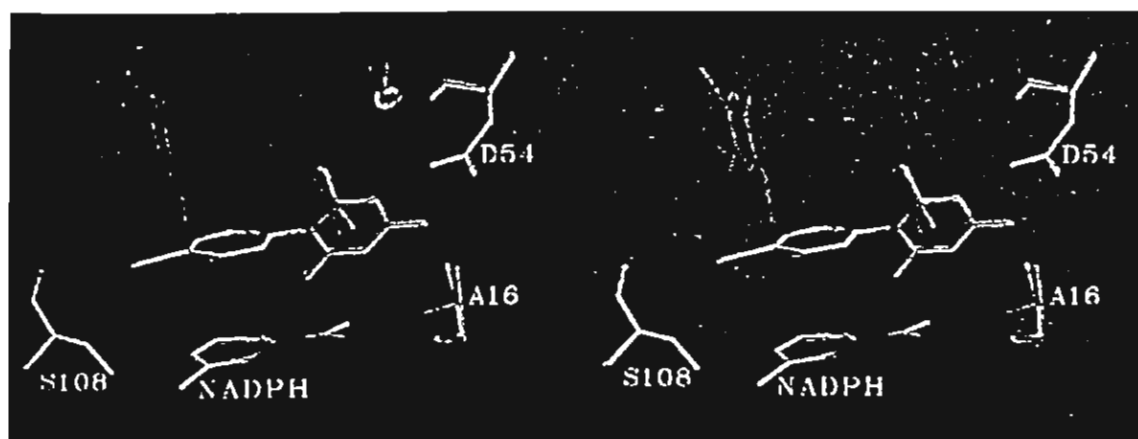


Figure 4. Stereoviews showing the superimpositions of inhibitors bound at the active site of wild-type pfDHFR. Residues A16, D54, S108, and the nicotinamide ring of NADPH are shown with the inhibitors Pyr (yellow), Cyc (green), and WR99210 (red).

with the same residues as those described for Cyc. However, while the chlorophenyl substituent of Cyc and Pyr stacks against the nicotinamide ring of NADPH and is directed toward S108, the flexible trichlorophenoxy propyloxy of WR99210 is almost perpendicular with the nicotinamide ring and is not in contact with the cofactor. Interestingly, one water molecule binds over the nicotinamide ring, making hydrogen bonds with the N1 nitrogen of the nicotinamide and the 2'-hydroxyl of the ribose. The trichlorophenoxy propyloxy substituent of WR99210 is in contact with C50, M55, S111, and P113 (see Fig. 3A and B).

The effects of single and double mutations at amino acid residues 16 and 108 on the binding of Pyr, Cyc and WR99210 were further investigated (Fig. 5). The molecular dynamics stimulation of the Pyr complexed with A16V mutant model revealed minimal changes of orientation and binding interactions with respect to that of the wild-type enzyme complex (Fig. 5A). By contrast, the complex between Cyc and the A16V mutant was significantly different from the complex with the wild-type enzyme; the inhibitor was displaced from its position in the wild-type complex due to the steric conflict with valine in A16V mutant (Fig. 5B). Superimposition between Cyc and Pyr in the wild-type complexes indicated that one of the two 2,2'-methyl substituents of Cyc was located much closer (1.4 Å) to the methyl of the A16 side chain with respect to the 6-ethyl substituent of Pyr, the former being in Van der Waals contact with the methyl of A16 whereas the latter is significantly more distant (data not shown). Therefore, the steric demand imposed by the A16V mutation should be much more pronounced for Cyc than for Pyr, although hydrogen bonding interactions with D54, I14, and I164 are still present. The displacement of Cyc in the structure of the A16V mutant is illustrated in Figure 5B. While the displacement of the triazine ring of Cyc did not seem to affect the ability to make hydrogen bond between the drug and D54, I14, and I164, the orientation of the chlorophenyl ring of the inhibitor bound to the A16V enzyme was found to be significantly different from that of the wild-type enzyme. This is due to Cyc being bound in a locked region comprising A16 and D54 for the dihydrotriazine ring, nicotinamide, and S108 for the chlorophenyl ring of the inhibitor. Interestingly, the stacking interactions with the nicotinamide ring of NADPH were completely lost in the A16V mutant and one water molecule was found in place of the chlorophenyl ring (data not shown).

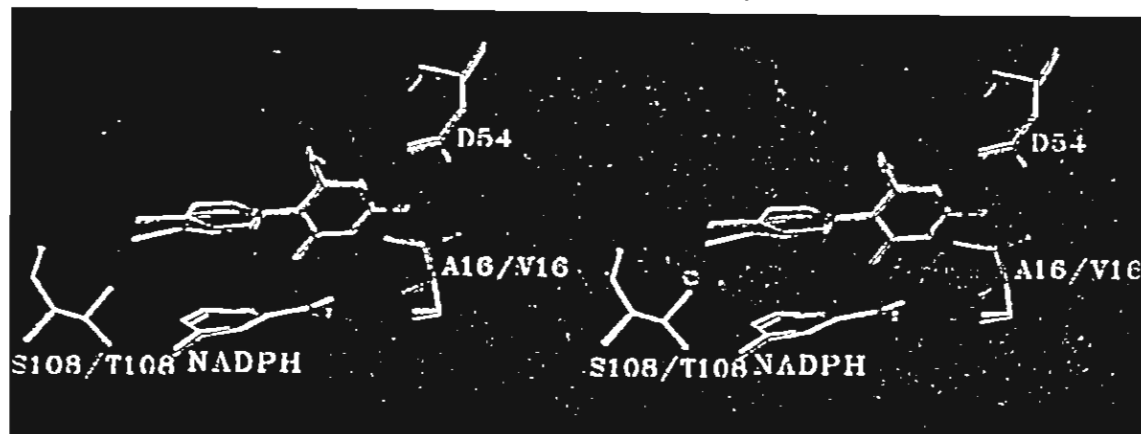
The dihydrotriazine rings of WR99210 and Cyc are identical and therefore are predicted to bind the A16V mutant pDHFR in the same orientation. As a consequence, the 2,2'-dimethyl substituents of WR99210 should also be in steric conflict with the valine side chain of the A16V mutant. Figure 5C illustrates the superimpositions of WR99210 bound to wild-type, A16V, S108T and A16V+S108T double mutant DHFR models. A slight displacement from its optimal position occupied in the wild-type complex was noted when the inhibitor was bound to the pDHFRs with A16V mutation. The model shows that the trichlorophenoxy

propyloxy substituent of WR99210 is not forced to change its orientation significantly in the A16V mutant. Instead, its flexible side chain allows it to be free from being locked between the residues A16, D54, nicotinamide, and S108 as observed in the structure of the wild-type complex. Therefore it is likely that the steric demand imposed by V16 would not significantly affect the orientation of the drug upon binding to the pDHFRs with the A16V mutation (Fig. 5C).

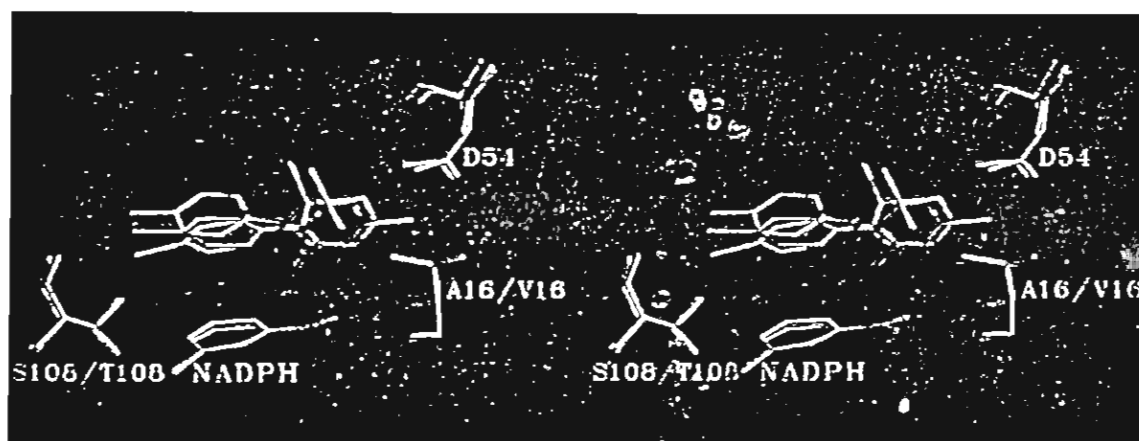
The bindings of Pyr, Cyc, and WR99210 to the mutant model of DHFR with single S108T mutation were also investigated. Generally, all the three inhibitors bind to the S108T mutant model with similar orientation as in the wild-type complex, albeit with some minor differences. The bulkier side chain of the T108 mutant slightly affects the position of the chlorophenyl ring of Pyr and Cyc (Fig. 5A and B). However, the orientation of the propyloxy substituent of WR99210 is not affected upon binding to pDHFR with the S108T mutation (Fig. 5C).

The model indicates that the 6-ethyl substituent of Pyr is not in steric conflict with A16V+S108T pDHFR despite the locking of the inhibitor between residues 16, 54, nicotinamide, and 108. Therefore, Pyr is predicted to bind the double mutant DHFR with no appreciable differences as compared to the binding to the wild-type, A16V, and S108T single mutant enzymes (Fig. 5A). As previously illustrated for the A16V and S108T mutant, the steric conflict with V16 must be relatively small and therefore no significant displacement of the inhibitors was detected. The situation is very different for Cyc. In the A16V+S108T double mutant, however, the chlorophenyl ring is further displaced from nicotinamide owing to steric conflicts with T108, while the position of the dihydrotriazine ring of Cyc is not different to that of the A16V mutant (Fig. 5B). Binding of WR99210 is similar to Cyc in its dihydrotriazine moiety but the inhibitor is not locked toward nicotinamide and is also much more flexible. As a result, slight displacement of the dihydrotriazine ring has no consequences for the overall binding in the double mutant enzyme (Fig. 5C). The steric clash between the V16 side chain of the A16V mutant pDHFR is clearly evidenced for Cyc binding, but not for Pyr binding (Fig. 6).

