

EXERCISE III.4**MEDICINAL CHEMISTRY AND MOLECULAR MODELING: AN INTEGRATION FOR THE TEACHING OF DRUG STRUCTURE–ACTIVITY RELATIONSHIP AND THE MOLECULAR BASIS OF DRUG ACTION****Ivone Carvalho***, **Áurea D. L. Borges**, and **Lílian S. C. Bernardes***Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Av. do Café s/n, Monte Alegre, Ribeirão Preto – SP, 14040-903, Brazil*E-mail: *carronal@usp.br*

Abstract: Molecular modeling is described as a tool for understanding fundamental concepts of drug structure–activity relationships in a medicinal chemistry course. The relevant molecular features of antimetabolite drugs are investigated by 3D visualization, their physical properties measured, and the molecular interaction pattern on target macromolecules illustrated by antineoplastic drugs. This approach became a high-quality computing and graphic tool to teach important aspects of biological molecules and drugs and the correlation of their structures and pharmacological actions. The students improve their perception and understanding of the molecular recognition process and may predict molecular properties by handling computer graphics and databases.

Keywords: molecular modeling; medicinal chemistry; drugs; molecular properties; molecular recognition.

Modern molecular modeling techniques are a remarkable tool in the search for potential novel therapeutic agents, by helping us to understand and predict the behavior of molecular systems. The powerful modeling components, including molecular graphics, computational chemistry, molecular database, and statistical modeling (QSAR) have assumed an important role in the development and optimization of leading compounds. Moreover, current information on the protein's 3D structure and functions opens up the possibility of understanding the relevant molecular interactions between a ligand and a target

macromolecule. As a consequence, a comprehensive study of drug structure–activity relationships (SARs) can be developed and provide the proper identification of a 3D pharmacophore model for a rational drug design. Although structure-based design is now a medicinal chemistry routine approach, there are still difficulties in teaching fundamental concepts to undergraduate pharmacy students, such as those related to the molecular recognition process [1,2].

An active-learning strategy in medicinal chemistry involves the incorporation of molecular modeling techniques to assist third-year undergraduate students in the understanding of structure–activity principles. This paper focuses on the use of computational chemistry and the protein data bank (PDB), accessed from the Web site <http://www.rcsb.org/pdb>, to understand and predict the chemical and molecular basis involved in the drug–receptor interactions. A comprehensive study of SARs comprises three approaches. The first one involves comparative analysis of antimetabolite drugs and the corresponding metabolites (named substrates), by representing, visualizing, and superimposing their 3D structures, obtained by minimization processes and molecular alignment techniques. Numerical properties of these molecules are then calculated, the most common being molecular energies and physical constants as partition coefficients, dipolar moment, molecular volume, and interatomic distance. Finally, particular structural features between substrate and antimetabolite are explored by assessing the electrostatic and geometric patterns required for chemical interaction in the active site of the target molecule, obtained from PDB. Table 1 lists therapeutic targets of interest, describing the enzymes and their corresponding substrates, some PDB files and antimetabolites currently used as antineoplastic, anti-HIV, antibiotic, antihypertensive, anti-inflammatory, cholinergic, and hypolipemic drugs [3].

Table 1 Antimetabolites, substrates, and their corresponding therapeutic targets and some PDB files of interest.

Group	Therapeutic class	Enzyme	Substrate	PDB file	Drug/ antimetabolite
I	Antineoplastic	Thymidylate synthase	deoxyuridylate monophosphate	1tls, 1tsn, 1hzw, 1kzi, 1hvy, 1bq1, 1cif, 1tsw	5-fluorouracil, trifluorothymidine
II	Anti-HIV	Reverse transcriptase	Deoxy-thymidylate monophosphate	1cot, 1cou, 1dtq, 1fkp, 1hmv, 1hvu	zidovudine, lamivudine, stavudine, didanosine
III	Anti-HIV	HIV Protease	Poliprotein	1hvp, 1kzk, 2aid, 2bpv, 2bpw, 7upj	saquinovir, indinavir, ritonavir
IV	Antibiotic	Transpeptidase and Carboxypeptidase	Acil-D-Ala-D-Ala	1cef, 1ceg, 1es2, 1es3	penicillin G, oxacillin, ampicillin, carbenicillin, cephalixin
V	Hipolipemic	HMGO-CoA reductase	HMGO-CoA	1dqa, 1hw8, 1hw9, 1hwj, 1qax, 1qay	lovastatin, simvastatin, fluvastatin, atorvastidina
VI	Antihypertensive	Dopa decarboxilase	L-Dihydroxy-phenylalanine	1js3, 1js6, 5pah, 4pah, 1phh, 1d7l	methyl dopa, methyl dopamine, methyl norepinephrine
VII	Antineoplastic	Dihydrofolate reductase	Dihydrofolic acid	1rb3, 1rg7, 1rh3, 1ra3, 7dfr, 3dfr	methotrexate, piritrexim, trimetrexate
VIII	Antiinflammatory	Cyclooxygenase	Arachidonic acid	1dcx, 1cqe, 1cvu, 1ddx, 1cx2, 1pgf	ibuprofen, indomethacin, naproxen, acetylsalicylic acid
IX	Cholinergic	Acetylcholinesterase	Acetylcholine	2ace, 1gqr, 2ack, 1dx4, 2clj, 1qti	neostigmine, tacrine, pyridostigmine, echothiophate, demecarium, pralidoxime

3D structure comparisons and overlays

Using the Molecular Modeling Pro [4] program, it is possible to construct 3D interactive drug pictures, optimized by reducing the energy of the molecules in some systematic way until a minimum energy conformer is found. Minimization processes can correct unfavorable bond lengths, bond angles, torsion angles, and nonbonded interactions in a starting structure, creating a more stable conformation. Mathematical models that perform geometry optimization are divided into classical, mechanical, and quantum mechanical approaches. In computational chemical simulations, the simplified description is a calculated potential energy surface, which applies classical mechanics equations to molecular nuclei, without considering electrons. A set of equations and parameters is called force field, and most molecular modeling programs can choose among several force fields, such as MM2 and AMBER [5]. The energy of any atomic arrangement can be stepwise calculated, by assessing how the energy of the system varies as the position of the atoms change. Molecular Modeling Pro can generate and examine automatically many molecular conformations, and their corresponding inter-conversion energy barriers graphically plotted. At the completion of the conformational analysis, the molecule is placed in its low energy conformation. Both procedures can be interactively performed to optimize the geometry of the molecules that now may be compared structurally by overlaying appropriate atoms or functional groups, previously aligned in the atomic coordinates. Rotation and different representation forms (by charge or lipophilicity) of the model allow a detailed investigation of the conformational and electronic properties of two structures, which could be the substrate and antimetabolite (inhibitor).

Two classes of antineoplastic drugs are chosen to exemplify the structure visualization and superposition processes, which involve compounds that act on target enzymes as thymidylate synthase (TS, Fig. 1) and dihydrofolate reductase (DHFR, Fig. 2). The purpose

of this exercise is to provide a geometrical and chemical overview of the substrate and inhibitor of great structural similarity. At the same time, it reinforces the importance of chemical transformations of a key substrate in a biochemical pathway when developing antimetabolite drugs.

Molecular dimensions and properties

Varying the physical and chemical properties of metabolite has often been used in drug design to produce a structure analog such as antimetabolite (metabolite antagonism), required for inhibition of cellular growth and response process. Bioisosterism strategies are frequently used to convert a key substrate into an inhibitor, allowing extra binding interaction to a target enzyme. The substituent modifications can affect various parameters in a drug molecule such as the partition coefficient, electronic density, conformation, bioavailability, and its capacity to establish direct interaction in the receptor domain. Thus, both drug and substrate physical properties can be calculated and compared to give useful information on SARs. Important properties can be estimated in the Molecular Modeling Program, exemplified by partial charges, bond order, dipole moment, ionization potential, electron density distribution, solubility, and thermodynamic properties. Some of the physical properties, including volume, surface area, density, molecular length, and dipole moment, are affected by geometry, which depends on the minimization process. Additionally, appropriate orientation of the molecule in the x-y plane is required for molecular length comparisons.

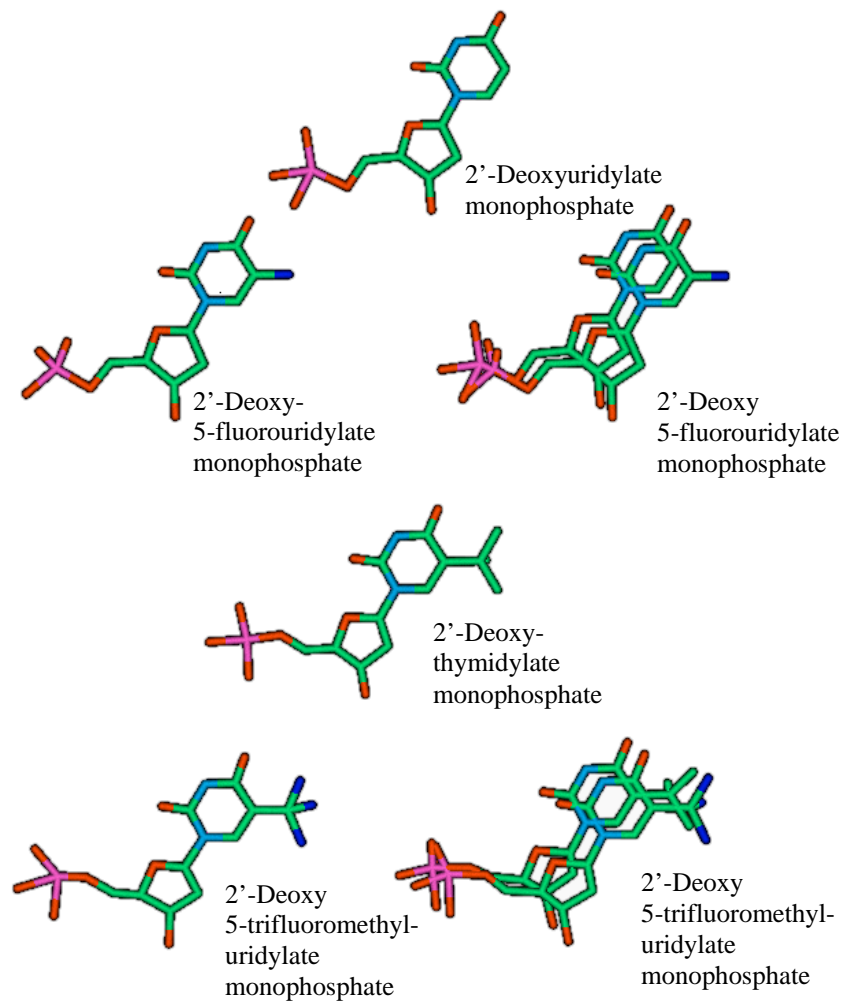


Fig. 1 Superposition and comparisons of nucleotide-derived antineoplastic drugs capable of inhibiting TS, with the corresponding substrates.

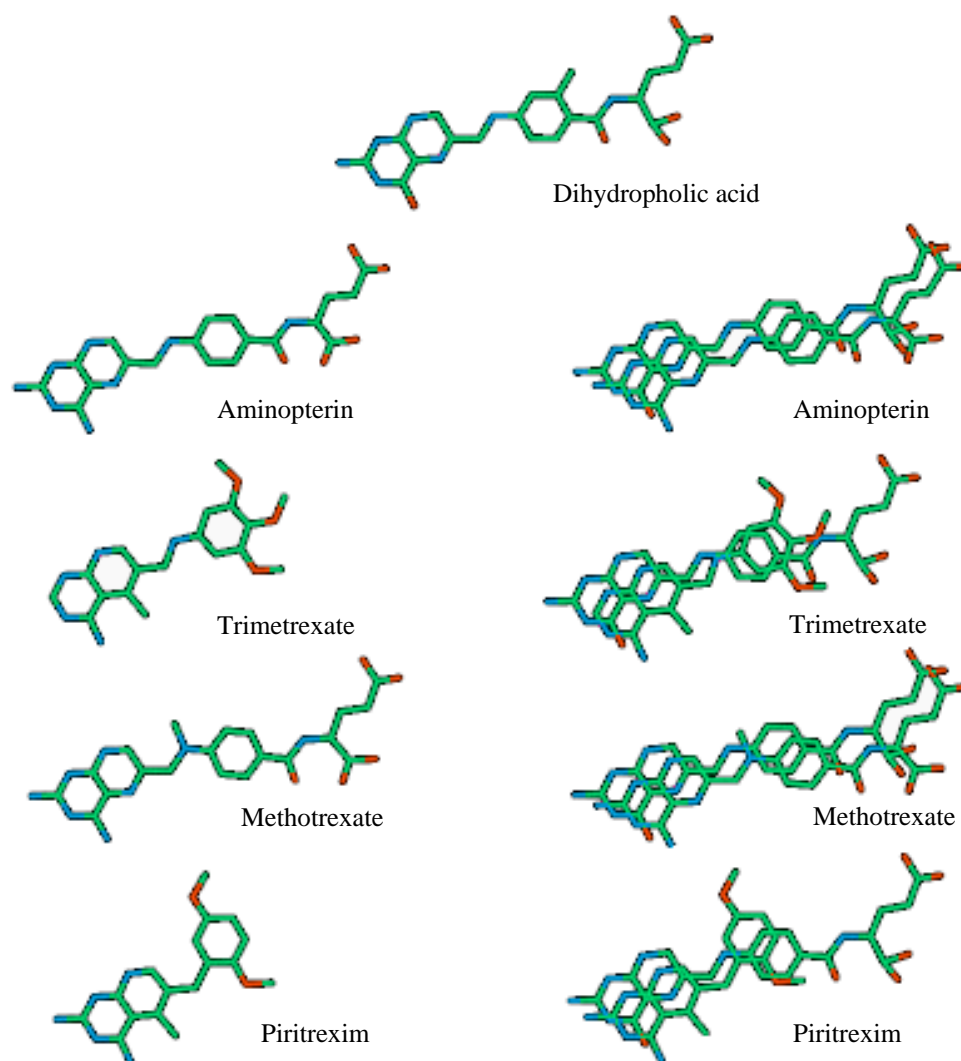


Fig. 2 Superposition and comparisons of folate cofactor-derived antineoplastic drugs capable of inhibiting DHFR, with the corresponding substrate.