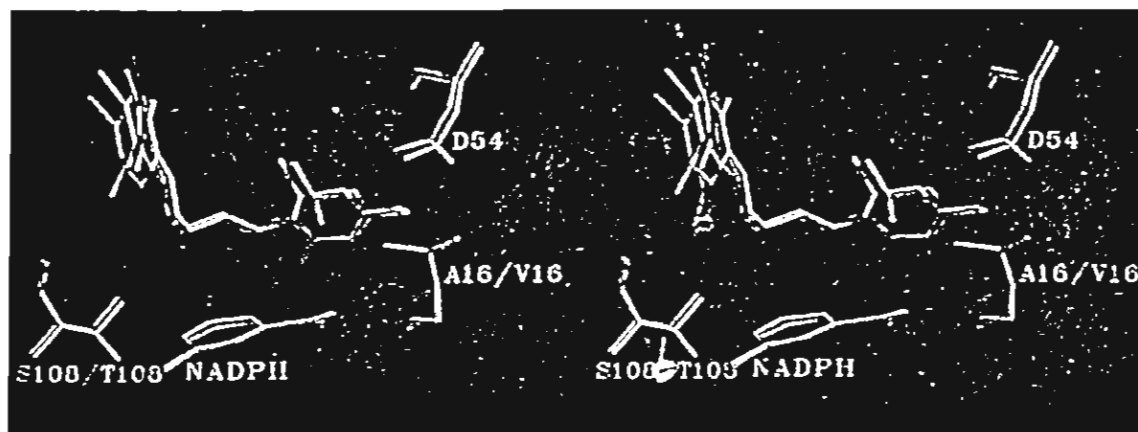
In order to verify the homology model of *P. falciparum* DHFR and to exploit the model for the prediction of binding of inhibitors and/or development of strategies for the design of the best effective inhibitor(s) targeted against antifolate resistant malaria, we have synthesized a number of Cyc derivatives and tested against both wild-type and A16V+S108T mutant pDHFRs. Our premise is that resistance to Cyc in the A16V+S108T mutant is the result of steric conflict imposed by a bulkier side chain of V16. To test the hypothesis, Cyc derivatives with one substituent (ethyl or methyl group) or no substituent at position R₂, and those with bulky substituents at both R₁ and R₂ positions of the dihydrotriazine ring were synthesized. The binding affinity (*K_i*) values were then determined and the ratios of *K_i*-mut/*K_i*-wt were compared. Table 1 summarizes the *K_i*



(A)



(B)



(C)

Figure 5. Superimpositions of inhibitors bound to wild-type (yellow), A16V (red), S108T (cyan) and A16V + S108T (white) pDHFRs. The inhibitors shown are (A) Pyr; (B) Cyc; (C) WR99210.

values for the wild-type and the A16V + S108T mutant DHFRs. The relative loss in binding affinity of the Cyc derivatives is reflected by the elevation of $K_i\text{-mut}/K_i\text{-wt}$ ratio. As shown in Table 1, the compounds II and III inhibited the wild-type pDHFR with K_i values 2–3 times higher than that of Cyc, but exhibited 7–10 times lower K_i values for the A16V + S108T mutant pDHFR

than the Cyc. As a result, the $K_i\text{-mut}/K_i\text{-wt}$ ratios for the compound II and III were ~ 17 – 28 times lower than that of the Cyc parent compound. Removal of both methyl groups at the 2- R_1 and 2- R_2 positions gave rise to compound IV (didesmethyl Cyc), which has a K_i value for the wild-type enzyme about 16-fold higher than that of Cyc, whereas the K_i value for the A16V + S108T enzyme

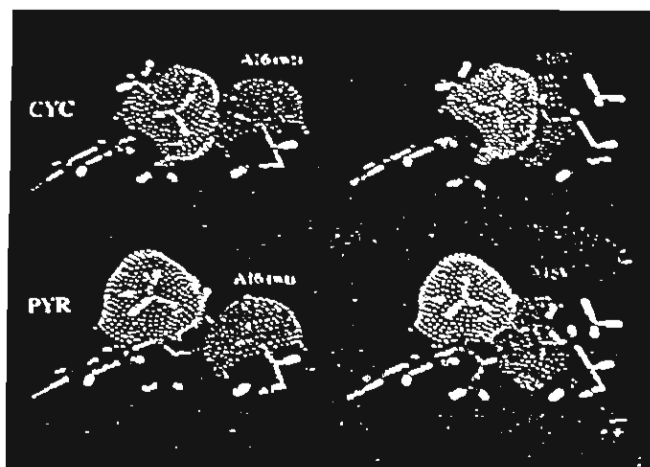


Figure 6. Display of the surface of the side chain residue of amino acid A16 (red) and A16V (purple) interacting with one of the 2,2'-methyl substituents of Cyc (yellow) and 6-ethyl substituent of Pyr (yellow).

was about 2-fold lower than Cyc. As a consequence, the ratio of K_i -mut/ K_i -wt for the compound IV was 34-fold lower than that for Cyc (Table 1).

Our studies were further extended to investigate the binding between Cyc derivatives with bulky 2,2' di-substituents and the wild-type and A16V+S108T mutant DHFRs. Nine compounds (compounds V–XIII) were selected for the evaluation. Except for compounds IX and XIII, which exhibited high K_i values, all other seven derivatives gave K_i values for the wild-type enzyme that were comparable to that for Cyc. When tested with the A16V+S108T enzyme, the K_i values of these derivatives except compound VIII were all higher than that of Cyc, resulting in relatively high ratios of K_i -mut/ K_i -wt.

The K_i values for WR99210 (compound XIV, Table 2) and other Cyc derivatives were compared with their 2-monomethyl derivatives (Table 2). Removal of one methyl from position 2-R₁ of WR99210 (compound

XV, Table 2) resulted in a 2- and 3-fold decrease in the K_i values for the wild-type and the A16V+S108T enzymes, respectively. The compound XVI with 4'-Br-C₆H₄ at 3-R position (Table 2) only slightly increased the K_i for the wild-type pDHFR by 2-fold, but dramatically elevated the K_i for the A16V+S108T enzyme by 800-fold. The steric conflict observed in the compound XVI was relieved by replacing the methyl group at position 2-R₁ with hydrogen. The resulting compound XVII was then found to exhibit a 10-fold lower K_i for the A16V+S108T enzyme as compared to the compound XVI, albeit the manipulation slightly affected the binding affinity of the inhibitor to the wild-type enzyme. Likewise, similar effect was also found for compounds XVIII and XIX, where changes were made at the positions 2-R₁ and 3-R (Table 2).

Discussion

Resistance to antifolates Pyr and Cyc in malaria parasites has been shown to be associated with point mutations in the *dhfr* gene.^{1,5,7,46} Cross resistance to both drugs is generally found to involve combinations of N51I, C59R, S108N, and I164L mutations, whereas parasites resistant to Cyc but susceptible to Pyr are associated with mutation at residues 16 and 108 (A16V+S108T). In the present study, we sought to understand the molecular interactions between the DHFR of *P. falciparum* and antimalarial antifolates Pyr, Cyc, and WR99210, to predict the effect of mutation(s) on the binding of the inhibitors to the enzyme, and to address the importance of the mutation(s) at different residues in conferring resistance to the antifolates.

We describe a homology model of *P. falciparum* DHFR based on sequence alignment of the DHFR sequences of *E. coli*, *L. casei*, human, chicken liver, and *P. carinii*. Our model was built using the conserved regions of a number of known DHFR templates, and the structures of the unique regions in pDHFR were subsequently assigned to avoid undesirable perturbation of

Table 1. K_i values for the cycloguanil derivatives with only one or no substituent at position R₁ (compounds I–IV) and bulky substituents at positions R₁ and R₂ (compounds V–XIII)^a

Compound	2-R ₁	2-R ₂	K_i -wt ^b (nM)	Rel to Cyc	K_i -mut ^c (nM)	Rel to Cyc	K_i -mut ^c / K_i -wt ^b
I	Me	Me	1.5 ± 0.3 ^d	1.0	1314 ± 165 ^d	1.0	876
II	H	Me	4.1 ± 0.0	2.7	127 ± 14	0.1	31
III	H	Et	3.6 ± 0.0	2.4	189 ± 37	0.14	53
IV	H	H	24.4 ± 4.3	16.3	646 ± 78	0.5	26
V	Me	<i>n</i> -Pr	3.5 ± 0.4	2.3	9229 ± 547	7.0	2637
VI	Me	<i>n</i> -Bu	1.8 ± 0.2	1.2	4217 ± 390	3.2	2343
VII	Me	<i>n</i> -Pen	1.2 ± 0.1	0.8	3594 ± 446	2.7	2995
VIII	Me	<i>n</i> -Hex	0.6 ± 0.1	0.4	956 ± 78	0.7	1593
IX	Me	<i>i</i> -Pr	36.5 ± 4.1	24.3	44,791 ± 5872	34.1	1227
X	Me	<i>i</i> -Bu	4.7 ± 1.9	3.1	15,263 ± 986	11.6	3247
XI	Me	<i>i</i> -Pen	1.7 ± 0.1	1.1	5755 ± 545	4.4	3385
XII	Et	Et	7.7 ± 1.2	5.1	16,151 ± 3599	12.3	2098
XIII	Et	<i>n</i> -Pen	19.3 ± 3.8	12.9	18,041 ± 1905	13.7	935

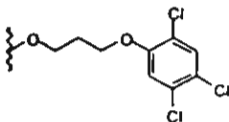
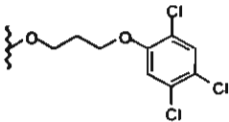
^aMe, methyl; Et, ethyl; Pr, propyl; Bu, butyl; Pen, pentyl; Hex, hexyl.

^bWild-type pDHFR.

^cpDHFR (A16V+S108T).