Some physical similarities and differences among antineoplastic compounds (Table 1) belonging to groups I (TS) and VII (DHFR) are illustrated, respectively, in Tables 2 and 3.

Compared with substrates uridylate and thymidylate, both antimetabolites 5-fluorouracil and trifluridine show additional electronegative groups, which contribute to better nucleophilicity toward the thymidylate synthase target (Table 2). Fluorine is often considered an isostere of hydrogen even though it does not have the same valence as

hydrogen. The atom is virtually the same size as hydrogen but more electronegative and thus can be used to vary the drug electronic properties without having any steric effect. Estimates of log *P* by using the fragment method show the greater fluorinated antimetabolites lipophilic character, as compared with the corresponding substrates (Table 2). Substituting fluorine for enzymically labile hydrogen can also disrupt the catalytic reaction since C–F bonds are not easily broken [6].

Table 2 Physicochemical similarities and differences of TS antimetabolites and substrate.

Compound	Molecular volume (cm ³ /mol)	Log <i>P</i> (fragments)	H bond acceptor	H bond donor	Dipole moment (debyes)
2'-Deoxyuridylate monophosphate	140.68	-4.13	3.43	1.05	4.60
2'-Deoxy-5-fluorouridylate monophosphate	140.33	-5.23	3.43	1.05	4.43
2'-Deoxy-thymidylate monophosphate	147.58	-3.48	3.44	1.05	3.05
2'-Deoxy-5-trifluoromethyl-uridylate monophosphate (trifluoridine MP)	161.64	-3.24	3.59	1.05	4.95

In contrast to the dihydrofolate substrate, the methotrexate antagonist has an extra pteridine ring amino group, which improves the hydrogen bond interaction on the active site. The replacement of the 4-oxo group of the substrate by the amino group will not appreciably change the size of the analog, but will have a marked effect on its polarity, electronic distribution, and bonding (Table 3). However, the *N*-methyl group containing methotrexate antagonist has a different shape and increased log *P* constant and liposolubility. The methyl group that generated steric hindrance may create constraints and impose particular favorable conformations for ligand and receptor interactions. Moreover, the *N*-methyl group inductive electron-donating effect disfavors ionization and gives rise to non-ionized forms, less soluble in water [7].

Table 3 Physical similarities and differences of DHFR antimetabolites and substrate.

Name	Molecular volume (cm ³ /mol)	Log <i>P</i> (fragments)	H bond acceptor	H bond donor	Dipole moment (debye)	Polar surface area
Dihydrofolic acid	224.90	-3.60	3.52	2.12	5.24	224.86
Aminopterin	223.21	-4.00	3.60	2.12	4.06	228.81
Methotrexate	234.42	-3.32	3.35	1.87	3.61	220.02
Trimetrexate	192.97	-0.02	1.57	0.82	1.53	104.41
Piritrexim	177.25	0.90	1.76	1.07	4.72	109.17

The development of resistance during acute leukemia therapy is probably due to the loss of the methotrexate cellular transport mechanism. Thus, investigation for a more lipophilic inhibitor led to trimetrexate and piritrexim, which are independent of the cell transport mechanism. They are analogs of methotrexate, in which one or two pteridine ring nitrogen atoms are replaced by carbon and a more lipophilic group replaces the benzoylglutamic acid chain. Table 3 lists log *P* values, showing the greater trimetrexate and piritrexim liposolubility and lower polar surface area, as compared to classical antimetabolites.

Identifying enzyme 3D pharmacophores

Many drugs are effective by interacting with biological macromolecules such as enzymes, DNA, glycoproteins, or receptors. The target enzyme-substrate, -inhibitor or -cofactor (ligand) 3D complexes, can be downloaded from PDB onto a computer program and studied by molecular modeling. Ligand and target interactions may be due entirely to nonbonded forces, but occasionally a covalent interaction may be involved. Tight-binding ligands often have a high degree of target complementarity, which can be assessed and measured. A 3D *pharmacophore* specifies the group spatial relationships, corresponding to a set of features

common to active molecules, such as hydrogen bond donors and acceptors, positively or negatively charged groups, and hydrophobic groups of an appropriate size. The correlation of these structures with pharmacological action and complementary molecular interaction analyses between biological molecules and substrate/drug were possible by the use of a web accessible tool, Protein Explorer, a freeware option under PDB View Structure. The steps involved in the manipulation of Protein Explorer are described in the supplemental material.

The students are asked to render different format and color 3D enzymes from PDB and search the Display File list to access the catalytic site. Although some target enzyme active sites are not available from PDB files, it is possible to estimate how strongly a molecule will bind to a catalytic site by selecting the ligand's surface contacts favorably interacting with specific functional groups of both ligand and macromolecule (see supplemental material). TS (group I) and DHFR (group VII) are conveniently chosen enzymes to illustrate these tutorials due to their concomitant action in the cell *de novo* biosynthesis of thymidilate nucleotides.

Both enzymes have long been recognized as a drug target for inhibiting DNA synthesis in rapidly proliferating cells such as cancer cells or in bacterial or malarial infections. Traditional inhibitors clinically used as antineoplastic and antimicrobial agents, have been modeled on dUMP or the cofactor N^5,N^{10} -methylenetetrahydrofolate, and thus are structurally related to natural substrate and cofactor [3].

Thymidylate synthase

Background

Thymidylate synthase catalyzes the reductive methylation of 2'-deoxyuridine monophosphate (dUMP) to 2'-deoxythymidylate monophosphate (dTMP), using N^5,N^{10} -methylenetetrahydrofolate cofactor, which is concomitantly converted to 7,8-dihydrofolate

(DHF). Dihydrofolate reductase and serine hydroxymethyltransferase are required by the cell to transform consumed coenzyme back to the active N^5, N^{10} -methylenetetrahydrofolate form. The whole pathway is illustrated in Fig. 3.