^dData from Sirawaraporn et al.¹⁰

Table 2. K_i values for the WR99210 and other cycloguanil derivatives compared with their 2-monomethyl derivatives

Compound	3-R	2-R ₁	2-R ₂	K_i -wt ^a (nM)	Rel to WR99210	K_i -mut ^b (nM)	Rel to WR99210	K_i -mut ^b / K_i -wt ^a
XIV (WR99210)		Me ^c	Me	0.5 ± 0.1	1.0	2.4 ± 0.4	1.0	4.8
XV		H	Me	0.3 ± 0.0	0.6	0.8 ± 0.1	0.3	2.7
XVI	4'-Br-C ₆ H ₄	Me	Me	1.1 ± 0.2	2.2	1948 ± 366	812	1771
XVII	4'-Br-C ₆ H ₄	H	Me	5.7 ± 0.5	11.4	202 ± 17	84	35
XVIII	4'-Me-C ₆ H ₄	Me	Me	1.8 ± 0.2	3.6	1584 ± 210	660	880
XIX	4'-Me-C ₆ H ₄	H	Me	23.4 ± 1.9	45.8	186 ± 22	78	8.0

^aWild-type pDHFR.^bpDHFR (A16V + S108T).^cMe, methyl.

the overall conformation of the structure. The overall topology of the pDHFR is similar to the DHFRs from other sources (Fig. 2), except for its two extra-loop insertions located at the enzyme surface. Unlike those previously reported by other workers,^{18,48} the model presented in this paper was built based not only on the alignments among the DHFR sequences, but also on their 3D-structural conservation. The homology noted at least from residue 10, therefore, makes the alignment more meaningful and enables analysis of residue 16 which is crucial for Cyc-resistance. In the present study, the structures of A16V and A16V + S108T mutant pDHFRs were also built from the wild-type template and docking of the antifolates Pyr, Cyc, and WR99210 was performed using both wild-type and mutant structures.

Our structural models show that the ethyl and methyl substituents of Pyr and Cyc, respectively, are directed towards the methyl side chain of A16 in the wild-type pDHFR. As a result, it is very likely that the bulkier side chain of valine in the A16V mutant of pDHFR would cause steric conflicts with these substituents, and therefore may explain Cyc resistance for these mutants. However, such effect does not occur in the case of Pyr; the inhibitor binds equally well with the same orientation and binding interactions to A16V + S108T mutant pDHFR as well as the wild-type enzyme, evidence which is in accord with the earlier observations by many groups that the A16V + S108T mutant is susceptible to Pyr.^{5,7,10,11}

Although Cyc is predicted to bind to the wild-type pDHFR in a similar way to Pyr, one of its two 2,2'-dimethyl substituents is significantly closer to the side chain of residue 16 than is the ethyl group of Pyr. This would inevitably generate steric constraint for binding to Cyc when the bulkier side chain of valine is present at position 16 as in the A16V and A16V + S108T mutants. Since Cyc is predicted to bind in a locked region which is composed of A16 and D54 for the dihydrotriazine

ring, and nicotinamide and S108 for the chlorophenyl ring of the inhibitor, A16V + S108T mutation would result in steric constraint and cause displacement of both the dihydrotriazine moiety and the chlorophenyl ring from its optimal position as depicted for the wild-type enzyme, giving rise to a remarkable loss of binding affinity for Cyc and hence resistance to Cyc for the A16V + S108T mutant pDHFR.

The displacement of WR99210, however, is relatively minor when compared with Cyc. This is likely due to the reason that the more flexible trichlorophenyl propoxy side chain of WR99210 is not locked between nicotinamide and S108 as observed for the chlorophenyl ring of Cyc. Therefore, the steric demand imposed by valine 16 only slightly affects the orientation of WR99210 and leaves unaffected binding interactions with other protein residues, in agreement with the finding that WR99210 is highly effective against the resistant mutants.

Our model predicts that binding of the derivatives with only one or no substituent at position 2 should be less affected by the mutation than Cyc. This would be reflected in the values of the ratios of the K_i for the A16V + S108T mutant enzyme and the wild-type enzyme (K_i -mut/ K_i -wt). In order to test this model, a number of Cyc derivatives were synthesized and the binding affinities to the wild-type and mutant enzymes were studied. It was found that K_i -mut/ K_i -wt values for the derivatives with only one or no substituent at position 2 are lower than for Cyc, as predicted (compounds II–IV, Table 1). Furthermore, 2,2'-disubstituted derivatives of Cyc with bulky substituents tend to have higher K_i -mut/ K_i -wt ratios, as expected (compounds V–XIII, Table 1). The favorable effect on binding of the removal of the 2-methyl group was also found for WR99210 (Table 2). However, this effect is far less than for Cyc derivatives with p-chloro-, bromo- or methyl-phenyl substituents, because WR99210 is not locked, and can already interact well with the mutant enzyme even with

both 2-methyl groups present. The overall results obtained from this study are in support of the steric constraint hypothesis. Further it reveals the important steric role in binding between the substituents at position 2 of Cyc to the active residue(s) of the p fDHFR , which could have profound implications for the development of new and effective inhibitors against antifolate resistant parasites.

Conclusion

The homology models of wild-type, A16V, S108T, and A16V+S108T *P. falciparum* DHFRs were built based on multiple sequence alignment and homology modeling procedures. The models were used to address the molecular interactions between the active site residues of the enzymes and antifolates Cyc, Pyr and WR99210. We proposed a 'steric constraint' hypothesis and demonstrated its validity by testing with inhibitor analogues using the wild-type and relevant mutant p fDHFR s. The models not only explains the Cyc resistance but also Pyr susceptibility in relation to the A16V mutation of the p fDHFR domain. In addition, it has allowed us for the first time to better understand the molecular basis of cross resistance between Cyc and Pyr, which have been known to involve mutations at residues 51, 59, 108 and 164 of the p fDHFR domain. Further, the modeling of the interaction between p fDHFR and WR99210 has also provided ideas that could lead to the development of new and more effective second generation antifolate antimalarials.

Acknowledgements

This research is supported by grants from World Health Organization (TDR) to WS, the European Union (INCO-DC IC18CT970223) to YY and GL, and The University of Modena (Ricerca Orientata) to GR. TV is supported by The Thailand Research Fund (PDF/57/2540). We thank Rachada Sirawaraporn and Kaneung Chitcholtan for their excellent technical assistance.

References

1. Cowman, A. F.; Morry, M. J.; Biggs, B. A.; Cross, G. A. M.; Foote, S. J. *Proc. Natl. Acad. Sci. USA* 1988, 85, 9109.
2. Snewin, V. A.; England, S. M.; Sims, P. F. G.; Hyde, J. E. *Gene* 1989, 76, 41.
3. Basco, L. K.; de Pecoulas, P. E.; Le Bras, J.; Wilson, C. M. *Exp. Parasitol.* 1996, 82, 97.
4. Basco, L. K.; de Pecoulas, P. E.; Wilson, C. M.; Le Bras, J.; Mazabraud, A. *Mol. Biochem. Parasitol.* 1995, 69, 135.
5. Foote, S. J.; Galatis, D.; Cowman, A. F. *Proc. Natl. Acad. Sci. USA* 1990, 87, 3014.
6. Hyde, J. *Parasitol. Today* 1989, 5, 252.
7. Peterson, D. S.; Walliker, D.; Wellems, T. E. *Proc. Natl. Acad. Sci. USA* 1988, 85, 9114.
8. Thaithong, S.; Chan, S.-W.; Songsomboon, S.; Wilairat, P.; Seesod, N.; Sueblinwong, T.; Goman, M.; Ridley, R.; Beale, G. *Mol. Biochem. Parasitol.* 1992, 52, 149.
9. Zolg, J. W.; Plitt, J. R.; Chen, G.-X.; Palmer, S. *Mol. Biochem. Parasitol.* 1989, 36, 253.
10. Sirawaraporn, W.; Sathitkul, T.; Sirawaraporn, R.; Yuthavong, Y.; Santi, D. *Proc. Natl. Acad. Sci. USA* 1997, 94, 1124.
11. Peterson, D. S.; Milhous, W. K.; Wellems, T. E. *Proc. Natl. Acad. Sci. USA* 1990, 87, 3018.
12. Childs, G. E.; Lambros, C. *Ann. Trop. Med. Parasitol.* 1986, 80, 177.
13. Hekmat-Nejad, M.; Rathod, P. K. *Exp. Parasitol.* 1997, 87, 222.
14. Sanchez, R.; Sali, A. *Curr. Opin. Struct. Biol.* 1997, 7, 206.
15. Toyoda, T.; Brobey, R. K. B.; Sano, G.-I.; Horii, T. *Biochem. Biophys. Res. Commun.* 1997, 235, 515.
16. McKie, J. H.; Douglas, K. T.; Chan, C.; Roser, S. A.; Yates, R.; Read, M.; Hyde, J. E.; Dascombe, M. J.; Yuthavong, Y.; Sirawaraporn, W. *J. Med. Chem.* 1998, 41, 1367.
17. Warhurst, D. C. *Drug Discovery Today* 1998, 3, 538.
18. Bzik, D. J.; Li, W.-B.; Horii, T.; Inselburg, J. *Proc. Natl. Acad. Sci. USA* 1987, 84, 8360.
19. Reyes, V. M.; Sawaya, M. T.; Brown, K. A.; Kraut, J. *Biochemistry* 1995, 34, 2710.
20. Bolin, J. T.; Filman, D. J.; Matthews, D. A.; Hamlin, R. C.; Kraut, J. *J. Biol. Chem.* 1982, 257, 13650.
21. Oefner, C.; D'Arcy, A.; Winkler, F. K. *Eur. J. Biochem.* 1988, 174, 377.
22. Lewis, W. S.; Cody, V.; Galitsky, N.; Luft, J. R.; Pangborn, W.; Chunduru, S. K.; Spencer, H. T.; Appleman, J. R.; Blakley, R. L. *J. Biol. Chem.* 1995, 270, 5057.
23. McTigue, M. A.; Davies II, J. F.; Kaufman, B. T.; Kraut, J. *Biochemistry* 1992, 31, 7264.
24. Champness, J. N.; Achari, A.; Ballantine, S. P.; Bryant, P. K.; Delves, C. J.; Stammers, D. K. *Structure* 1994, 2, 915.
25. Sali, A.; Blundell, T. L. *J. Mol. Biol.* 1993, 234, 779.
26. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* 1990, 215, 403.
27. Laskowski, R. A.; McArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* 1993, 26, 283.
28. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatman III, T. E.; Ferguson, D. M.; Kollman, P. A. *AMBER 4.1*. University of California: San Francisco, 1995.
29. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* 1995, 117, 5179.
30. van Gusteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* 1977, 34, 1311.
31. Penner, M. H.; Frieden, C. *J. Biol. Chem.* 1985, 260, 5366.
32. Fierke, C. A.; Johnson, K. A.; Benkovics, S. J. *Biochemistry* 1987, 26, 4085.
33. Appleman, J. R.; Howell, E. E.; Kraut, J.; Blakley, R. L. *J. Biol. Chem.* 1990, 265, 5579.
34. Costantino, L.; Rastelli, G.; Vescovini, K.; Cignarella, G.; Vianello, P.; Corso, A. D.; Cappiello, M.; Mura, U.; Barlocco, D. *J. Med. Chem.* 1996, 39, 4396.
35. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* 1993, 97, 10269.
36. Cieplak, P.; Bayly, C. I.; Cornell, W. D.; Kollman, P. A. *J. Comput. Chem.* 1995, 16, 1357.
37. Jørgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* 1983, 79, 926.
38. Cocco, L.; Roth, B.; Temple, Jr., C.; Montgomery, J. A.; London, R. E.; Blakley, R. L. *Arch. Biochem. Biophys.* 1983, 226, 567.
39. Cocco, L.; Groff, J. P.; Temple, Jr., C.; Montgomery, J. A.; London, R. E.; Matwiyoff, N. A.; Blakley, R. L. *Biochemistry* 1981, 20, 3972.
40. Varnex, A.; Helissen, S.; Wipff, G.; Collet, A. *J. Comput. Chem.* 1998, 19, 820.