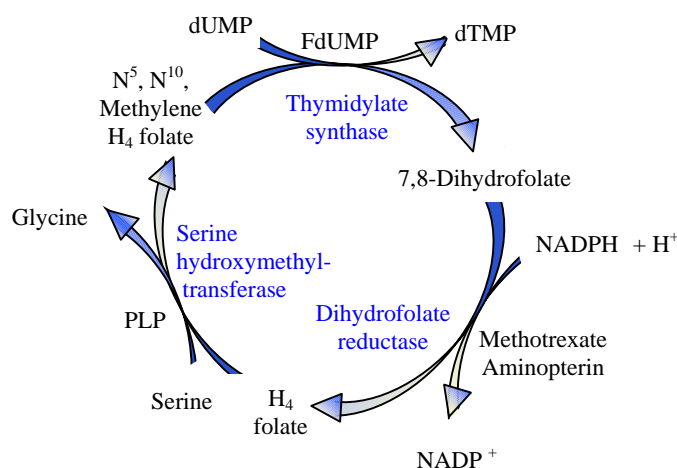
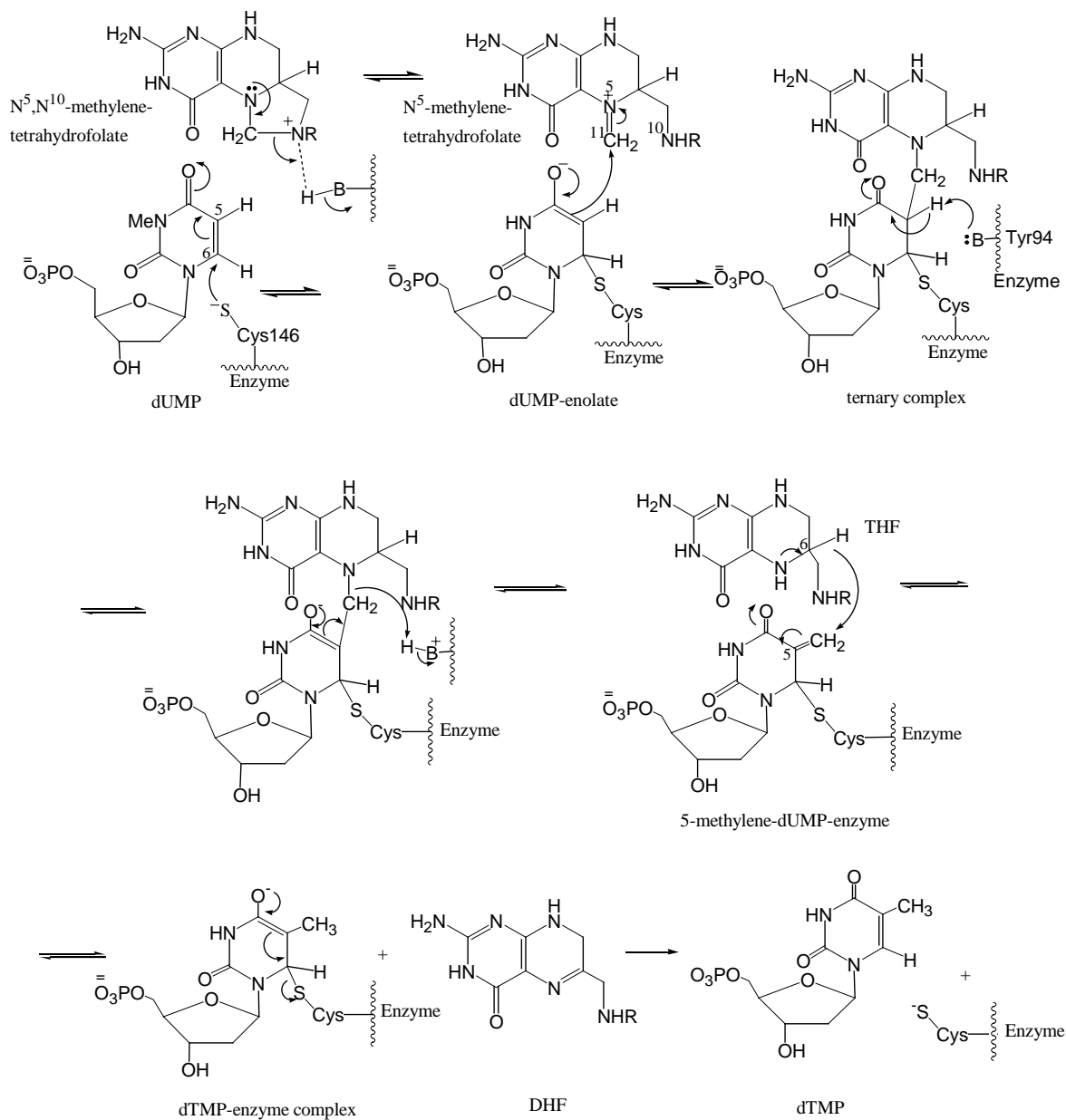


Fig. 3 Cell thymidilate biosynthetic pathway involving the enzymes TS, DHFR, and serine hydroxymethyltransferase.

Based on background information, the student draws the initial steps of the catalytic mechanism as shown in Scheme 1 [8]. By this mechanism, the reaction proceeds by the catalytic Michael addition of a cysteine residue (Cys 146) to C6 of dUMP, generating the dUMP enolate, able to attack either the cyclic or open form of methylenetetrahydrofolate. Apparently, the preferred way is the much more reactive Mannich basic form of the N^5 -iminium ion of methylenetetrahydrofolate, allowing the formation of a dUMP and cofactor containing ternary complex, covalently binding the enzyme. Deprotonation of the C5 proton of dUMP, assisted by a basic tyrosine residue (Tyr 94), is the next step originating tetrahydrofolate (THF) and a C5 exocyclic methylene group containing dUMP, still bound to the enzyme. Finally, the C5 exocyclic methylene group may be reduced by the transfer of an

hydride from THF C6 to give a dTMP-enzyme enolate and dihydrofolate (DHF), probably due to the proper co-enzyme C6 hydrogen orbital alignment with the *exo*-methylene group (Scheme 1).



Scheme 1 Proposed mechanism for the thymidylate synthase-catalyzed reaction.