41. Carrington, H. C.; Crowther, A. F.; Stacey, G. J. *J. Chem. Soc.* 1954, 1017.
42. Colebrook, L. D.; Giles, H. G.; Rosowsky, A.; Bentz, W. E. *Can. J. Chem.* 1976, 54, 3757.
43. Modest, E. J. *J. Org. Chem.* 1956, 21, 1.
44. Modest, E. J.; Levine, P. J. *J. Org. Chem.* 1956, 21, 14.
45. Mamalis, P.; Green, J.; Outred, D. J.; Rix, M. *J. Chem. Soc.* 1962, 3915.
46. Sirawaraporn, W.; Prapunwattana, P.; Sirawaraporn, R.; Yuthavong, Y.; Santi, D. V. *J. Biol. Chem.* 1993, 268, 21637.
47. Blakley, R. L.; Benkovics, S. J. In *Folates and Pteridines*; Blakley, R. L.; Benkovics, S. J., Eds.; Wiley: New York, 1984; Vol. 1, pp 191–253.
48. Lemcke, T.; Christensen, I. T.; Jørgensen, F. S. *Bioorg. Med. Chem.* 1999, 7, 1003.
49. Smith, D. R.; Calvo, J. M. *Nucleic Acids Res.* 1980, 8, 2255.
50. Freisheim, J. H.; Bitar, K. G.; Reddy, A. V.; Blankenship, D. T. *J. Biol. Chem.* 1978, 253, 6437.
51. Masters, J. N.; Attardi, G. *Gene* 1983, 21, 59.
52. Chen, M.-J.; Shimada, T.; Moulton, A. D.; Cline, A.; Humphries, R. K.; Maizel, J.; Nienhuis, A. W. *J. Biol. Chem.* 1984, 259, 3933.
53. Kumar, A. A.; Blankenship, D. T.; Kaufman, B. T.; Freisheim, J. H. *Biochemistry* 1980, 19, 667.
54. Edman, J. C.; Edman, U.; Cao, M.; Lundgren, B.; Kovacs, J. A.; Santi, D. V. *Proc. Natl. Acad. Sci. USA* 1989, 86, 8625.

Molecular Dynamics Simulation of the Human Apo-dihydrofolate Reductase: An Investigation of an Unstable Enzyme

PORNTHIEP SOMPORNPIST^{a,*}, ATCHARA WIJITKOSOOM^a, VUDHICHAI PARASUK^a and WORACHART SIRAWARAPORN^b

^aDepartment of Chemistry, Faculty of Science, Chulalongkorn University, Phayathai Road, Bangkok 10330, Thailand; ^bDepartment of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

(Received April 2001; In final form July 2001)

Structure and dynamics data of the *Homo sapiens* apo-dihydrofolate reductase (hDHFR) have been investigated by molecular dynamics (MD) simulation. The MD simulation for a total of 300 ps was performed for the system containing the apo-hDHFR with explicit water molecules. The last 250 ps MD trajectory of the equilibrated system was generated and the 25 sampling structures of the apo-hDHFR were analyzed. The statistical quantities of the simulation reveal significant differences between the apo-enzyme and the complex form. The differences include the mainchain mobility, the solvent accessibility, the secondary structures, the phenylalanine movement and the hydrogen bond may relate to the stability of the enzyme. The tertiary folding from the secondary structure motif reveals a less compact for the apo-enzyme structure with respect to the X-ray structure of the hDHFR-folate complex. In addition, intramolecular hydrogen bond involving the backbone protons dramatically decreases. Comparisons between the unbound and the ligand-bound enzyme obtained from the X-ray, NMR and MD data are discussed.

Keywords: Molecular dynamics; Dihydrofolate reductase; Apo-enzyme; Root mean square deviation

1. INTRODUCTION

Knowledge of the structure and dynamics of biological macromolecules at an atomic level is essential for understanding their functional mechanism in biochemical process. To reach such fundamental relationship, molecular dynamics (MD) calculations have considerably become one of the powerful techniques for investigating dynamics

information from atomic coordinates [1,2]. This computational tool provides structural statistics that link to its biological task in physiological conditions. This technique has also great potential to extend details that the experimental observations are missing due to technical difficulty or impracticality. The MD applications have successfully revealed the structure-function relationship for numerous proteins, including the enzyme dihydrofolate reductase (DHFR:E.C.1.5.1.3).

DHFR has been a remarkable attention in molecular and structural research because of its biological and clinical importance. DHFR is a crucial enzyme responsible for the biosynthesis of nucleotide in living organisms. The biological role of this enzyme is to catalyze the NADPH-dependent reduction of 7, 8-dihydrofolate or folate to 5,6,7,8-tetrahydrofolate required in several biochemical processes including biosynthesis of thymidylate, purines, and certain amino acids. For the clinical usefulness, DHFR has been significantly useful as a drug target of antifolate inhibitors for a number of diseases. This includes anticancer (methotrexate, MTX), antibacterial infection (trimethoprim, TMP) and antimalaria (pyrimethamine, PYR) and cycloguanil, CYC, etc [3].

Because of the appropriate size of the enzyme molecule (15–30 kD), it is, therefore, amenable for the structural study using X-ray crystallography and NMR spectroscopy. The three-dimensional structures of the DHFR wildtype and the drug resistant

*Corresponding author. Address: Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330 Thailand. Tel.: 662-2185224. Fax: 662-2541309. E-mail: spornthe@chula.ac.th

mutants in protozoal, bacterial and vertebrate DHFRs have been reported. These include *Leishmania major* [4], *Lactobacillus casei* [5,6], *Escherichia coli* [5,7,8], *Pneumocystis carinii* [9], chicken liver [10–13] and human [14–16]. The atomic coordinates of those DHFR structures have shown that the tertiary folding pattern of the enzyme is similar and the secondary structures are very well conserved. Despite the similarity of their overall folding topology, subtle differences do exist, particularly residues surrounding the substrate-binding site. Such differences are crucial for enzyme selectivity and specificity to small organic molecules such as the substrate or antifolate inhibitors [17]. This requires atomic details associated with the dynamics information in order to design and develop novel and more effective inhibitors.

Most of DHFR studies reported so far were conducted in the presence of small ligand molecules such as the folate substrate or antifolate inhibitor. The fact is that the binary or the ternary enzyme complexes are the most interest for studying the protein-ligand interactions. Moreover, the binding of these small molecules is strong enough to stabilize the enzyme throughout experimental measurements. Without the substrate/inhibitor or, many experimental techniques seem to fail due to protein instability [18]. What causes the enzyme in the absence of ligands to be unstable? Are there any conformational changes that could lead the apo-enzyme to unfold?

Here, we have performed the MD simulation of the apo-enzyme DHFR from *Homo sapiens* (hDHFR) in aqueous solution at 300 K, with the objectives (1) to obtain dynamics information of the apo-enzyme; (2) to extend the knowledge of the crystal structure of the human DHFR; (3) to relate some specific properties of the calculation to experimentally observable quantities for the system and (4) to find out possible evidence influencing the stability of the enzyme. The hDHFR is a monomeric protein consisting of 186 amino acid residues (21.5 kD) that make possible to pursuing MD study with a timescale of hundreds picosecond. Its 3D crystal structures with the resolution ranging from 2.0 to 2.5 Å have been solved by the X-ray diffraction techniques. The resonance assignments of the backbone and the side-chain protons of the hDHFR complexed with methotrexate in aqueous solution were achievable using multidimensional NMR spectroscopy. The NMR structure of the hDHFR-MTX-NADPH complex has been obtained from NOEs restraints around protein-ligand binding region [19]. Due to the poor stability of the apo-hDHFR, it is therefore not possible for structural studies through X-ray and NMR techniques. Availability from the X-ray and NMR data of the hDHFR

complex is helpful for the purpose of comparison with the MD results.

2. COMPUTATIONAL DETAILS

The single crystal structure (Fig. 1) of the mammalian DHFR complexed with the folate refined at 2.0 Å resolution [14] was taken from the Brookhaven Protein Data Bank (<http://www.rcsb.org> and pdb ID: 1drf). The structure was then used as the starting conformation of the study. The atomic coordinates belonging to the folate, water and sulfate molecules were removed. In order to obtain the full protein structure, the missing sidechain atoms of Gln12 were built on the basis of the internal coordinates of glutamine residue in the Amber amino acid library. Addition of hydrogen coordinates was considered the ionization-state of the charged residues including Glu, Asp, Arg and Lys as well as the positive and negative charges at N-terminal Val1 and C-terminal Asp186. Energy minimization of the hDHFR molecule was subsequently performed to improve the poorly defined regions present in the protein structure as a consequence of the geometrically inappropriate construction of added atoms.