Tutorial

A Protein Explorer contact analysis of the enzyme-substrate complex, PDB entry 1kzi from *E. coli*, revealed both the ligand binding place and the binding interactions between substrate and cofactor [9]. Indeed, long-range recognition factors between drug and receptor represented by hydrogen bond may be assessed and inferred in terms of electrostatic character and steric orientation. Figure 4a shows the bent conformation tetrahydrofolic acid with its PABA residue above the pteridine ring, allowing tight enzyme and nucleotide binding. In this thymidylate synthase-catalyzed reaction, it is possible to visualize the good proximity of cysteine residue (Cys146) sulphhydryl group to the C6 of dUMP pyrimidine moiety (1.94 Å), and the close contact of the coenzyme, that favors the attachment of the dUMP enolate at C5. Hydrogen bonds between Cys 146, Arg 166, and Tyr 94 (Fig. 4b) are also included in the TS structural domain and probably involved in the bond instability of the enzyme covalent adduct for further cleavage. The students are reminded to look at the potential base Tyr 94 and surrounding water that seem involved in the abstraction of the pyrimidine ring C5 proton of the ternary complex, preceding DHF elimination. In addition, the optimal positioning of C6 THF cofactor and the C5 dUMP exocyclic methylene group for hydride transfer (2.5 Å) and the active site shielding from bulk solvent are influenced by Trp80 and Leu143 residues. Although there are other interactions at the catalytic site, the hydrogen bonds between dUMP and the highly conserved amino acid Asn 177, which encode dUMP specificity over other nucleotides should be pointed out.

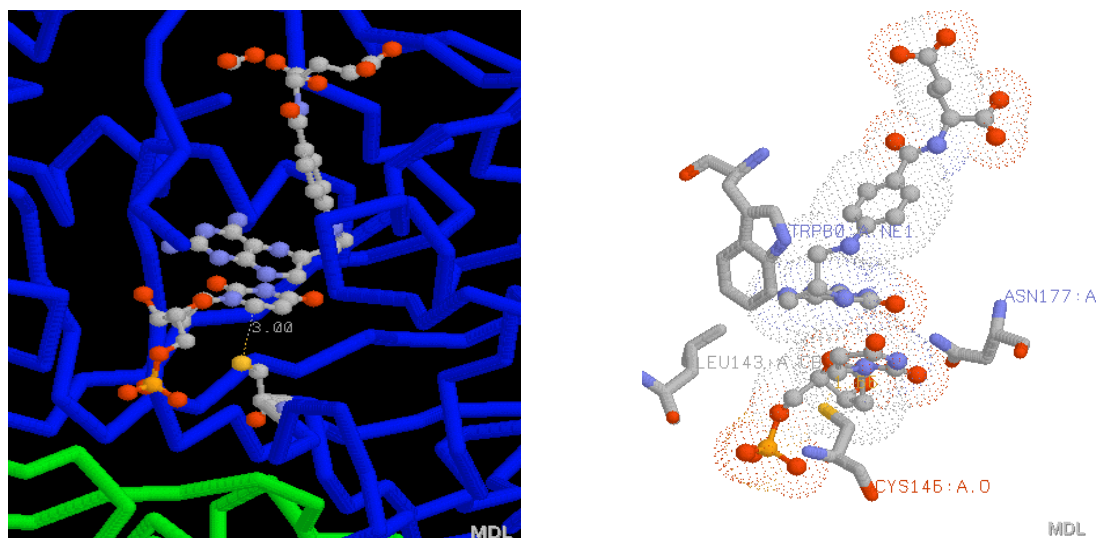
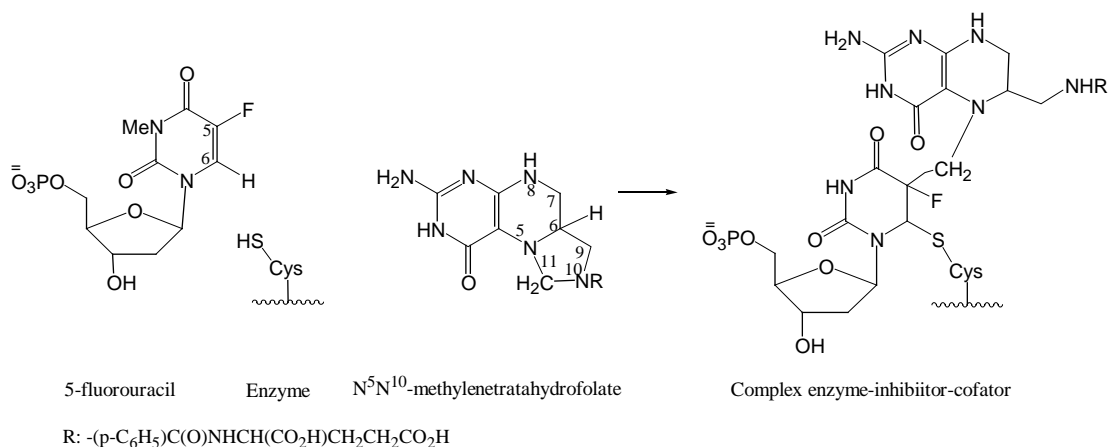


Fig. 4 Images obtained from PDB file 1kzi, showing tight contact between the cofactor, tetrahydrofolate (top), and the substrate, 2-deoxyuridylate monophosphate (middle). Cysteine residue (Cys146) is represented at the bottom, positioned for attack on substrate C4. In the left side picture, cofactor and substrate are drawn in balls and sticks, and the amino acid residues in sticks.

The elucidation of the mechanism of thymidylate synthase was through the structure determination of the complex formed when 5-fluoro-2'-deoxyuridylate (5FdUMP) inactivates the enzyme; the reaction involved is shown in Scheme 2 and can be viewed by the structures from PDB entry 1tfs, also from *E. coli* (10), Fig. 5.



Scheme 2 Inactivation of thymidylate synthase by 5-fluoro-2'-deoxyuridylate.

The presence of fluorine in place of labile hydrogen may disrupt the enzymatic reaction since C–F bonds are not easily broken. Its target enzyme accepts antineoplastic 5-fluorouracil, metabolically converted to 5-fluoro-dUMP (FdUMP), since it appears little different from the normal uracil substrate. However, the mechanism of the enzyme-catalyzed reaction is totally disrupted since fluorine cannot leave the molecule as the original hydrogen of C5 dUMP, which leaves as a positively charged species. As a result, a stable covalent complex is formed, linking inhibitor, cofactor, and enzyme, due to the strong fluorine electron-withdrawing potential, Scheme 2. The covalently linked complex undergoes an extensive conformational change, resulting in a structure represented by the C11–C5 methylene bridge between methylenetetrahydrofolate and FdUMP, which lies axial to the pyrimidine ring, in a similar configuration of thiol linkage between Cys146 and C6 of FdUMP [10].