All energy calculations were conducted using the Cornell *et al.* all-atom force field model available in the program Amber [20]. The steepest descent algorithm was applied through the first 5000 steps and then followed by the conjugated gradient algorithm for the remaining minimization. The convergence criterion of $1.0 \times 10^{-4} \text{ kcal mol}^{-1} \text{ Å}^{-1}$ was chosen for the energy gradient otherwise the maximum minimization step is reached. The treatment of

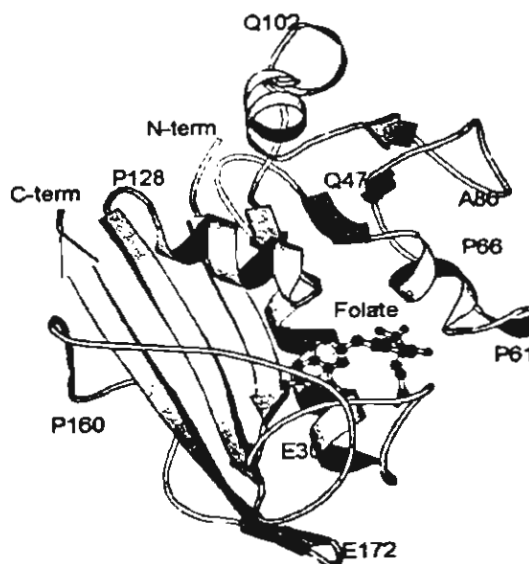


FIGURE 1 Schematic representation of the X-ray structure (1drf) showing the secondary structure elements and the substrate folate. The figure was generated using MOLSCRIPT [27].

TABLE I Relevant simulation parameters for the MD simulation

<i>Protein-solvent simulation condition</i>	<i>Simulation data</i>
<i>Energy minimization</i>	
In vacuo	60,000 steps (5,000 steps steepest descent and 55,000 conjugated gradient)
In water	60,000 steps (5,000 steps steepest descent and 55,000 conjugated gradient)
<i>Molecular dynamics</i>	
Number of atoms	19933 (atoms) 3013 protein atoms and 5640 TIP3P water molecules
Thermodynamic ensemble	Isobaric-isothermal ensemble (NPT)
Integration algorithm	Leapfrog
Integration time step	0.001 ps
Constraints	No
Long-range interactions	10 Å residue cut off distance
Periodic boundary conditions	Box dimension 64.8 × 57.6 × 52.8 Å ³
Thermalization	15 ps
Equilibration	35 ps
Dynamics trajectory for structural evaluation	250 ps
Collection of coordinate sets	2500 substructures (every 0.1 ps)

long-range electrostatic interactions was employed by a residue-based cutoff value of 10.0 Å.

The energy-minimized coordinates were placed afterwards in a rectangular box of water. To employ the method of periodic boundary conditions, the cubic dimension of 64.8 Å × 57.6 Å × 52.8 Å was defined with a total of 19933 atoms, including 3013 protein atoms solvated by 5640 explicit water molecules. This represents sufficient numbers of solvent molecules as to make the results obtained meaningful. The interactions involving the solvent molecules were modeled by the well-known three-center charge TIP3P model [21]. Since the simulation of this step was carried out in the explicit solvent water molecules, the dielectric constant of 1 was used. The periodic boundary conditions were employed throughout the simulations. Another energy minimization of the protein in aqueous solution was performed.

The MD simulation was carried out in subsequent calculations. An isobaric-isothermal ensemble (NPT) was employed using constant pressure of 1 atm and constant temperature of 300 K. The pressure was kept constant by coupling to an external pressure bath [22] with a coupling constant of 0.2 ps. The velocities of the atoms were assigned according to a Maxwellian velocity distribution at 300 K. The integration of the equation of motion was done using a leapfrog algorithm employing the time step of 1 fs. For the first 50 ps of the MD simulation, an equilibration of the protein in solution was reached after warming up the system for 15 ps. During the thermalization period, the temperature was adjusted with the SHAKE option using the Berendsen coupling algorithm. A single scaling factor was used for all atoms at 300 K. A set of the coordinates containing enzyme structure with bulk solvent water molecules, so called MD trajectory, was collected every 0.1 ps over the last 250 ps and was used to

monitor the dynamics behavior of the simulation. The overall behavior of the MD simulation was analyzed by assessments of the fluctuation of total energy, kinetic energy, potential energy, pressure, volume, temperature of the solution and the deviation of the atomic positions (or a root-mean-square deviation, RMSD) as a function of time. RMSD calculation for hDHFR did not take into consideration the effect of the N- and C-terminal, neglecting the first and the last three residues. Relevant simulation parameters were summarized in Table I. All MD simulation and analysis of MD trajectory were performed using the program Amber 5 [23]. The three-dimensional structures were visualized using the program Swiss-PdbViewer [24], Rasmol [25], Molmol [26] and Molscrip [27]. Evaluation of the quality of the protein structure was performed using Procheck [28]. The simulations were performed on an SGI Power Challenge XL 8 × R10000.

3. RESULTS AND DISCUSSION

3.1. Energy Minimization

The energy minimization profiles of the apo-enzyme hDHFR in the absence of water molecules reveal a rapid drop from 4.60×10^3 kcal mol⁻¹ to -6.0×10^3 kcal mol⁻¹ and then insignificantly alter. The final potential energy of the minimized enzyme structure was -6.17×10^3 kcal mol⁻¹. The potential energy of the system containing the apo-enzyme and water molecules starts with -5.33×10^4 kcal mol⁻¹ and reached a value of -7.82×10^4 kcal mol⁻¹ when minimized. The energy profile curves of the apo-enzyme both in the presence and absence of water were similar, implying that the minimization significantly decreased the strain of the system.

The system was gradually relaxed and ready to perform MD simulation for the next investigation.

3.2. Molecular Dynamics

3.2.1. Analysis of the MD Trajectory

In a molecular dynamics simulation, the trajectory of coordinates of hundreds picosecond provides useful information that some statistical quantities can be extracted. Figures 2A and B illustrates the energy profiles (total energy, kinetic energy, potential energy) and temperature of the simulated system for 300 ps, respectively. As a typical MD simulation, the results are well behaved. We observed rapid initial equilibration and slight fluctuations over a longer simulation time scale. The overall MD profiles obtained were most common to those observed in other protein simulations [29,2]. Analysis of properties of the system performed after the first 50 ps and the final MD parameters averaged over the last 250 ps are summarized in Table II.

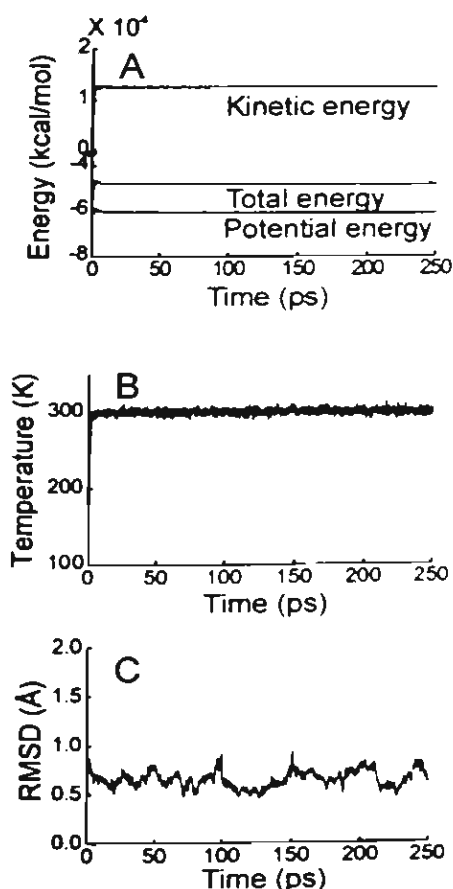


FIGURE 2 MD simulation profiles for the total energy, the potential energy, the kinetic energy (A), the temperature (B), the mainchain fluctuation of the enzyme with respect to that of the average energy-minimized structure (C).

TABLE II The average values from the MD simulation

MD parameters	Mean values
Total energy ($\times 10^4$ kcal mol $^{-1}$)	-4.83 ± 0.009
Kinetic energy ($\times 10^4$ kcal mol $^{-1}$)	1.26 ± 0.006
Potential energy ($\times 10^4$ kcal mol $^{-1}$)	-6.09 ± 0.011
Temperature of the enzymes (K)	301 ± 4

For the last 250 ps of the MD trajectory, the apo-hDHFR coordinates of all collected configurations were used to generate an average structure, which is in turn used to calculate the RMSD of atomic coordinates. Figure 2C shows a plot of the mean global RMSD with respect to the ensemble average structure versus the simulation time of 250 ps. The RMSD for backbone atoms of residues Ser3-Glu 183 over the MD trajectory was in a range of 0.50–0.90 Å. This structural fluctuation is not uncommon in the typical MD simulation of protein, indicating the reliable equilibration of the system in this study.

3.2.2. Sampling Configurations and Evaluation

A set of 25 snapshot structures taken every 10 ps from the 250 ps MD trajectory is shown in Fig. 3A. A superposition of the 25 substructures manifests the mean global RMSD of 1.38 ± 0.17 Å for the backbone atoms and 1.72 ± 0.15 Å for the heavy atoms. As shown in Fig. 3B, the conformations of 25 MD structures superimposed on each other were considerable similar. Three segments with the largest

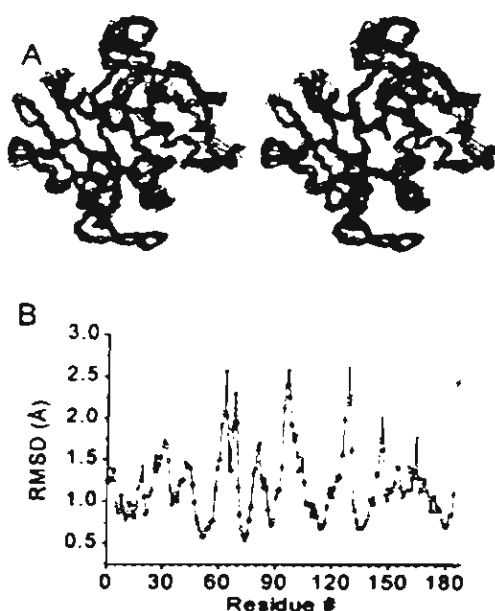


FIGURE 3 Stereo views showing the backbone α -carbon of the hDHFR (A) and its corresponding RMSD (B). The structures were taken every 10 ps from the MD trajectory. The figure was prepared with the program MOLMOL [26].

backbone fluctuation ($\text{RMSD} > 2.0 \text{ \AA}$) encompass residues Glu62-Asn64, Asp94-Leu97 and His 127-Gly129.

The 25 MD sampling configurations were evaluated. The calculation of mainchain torsion angles of the apo-hDHFR showed that no geometrically distortion of the backbone polypeptide was found, except for Ser 118. The backbone dihedral angles, ϕ and ψ , of the apo-enzyme fall within the structurally favorable regions of Ramachandran diagram. This is an indication of a good quality of regular protein structures.

The elements of the protein secondary structure were computed using the DSSP algorithm [30]. The tertiary folding and secondary structure elements of the solvated apo-hDHFR are mostly comparable to that of the hDHFR-folate complex in the crystallographic state. Significant discrepancy between the X-ray diffraction and the MD calculations was found for the two secondary structures as follows: the β -strand (Phe88-Ser90) and the short α -helical structure (Pro61-Arg65). From the 25 configurations sampling from the MD trajectory, the β -strand changed to a coil structure, while the disruption of the α -helix was observed with 65% occurrence. Among the known 3D structures of DHFR family including from *L. casei*, *E. coli*, chicken liver, *P. carinii* and *L. major*, the strand is one of the seven conserved β -structures while the helix is not. Interestingly, these two regions are short in numbers of residue and close in contact with the bulk solvent. Perhaps these two segments involve in the early event of denaturation of the mammalian apo-enzyme.

3.2.3. Mobility from RMSD and the Solvent Accessibility

In this study, the RMSD was used to indicate the local movement along the mainchain polypeptide relative to the molecular tumbling rate in solution. The RMSD values for the backbone atoms of the 25 sampling coordinates were thus characterized the amino acid residues into three categories: immobilized to slightly mobile residue (for residues $\text{RMSD} < 1.0 \text{ \AA}$); fairly mobile residue ($1.0 \text{ \AA} \leq \text{RMSD} \leq 2.0 \text{ \AA}$); and very mobile residue ($\text{RMSD} > 2.0 \text{ \AA}$). From the simulation, each category was composed of 36, 56 and 8% amino acids from a total of 186 residues, respectively.

It should be noted that the highly mobile residue has the greatest potentials of the conformational perturbation in the absence of the cofactor of the enzyme. Residues of the third category are, therefore, of interest because they may involve either the protein-ligand contact or the denaturing event or both. It should, however, be careful to not over-interpret the results of

hundreds picosecond timescale. In addition, the residue experiencing rapid conformational movement is typically located at the protein surface. As can be seen (Fig. 3), this category composes of residues present in the helical structures (Glu62-Asn64 and Asp94-Leu97) and in random coil segments (Lys68, His127-Gly129, Thr146). Most of these residues are solvent exposed.

It is obvious that a hydrophilic residue that is exposed to the protein surface has more favor of the conformational mobility, thus raising the RMSD value. To analyze this more quantitatively, the solvent accessible surface area (SASA) was computed using an algorithm developed by Connolly [31] with the probe radius of $1.4 \text{ \AA}$. The total SASA values of all 25 MD sampling structures were approximately 10% higher than that of the crystal structure. Here a term called "the solvent accessibility (SA)" was defined as the calculated SASA divided by a total surface area of a residue being considered. A higher value of SA of a residue, the more chance of residue interacts to the solvent. For instance, if the SA value of a residue is 0 (dimensionless), this residue is completely buried. It is unlikely that the solvent molecule can contact to the residue. Figure 4A shows a plot of the average SA per residue obtained from the accessibility calculation of the 25 MD structures. Distribution of the SASA values (Fig. 4A) of all residues reveals the protein is compact. Most residues are embedded in the protein. About 50% of residues are below 0.2 of the SA value (Fig. 4B). However, Fig. 4C is a plot of the difference of the solvent accessibility (ΔSA) between the average 25 MD structures and the X-ray structure along the entire sequence. Apparently, most residues of the solvated apo-enzyme have positive ΔSA , suggesting a more contact to the environment with respect to those of the crystal structure.

3.2.4. Comparison to the X-ray Structure

We have so far discussed the comparisons of the tertiary and secondary structures and the accessibility between the free enzyme in the aqueous environment and the folate bound enzyme in the crystal structure. In order to determine the magnitude of the MD structures drifted from the crystal coordinates, RMSD calculation was performed using the reference X-ray structure and the 25 sampling structures of the MD simulation. This comparison may indicate conformational differences between the bound (hDHFR-folate) and free (apo-hDHFR) enzyme. The global RMSD per residue and the superposition structures between the X-ray hDHFR-folate and the 25 structures apo-hDHFR are shown in Fig. 5. The RMSD values suggest a significant

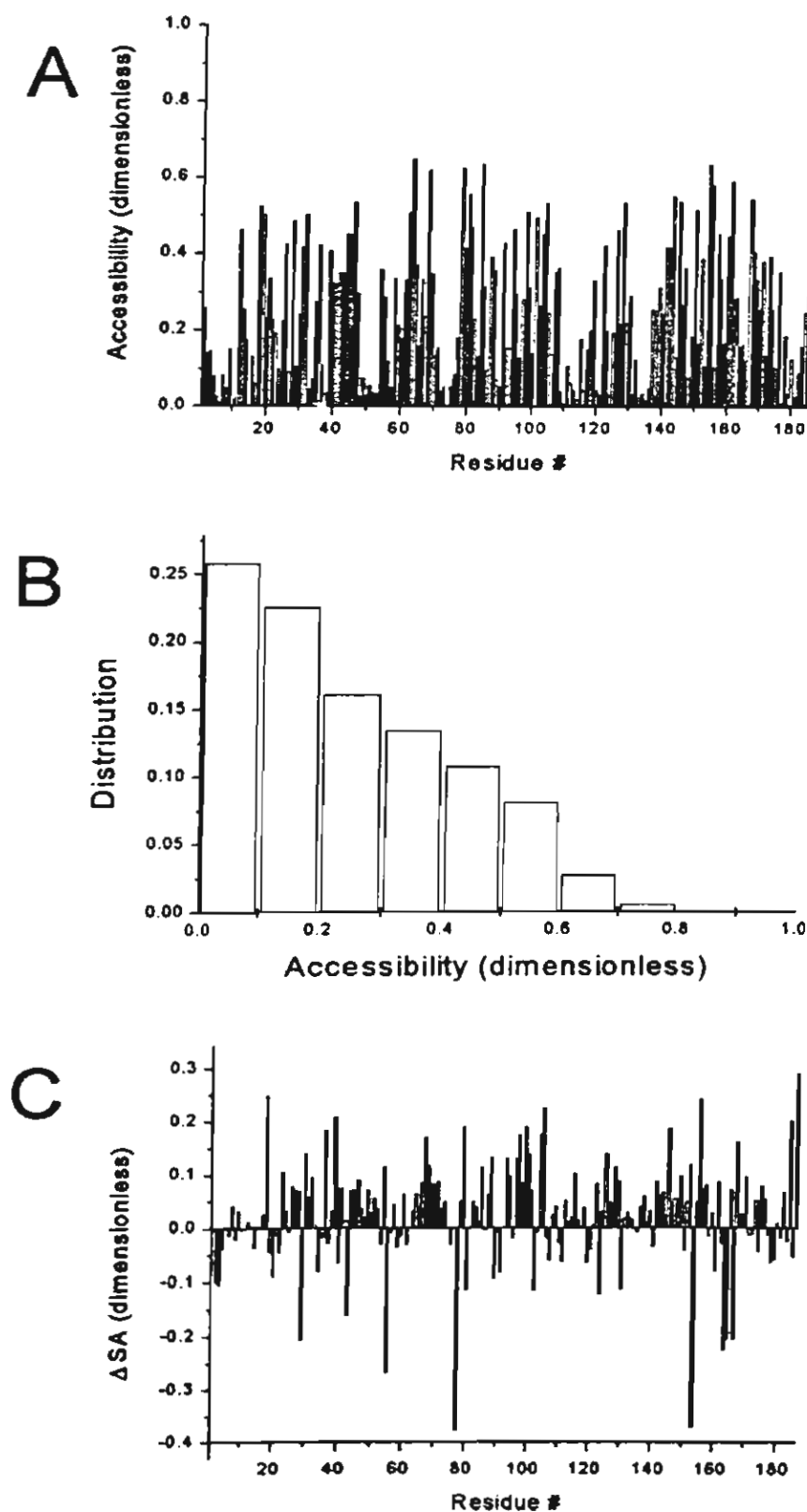


FIGURE 4 The solvent accessibility per residue (A), the residue, distribution of the accessibility (B), and the accessibility difference between the MD and the X-ray structures (C). The solvent accessibility was calculated with the program MOLMOL.

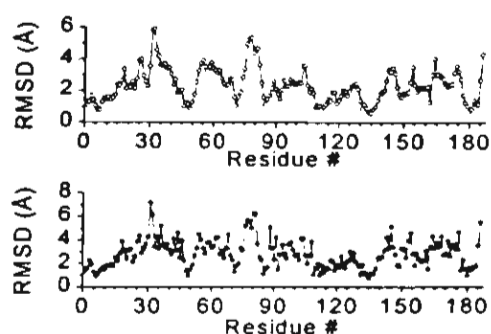


FIGURE 5 RMSD per residue of the 25 MD structures with respect to the X-ray structure: for the backbone atoms (A) and for the heavy atoms (B).

difference between the two structures (2.64 Å for the backbone atoms and 3.20 Å for the heavy atoms). Despite the similar folding patterns between the X-ray and the average structure, the RMS differences of some residues are sizeable to reflect the conformational differences. In this study, the backbone RMSDs of residues Pro26, Phe31-Phe34, Ser76-Glu81, and Gly164 were in a range of 4–6 Å. Interestingly, residues Phe31-Phe34 are located within 5 Å accounting from the folate coordinates.

In the unbound state, the residues surrounding the binding site may possess a more rapid conformation movement with respect to that in the complex form. This results in an increase of the mainchain mobility. From an observation of six isomorphous crystallographic structures of *E. coli* DHFRs, the loop and domain motions were influenced by the folate and NADPH [32]. Such conformational rearrangements occur when the enzyme releases the folate product (<http://chem-faculty.ucsd.edu/kraut/dhfr.html>). From the X-ray structure of the folate-DHFR complex, 24 residues that are located within 5.0 Å accounting from the atomic coordinates of the folate were identified. These residues are summarized with the RMSD in Table III. Four out of the 24 ligand-encountered residues exhibiting the RMSD over 4.0 Å were located in the α -helix conformation. However, we remark on the possibility that there might be a considerable difference between both the conformations for the RMSD in a range of 3–4 Å.