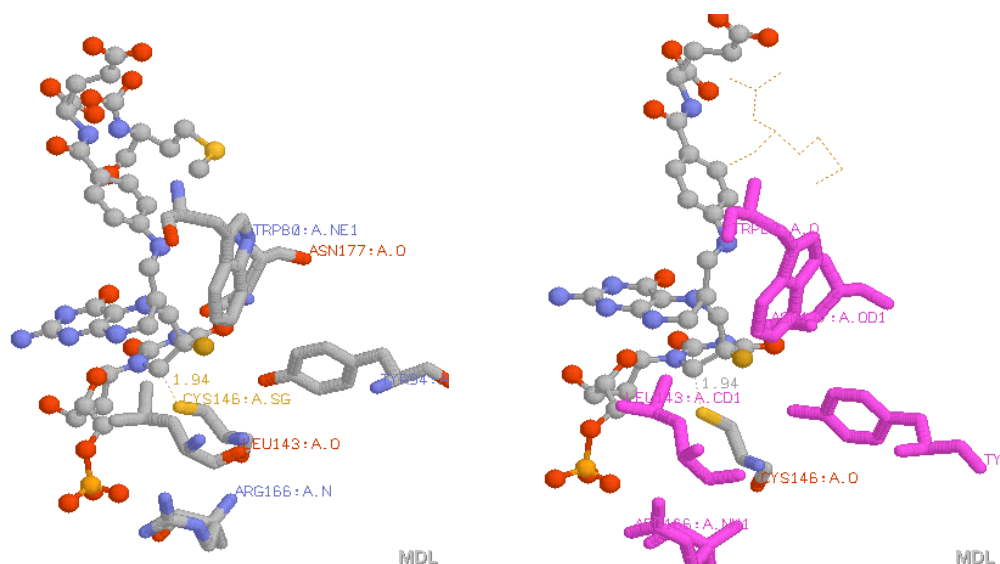


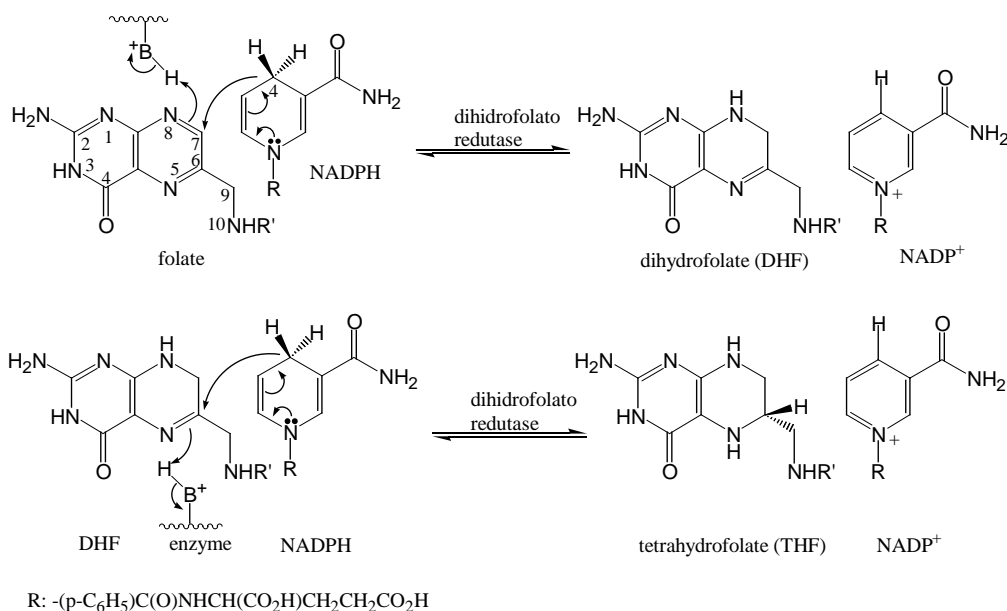
Fig. 5 Images obtained from PDB file 1t1s, showing the dihydrofolate cofactor (top), and the inhibitor, 5-fluoro 2'-deoxyuridylylate monophosphate (middle), bound by covalent linkage in the active site of thymidylate synthase. The cysteine residue (Cys146) is represented at the bottom, positioned to attack C4 of FdUMP. Cofactor and inhibitor are drawn in balls and sticks, and the amino acid residues in sticks. The left picture shows the target amino acids in magenta sticks.

Dihydrofolate reductase

Background

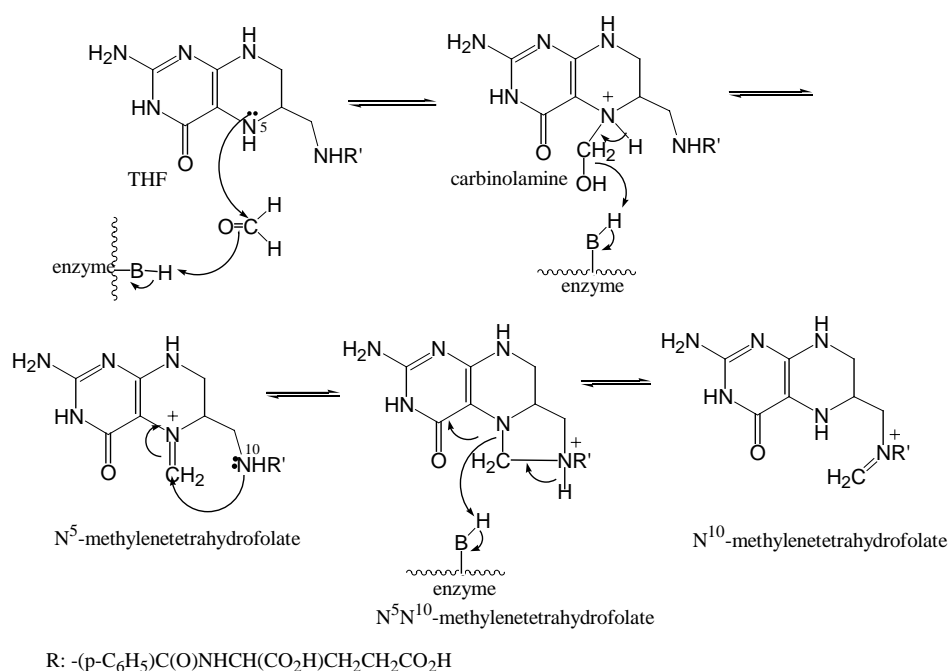
The effect of dihydrofolate antagonists on DNA synthesis results partially from dihydrofolate reductase (DHFR) inhibition, which depletes the N^5,N^{10} -methylenetetrahydrofolate pool, limiting thymidylate (dTMP) biosynthesis. Furthermore, the inhibition of enzyme additionally hinders nucleic acid biosynthesis by lowering the level of N^{10} -formyltetrahydrofolate, the formyl donor to glycinamide ribonucleotide in purine biosynthesis.

Two of the folic acid C–N double bonds are reduced stepwise by the nicotinamide adenine dinucleotide phosphate (NADPH)-dependent enzyme dihydrofolate reductase to produce tetrahydrofolate, which is recycled by reoxidation. The students draw the 2D stepwise reduction of folic acid to dihydrofolic and tetrahydrofolic acids in the presence of NADPH-dependent enzyme dihydrofolate reductase [8], as illustrated in Scheme 3.



Scheme 3 Dihydrofolate reductase-catalyzed reductions of folate and dihydrofolate, in the presence NADPH cofactor.