Although the RMSD and SA values seem arbitrarily to relate with the experimental data, our data are consistent with the previous proposal from the crystallographic study of *E. coli* DHFR (ecDHFR). Structural invariance was commonly seen in the vertebrate DHFRs, while it is not for the case of bacterial DHFRs [32]. In ecDHFR, the loop and domain movements encompass Met16-Asn23 and His45-Trp47, respectively. The alignment residues Lys18-Pro25 and Lys55-Trp57 of the hDHFR that are homologous to these two segments

TABLE III Comparisons of RMSD and the solvent accessibility between the X-ray and the average MD structures. Residues within 5 Å from the folate are reported

Residue	RMSD (Å)	
	Backbone	Heavy
Ile7	0.77	1.14
Val8	1.29	1.44
Ala9	1.52	1.65
Leu22	2.57	3.24
Arg28	2.56	3.26
Glu30	3.55	4.26
Phe31	5.85	7.19
Arg32	5.9	6.14
Tyr33	4.58	3.45
Phe34	4.2	4.27
Gln35	3.65	3.26
Thr38	3.43	3.35
Thr56	3.53	3.3
Ser59	3.79	3.83
Ile60	3.42	3.37
Pro61	3.28	3.46
Asn64	2.78	2.81
Leu67	2.38	3.31
Lys68	2.78	4.49
Arg70	1.59	2.33
Val115	1.48	1.65
Gly116	1.35	1.54
Val121	1.67	1.7
Thr136	0.82	1.18

were, however, slightly mobile (RMSDs within the ensemble were in a range of 0.7–1.4 Å). The RMSD values for the heavy atoms of this segment scattered between 2.2–3.9 Å. The SASA values calculated from the crystal structure of the ternary complex of ecDHFR were about twofold greater than those of human DHFR apo-enzyme. However, there is still inconclusive to explain why the mobility of the mammalian and the bacterial DHFRs are different.

3.2.5. Deviation of Aromatic Rings of Phenylalanine Side-chain

The dynamics information of the phenylalanine sidechain were extracted from the MD trajectory. As depicted in Fig. 6, we monitored the dynamics profile of phenylalanine rings (the χ_2 torsion angles of C α -C β -C γ -C δ) along the 250 ps trajectory. The residues Phe58, Phe88, Phe134, Phe142, Phe147, Phe148 and Phe179 were also included in the analysis. The profiles indicated that the χ_2 values of most phenylalanine residues including Phe34, Phe58, Phe134, Phe142, Phe147 and Phe148 were between 80–100° or equivalent to –100–80° (because of the presence of the symmetry of the aromatic ring), with an average fluctuation of 20°, implying the similar orientation of the aromatic ring with respect to their C α -C β -bond. A slight difference of the phenyl ring orientation was found for Phe88 and Phe179, of which the χ_2 values were around –120 and +120°, respectively. The sidechain of Phe31 exhibits the most

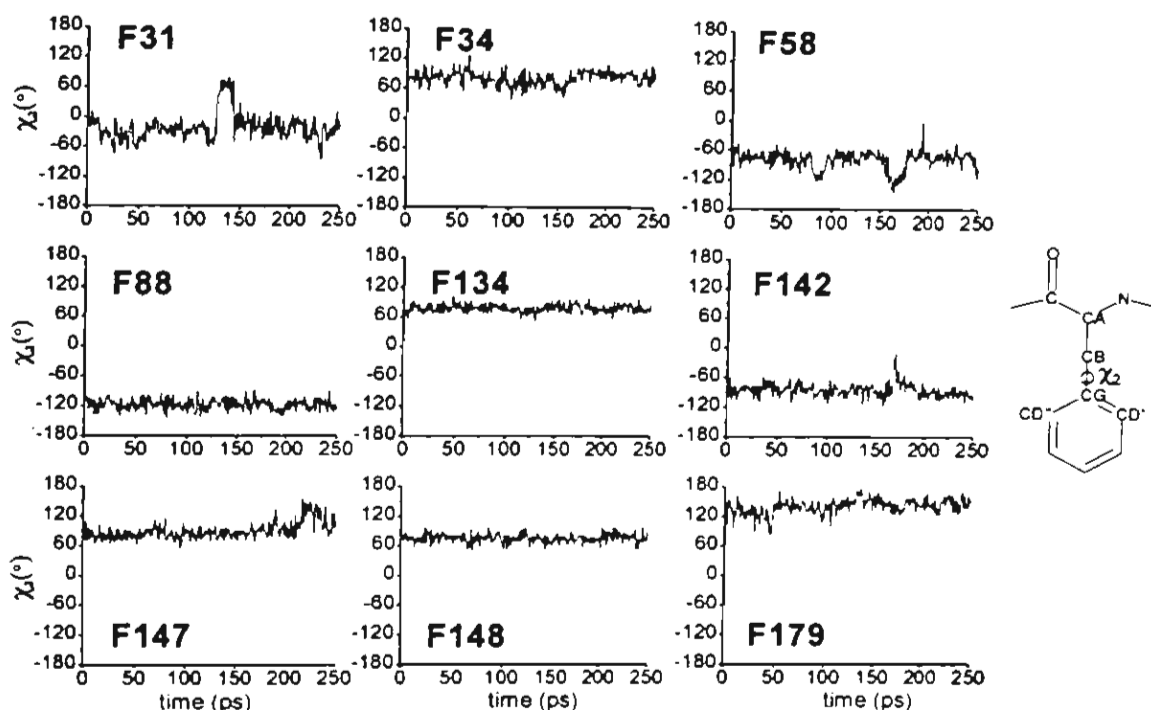


FIGURE 6 Dihedral angle χ_2 fluctuation of the sidechain of nine phenylalanines versus the period of the MD simulation.

different orientation as well as the fluctuation relative to the others.

According to the results from site-directed mutagenesis experiments, Phe31 and Phe34 of the wild-type protein, were previously reported to play a critical role for the binding efficiency of the ligand in the enzyme catalysis [33,34]. In the binary and ternary complexes of hDHFR, the phenyl ring of Phe34 forms a π - π stacking with the pteridine ring of the folate, but such interaction was unlikely in the case of Phe31 (Fig. 7). In the deficiency condition of the folate, the movement of aromatic ring of the side chain of Phe31 and Phe34 might relate to the conformational change of the apo-enzyme. From Fig. 6A, the flipping of the Phe31 ring was clearly observed and hence the most sensitive residue with respect to the others. The χ_2 value changed from -30° to $+60^\circ$ and stabilized for a period of 30 ps and returned back again to the original equilibrium value. In contrast, χ_2 of Phe34 suggested only a minor movement of the side chain. The differences in orientation and the movement of the side chain between Phe31 and Phe34 in the apo-enzyme were likely affected by the unoccupied space in the binding pocket. It appears that there is a strong relationship between the mobility of the sidechain and the surface area of molecular contact. Phe34 is considerably buried in the protein with a low solvent accessibility (0.07) whereas the SA of Phe31 was 0.42 (Fig. 4A).

3.2.6. Comparison of Interamolecular Hydrogen Bond

It has been known that the effect of the interresidue hydrogen bond network plays an important role in folding and stabilizing protein conformations. Here,

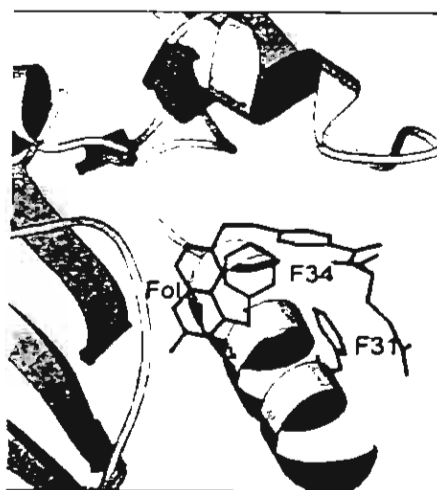


FIGURE 7 Schematic representation showing an orientation of Phe31, Phe34 and the folate in the X-ray structure. For clarification, some segments of the protein were not shown

- [30] Kabsh, W. and Sander, C. (1983) "Dictionary of protein secondary structure; pattern recognition of hydrogen-bonded and geometrical features", *Biopolymers* 22, 2577.
- [31] Connolly, M.L. (1983) "Solvent-accessible surfaces of proteins and nucleic acids", *Science* 221, 709.
- [32] Sawaya, M.R. and Kraut, J. (1997) "Loop and sundomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: crystallographic evidence", *Biochemistry* 36, 586.
- [33] Chunduru, S.K., Cody, V., Luft, J.R., Pangborn, W., Appleman, J.R. and Blakley, R.L. (1994) "Methotrexate-resistant variants of human dihydrofolate reductase: effects of Phe31 substitutions", *J. Biol. Chem.* 269, 9547.
- [34] Nakano, T., Spencer, H.T., Appleman, J.R. and Blakley, R.L. (1994) "Critical role of phenylalanine 34 of human dihydrofolate reductase in substrate and inhibitor binding and in catalysis", *Biochemistry* 33, 9945.

significantly different. As can be seen from Fig. 8, many vanishing NH protons, particular in the conserved α -helix encompassing residue 32–39, are rather not used to maintain regular secondary structures. This might suggest the apo-enzyme is less compact and eventually unable to sustain the conformation without the ligand molecule. Perhaps, this causes the enzyme to be unstable without the ligand.

4. CONCLUSIONS

The present study demonstrates the dynamics structures of the apo-enzyme in the absence of ligand. The MD simulation provides a tendency of conformational changes from the crystal structure of the enzyme-folate complex to the apo-enzyme in aqueous solution. The changes including the secondary structure, the solvent accessibility, the mainchain mobility and the hydrogen bond could play a critical role that relate to the stability of the apo-enzyme.

Acknowledgements

Part of this work was supported by the Thailand Research Fund Rachadapisek Sompoj Endowment, and National Science and Technology Development Agency. Thanks to Dr Banci and Dr Bates for critical review of the manuscript

References

- [1] McCammon, J.A. and Harvey, S.C. (1987) *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge).
- [2] Brooks, C.L., K., M. and Pettitt, B.M. (1988) *Protein: A Theoretical Perspective of Dynamic Structure and Thermodynamics* (Wiley, Canada).
- [3] Schwentzer, B.L., Dicker, A.P. and B., J.R. (1990) "Dihydrofolate reductase as a therapeutic target", *FASEB J.* 4, 2441.
- [4] Knighton, D.R., Kan, C.-C., Howland, E., Janson, C.A., Hostomska, Z., Welsh, K.M. and Matthews, D.A. (1994) "Structure and kinetic channelling in bifunctional dihydrofolate reductase-thymidylate synthase", *Nat. Struct. Biol.* 1, 186.
- [5] Bolin, J.T., F.D., J., Matthews, D.A., Hamlin, R.C. and Kraut, J. (1982) "Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. I. General features and binding of methotrexate", *J. Biol. Chem.* 257, 13650.
- [6] Birdsall, B., F., J., Tendler, S.J., Hammond, S.J. and Roberts, G.C. (1989) "Dihydrofolate reductase: multiple conformations and alternative modes of substrate binding", *Biochemistry* 28, 2297.
- [7] Bystroff, C., O., S.J. and Kraut, J. (1990) "Crystal structures of *Escherichia coli* dihydrofolate reductase: the NADP⁺ holo-enzyme and the folate.NADP⁺ ternary complex. Substrate binding and a model for the transition state", *Biochemistry* 29, 3293.
- [8] Bystroff, C. and J., K. (1991) "Crystal structure of unliganded *Escherichia coli* dihydrofolate reductase. Ligand-induced conformational changes and cooperativity in binding", *Biochemistry* 30, 2227.
- [9] Champness, J.N., A., A., Ballantine, S.P., Bryant, P.K., Delves, C.J. and Stammers, D.K. (1994) "The structure of *Pneumocystis carinii* dihydrofolate reductase to 1.9 Å resolution", *Structure* 2, 915.
- [10] Volz, K.W., M., D.A., Alden, R.A., Freer, S.T., Hansch, C., Kaufman, B.T. and Kraut, J. (1982) "Crystal structure of avian dihydrofolate reductase containing phenyltriazine and NADPH", *J. Biol. Chem.* 257, 2528.
- [11] Matthews, D.A., Bolin, J.T., Burridge, J.M., Filman, D.J., Volz, K.W. and Kraut, J. (1985) "Dihydrofolate reductase: the stereochemistry of inhibitor selectivity", *J. Biol. Chem.* 260, 392.
- [12] McTigue, M.A., D., J.F., Kaufman, B.T. and Kraut, J. (1992) "Crystal structure of chicken liver dihydrofolate reductase complexed with NADP⁺ and biopterin", *Biochemistry* 31, 7264.
- [13] McTigue, M.A., D., J.F., Kaufman, B.T. and Kraut, J. (1993) "Crystal structures of chicken liver dihydrofolate reductase: binary thioNADP⁺ and ternary thioNADP⁺+biopterin complexes", *Biochemistry* 32, 6855.
- [14] Oefner, C., D., A.A. and Winkler, F.K. (1988) "Crystal structure of human dihydrofolate reductase complexed with folate", *Eur. J. Biochem.* 174, 377.
- [15] Davies, J.F., Delcamp, T.J., Prendergast, N.J., A., V.A., Freisheim, J.H. and Kraut, J. (1990) "Crystal structure of recombinant human dihydrofolate reductase complexed with folate and 5-deazaflavone", *Biochemistry* 29, 9467.
- [16] Cody, V., Galitsky, N., Luft, J.R., Pangborn, W., Blakley, R.L. and Gangjee, A. (1981) "Comparison of ternary crystal complexes of F31 variants of human dihydrofolate reductase with NADPH and a classical antitumor furo-pyrimidine", *Anti-Cancer Drug Des.* 13, 307.
- [17] Matthews, H.R., Matthews, K.S. and Opella, S.J. (1977) "Selectively deuterated amino acid analogues: synthesis, incorporation into proteins and NMR properties", *Biochim. Biophys. Acta* 497, 1.
- [18] Stockman, B.J., N., N., Wagner, G., Delcamp, T.J., De Yarman, M.T. and Freisheim, J.H. (1992) "Sequence-specific ¹H and ¹⁵N resonance assignments for human dihydrofolate reductase in solution", *Biochemistry* 31, 218.
- [19] Johnson, J.M., Meiering, E.M., Wright, J.E., Pardo, J., Rosowsky, A. and Wagner, G. (1997) "NMR solution structure of the antitumor compound PT523 and NADPH in the ternary complex with human dihydrofolate reductase", *Biochemistry* 36, 4399.
- [20] Cornell, W.D., Cieplak, P., Bayly, C.J., Gould, I.R., Merz, Jr., K.M., Ferguson, Jr., D.M., Spellmeyer, Jr., D.C., Fox, Jr., T., Caldwell, Jr., J.W. and Kollman, Jr., P.A. (1995) "A second generation force field for the simulation of proteins, nucleic acid, and organic molecules", *J. Am. Chem. Soc.* 117, 5179.
- [21] Jorgensen, W.L., Chandrasekhar, J., Madura, J., Impey, R.W. and Klein, M.L. (1983) "Comparison of simple potential functions for simulating liquid water", *J. Chem. Phys.* 79, 926.
- [22] Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R. (1984) "Molecular dynamics with coupling to an external bath", *J. Chem. Phys.* 81, 3684.
- [23] Case, D.A., et al. (1997) AMBER 5 (University of California, San Francisco).
- [24] Guex, N. and P., M. (1997) "SWISS-MODEL, and the Swiss-PdbViewer: an environment for comparative protein modelling", *Electrophoresis* 18, 2714.
- [25] Sayle, R. (1994) Rasmol 2.6 edit. (Biomolecular Structure Department, Glaxo Research and Development, Greenford, Middlesex, UK).
- [26] Koradi, R., Billeter, M. and Wuthrich, K. (1996) "MOLMOL: a program for display and analysis of macromolecular structures", *J. Mol. Graph.* 14, 51.
- [27] Kraulis, P.J. (1991) "MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures", *J. Appl. Crystallog.* 24, 946.
- [28] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) "PROCHECK: a program to check the stereochemical quality of protein structures", *J. Appl. Crystallog.* 26, 283.
- [29] McCammon, A.J., H.S., Ed. (1987) *Dynamics of Proteins and Nucleic Acids* (Cambridge university press, Cambridge UK).

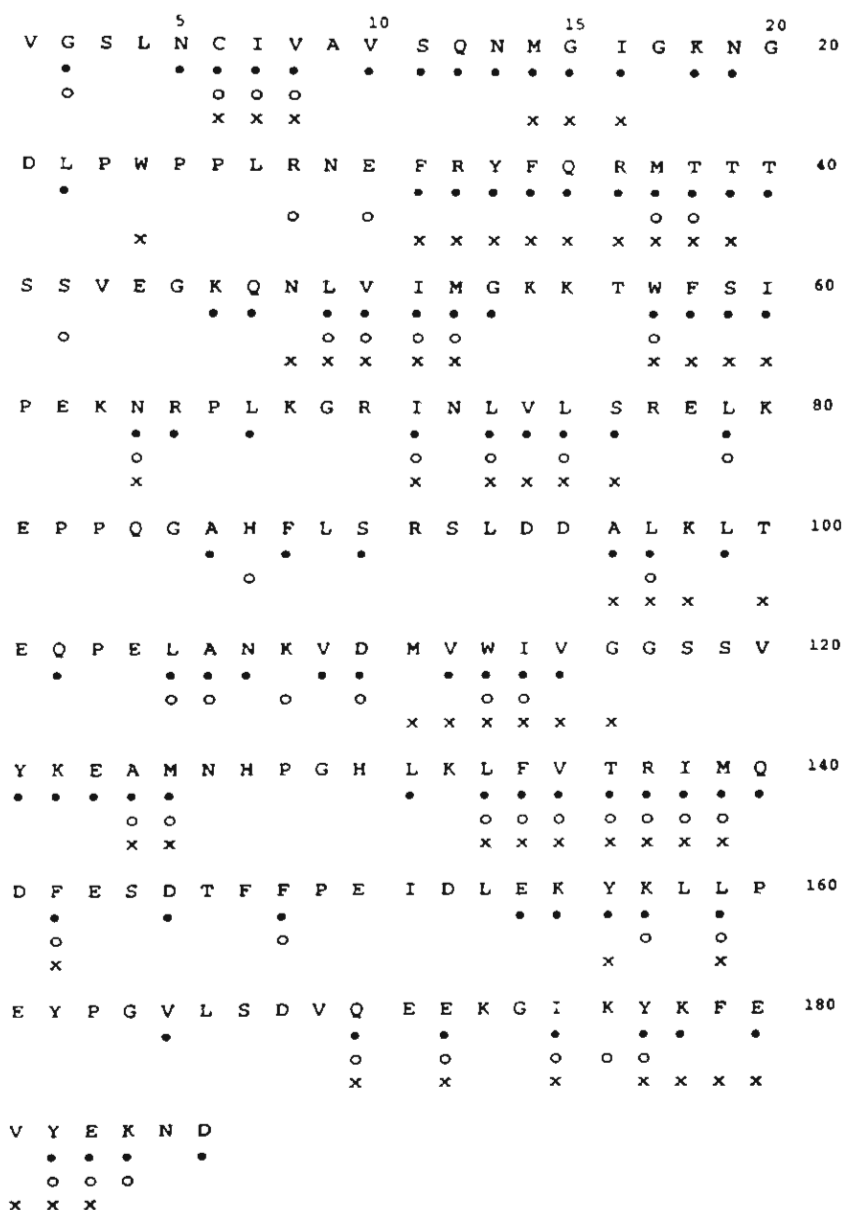


FIGURE 8 Comparison of the hydrogen bond involving the NH protons: (•) predicted from the X-ray structure, (○) calculated from the MD structures and (x) taken from the NMR data [18].

additional information from the dynamics trajectory was further investigated by identification of possible hydrogen bond involving backbone hydrogen atoms of the apo-hDHFR. A principle criteria used to identify a possible formation of the hydrogen bond are: (1) a distance between a HN atom and the hydrogen bond acceptor should be below 2.5 Å and (2) the angle of N-HN the acceptor should not exceed 35°. About 94 out of 174 backbone amide protons (excluding 12 prolines) were found in the X-ray structure (Fig. 8). On the other hand, analysis of the 25 MD structures shows that 48 NH protons

appearing consistently more than 80% of the ensemble are capable of forming the intramolecular hydrogen bond. It should be noted that the resonance assignment from the homonuclear and heteronuclear NMR spectra of the hDHFR complexed with methotrexate (MTX) identified 63 NH protons corresponding to the slow exchangeable proton [18]. Almost half of backbone hydrogen bond disappears from the simulation. Indeed, the number of the hydrogen bond involved with the NH protons between the MD simulation of the apo-enzyme and the NMR data of the hDHFR-MTX complex is