Thus, the students learn that tetrahydrofolate is not the full coenzyme; the complete coenzyme form contains an additional carbon atom between the N⁵ and N¹⁰ positions, provided by a formaldehyde molecule derived from the L-serine methylene group. Serine hydroxymethyltransferase catalyzes the attack of the more basic tetrahydrofolate N⁵-nitrogen to the formaldehyde group, originating N⁵-methylenetetrahydrofolate, by dehydration of the carbinolamine intermediate. N⁵,N¹⁰-methylenetetrahydrofolate and N¹⁰ and N⁵-methylenetetrahydrofolate cofactors are in equilibrium but the former is more abundant due to the closer proximity between the N¹⁰-amino group and the N⁵-methylene group (Scheme 4).



Scheme 4 Serine hydroxymethyltransferase-catalyzed reaction of formaldehyde and tetrahydrofolate giving methylenetetrahydrofolate.

Tutorial

Dihydrofolate reductase has at least 90 3D rendering structures in the PDB, many complexed to folic acid (substrate), NADP⁺ (cofactor) or inhibitors. Different organisms show slight differences in the structural features of the enzyme, but the DNA building function remains

intact. To understand folate binding to the active site and the important differences between substrate and inhibitor binding interactions, the students search PDB enzyme structures bound to folate and NADP^+ and, for comparison, the same structure bound to methotrexate, as an inhibitor (antimetabolite) that replaces and mimics folate. The major amino acid residues of the protein, which make nonbonded contacts with NADP^+ -folate and NADP^+ -methotrexate, may be visualized and inferred in terms of electrostatic and steric properties.

The folate- NADP^+ -enzyme complex, obtained from PDB entry 7dfr as a crystal structure of *E. coli*, has a long groove showing folate bound at one end, and NADP^+ at the other end [11]. By using the Protein Explorer tools, it is possible to visualize most of the energetically favorable binding interactions in the ternary NADP^+ -folate complex. Figure 6 illustrates the main interactions between folate and NADP^+ , such as the van der Waals contacts between the pteridine ring and the nicotinamide moiety. DHFR side-chains are wrapped around these two molecules, positioning them tightly next to one another (3.3 \AA) and making it easier for the hydride to be transferred from C4 in the nicotinamide NADPH moiety to C6 of the folate dihydropteridine ring. The tight contact cannot be relieved due to the close proximity of Phe31 to the pteridine ring in one side, forming a strong hydrophobic interaction above the binding site (3.2 \AA), and of Tyr100 close to the NADP-nicotinamide on the other side [11].

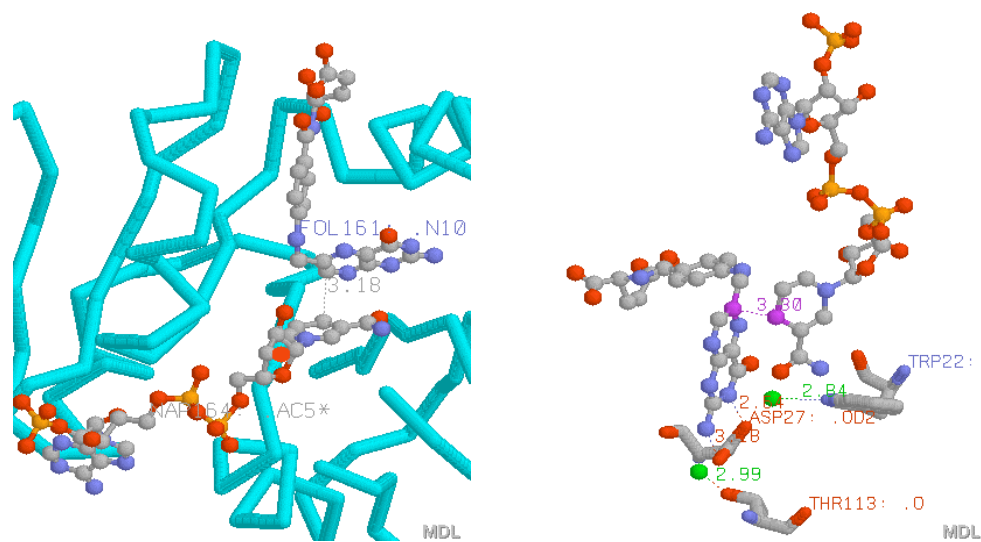


Fig. 6 Images from PDB file 7dfr, showing tight contacts between substrate (folate) and cofactor (NADP⁺) on the dihydrofolate reductase active site. Cofactor and substrate are drawn in balls and sticks, amino acid residues in sticks and water in green balls.

As can be seen from Fig. 6, dihydrofolate in a bent conformation in a deep hydrophobic cleft, binds to a single polar residue, Asp27, originating two distorted H-bonding with N3 (2.6 Å) and 2-amino group (3.2 Å) of folate. These interactions confer rigidity to the bound substrate and promote hydrophobic interactions with residues, such as Phe31, Ile5, and Met20. Asp27 is also important for substrate N5 protonation, via a water molecule, which increases the carbonium ion character at C6 NADPH, facilitating the hydride transfer. Figure 6 also shows the substrate positioned at the catalytic site by H-bonding the 2-amino and 4-oxo groups to two ordered water molecules, Wat301 and Wat206, which are, respectively, hydrogen-bonded to the conserved residues Thr113 and Trp22, and acting as a hydrogen-bonding bridge to Asp27. The greater specificity of the enzyme for dihydrofolate reduction as compared to folate can be explained by the extra hydrogen-bonding interaction of the N8 donor proton of dihydrofolate, not present in the folate, with the closer carbonyl group of Ile 5 (Fig. 7A). The *p*-aminobenzoyl moiety of the substrate occupies a hydrophobic pocket formed by residues Phe 31, Leu 28, and Leu 54. Finally, a

pair of hydrogen bonds completes the binding of the glutamyl group involving its carboxyl and the Arg 57 guanidinium groups.

On the other hand, the oxidized coenzyme NADP⁺ in an extended conformation appears to have more favorable interactions with DFHR through the pyrophosphate and the ADP 2'-phosphate groups. These are illustrated by 10 hydrogen bonds, 1 ionic interaction (His45), and 2 helix dipoles in the case of pyrophosphate and 5 hydrogen bonds and 1 ion pair (Arg44) with ADP 2'-phosphate [11]. It should be clear that the nicotinamide mononucleotide is held in the enzyme by hydrogen bonds (Arg7 and Ile14), dipole and van der Waals interactions (Ile94, Tyr100, and Thr46), and the 14-19 backbone of the Met20 loop closing over the bound nicotinamide ribose. The Met20 loop changes from the closed to the occluded conformation during the catalytic cycle, in movements that allow tetrahydrofolate release, assisted by NADPH binding [12].

Both, the 4-amino-4-deoxy antimetabolites of folic acid, aminopterin, and its N¹⁰-methyl homologue, methotrexate, bind tightly to the enzyme site through the 2,4-diaminopteridine ring, protonated at the physiological pH. These inhibitors are folate mimetics because of their similar size, shape, chemical composition, and binding position. Although the inhibitor complex has almost identical overall geometry, it binds to dihydrofolate reductase 1000 times more tightly than folate, blocking the action of the enzyme. The substrate binding interactions around the pteridine portion are inverted due a rotation of approximately 180° about an axis through the N5 and 2-amino group, and distorted when compared to the bound methotrexate. This can be seen by assessing PDB files entries 7dfr (dihydrofolate) [11] and 1ra3 (methotrexate) [13], both obtained from *E. coli* (Fig. 7). The 2-amino groups of both ligands form two hydrogen bonds with Asp27 and Wat301 (dihydrofolate) or Wat302 (methotrexate). The distortion of pteridine substrate allows its N3 group to be H-bonded to Asp27 (Fig. 7A), while methotrexate prefers to

position N1, allowing the H-bonding of N8 to Wat372, instead of binding to Ile5. As a result, the extra 4-amino group on methotrexate takes the place of N8 of the substrate and favors better contacts with the enzyme, involving two concomitant H-bindings to Ile5 and Ile94 (Fig. 7B).

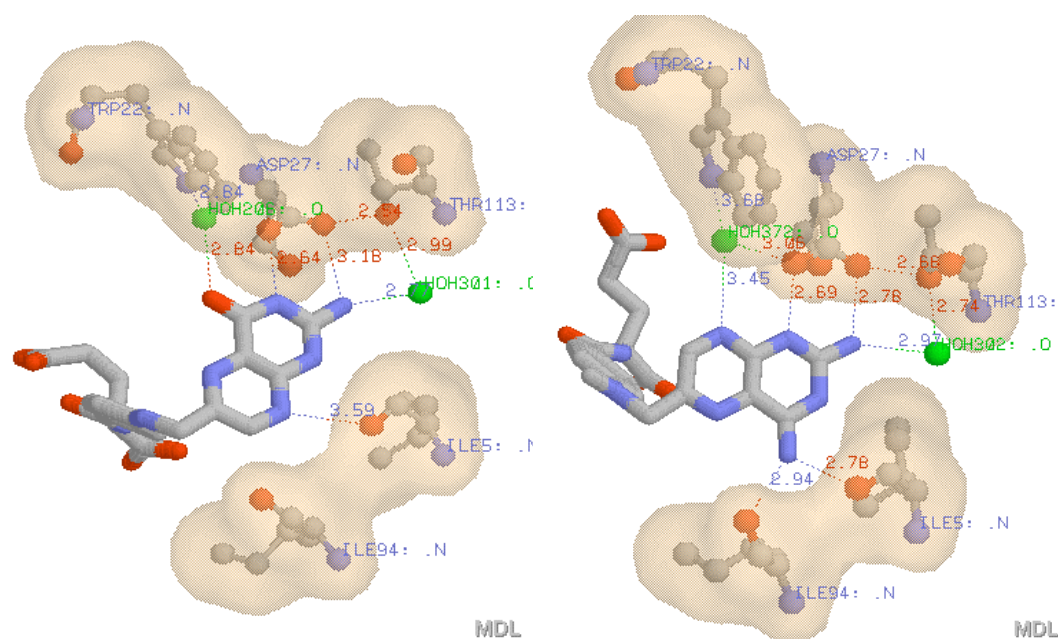


Fig. 7 Main interactions of (A) substrate (dihydrofolate) and (B) inhibitor (methotrexate) on the dihydrofolate reductase active site. Images from PDB entries 7dfr and 1ra3, using Protein Explorer tools. Substrate and inhibitor are drawn in stick, amino acids in ball and stick, colored by contact surface, and water in green balls.

A great deal of information is provided by these overall visualizations, concerning ligand-enzyme interactions and comparisons between substrate and inhibitor specificity. It is assumed that the *p*-aminobenzoyl-L-glutamate group of both substrate and methotrexate is bound in the same manner at the DHFR catalytic site.

Concluding remarks

Molecular modeling has been an essential tool in the teaching of relevant aspects of medicinal chemistry such as SARs and the molecular bases of antimetabolites/drug action. The 3D visualization and identification of enzymic 3D pharmacophores, provided by advanced graphic systems (Protein Explorer) and data bank (PDB) improve the students' understanding of the biological event by correctly predicting the binding affinity of a ligand to its target macromolecule. Also, by exploring the steric and electrostatic patterns involved in drug action, the students enhance their perception and appreciation of molecular recognition processes and their applications to drug design. Furthermore, computational chemistry offered visual insights, and the great deal of information obtained from PDB help the students to improve their scientific communication skills, such as 3D illustrations and animations in the final seminar presentations.

Supplemental material

Tutorials of Molecular Modeling Pro and Protein Explorer to perform the described exercises are available, respectively, as Supplemental Material 1 and 2.

Acknowledgments

The authors thank Prof. Dra. Zuleika Rotschild for valuable comments on the manuscript and the investigators who elucidated the 3D structures of target enzymes and deposited in the Protein Data Bank.

Literature cited

1. Cohen Claude N. *Guidebook on Molecular Modeling in Drug Design*, Academic Press: San Diego, 1996.

2. Leach, Andrew R. *Molecular Modeling: Principles and Applications*, Longman, Essex, 1996; pp 543-585.
3. Gringauz, A. *Introduction to Medicinal Chemistry: How Drugs Act and Why*, Wiley-VCH: New York, 1997.
4. Quinn, James A. *Molecular Modeling ProTM4.05: Computational Chemistry Program*, ChemSW[®] Inc, Fairfield, 2001.
5. Sansom, Clare E.; Smith, Christopher A. *Biochem. Educ.* **1998**, *26*, 103-110.
6. Wermuth, Camille G. *The Practice of Medicinal Chemistry*; Wermuth, Camille G Ed.; Academic Press: San Diego, 1996, pp 226-228.
7. Bajorath, J; Kraut, J.; Li, Z.; Kitson, D. H.; Haglet, A. T. *Proc. Natl. Acad. Sci.* **1991**, *88*, 6423-6426.
8. Silverman, Richard B. *The Organic Chemistry of Enzyme-Catalyzed Reactions*, Academic Press: San Diego, 2000, pp. 479-500.
9. Fritz, Timothy A.; Liu, Lu; Finer-Moore, Janet S.; Stroud, Robert M. *Biochemistry* **2002**, *41*, 7021-7029.
10. Hyatt, David C.; Maley, Frank; Montfort, William R. *Biochemistry* **1997**, *36*, 4585-4594.
11. Bystroff, Christopher; Oatley, Stuart J.; Kraut, Joseph *Biochemistry* **1990**, *29*, 3263-3277.
12. Sawaya, Michael R.; Kraut, Joseph *Biochemistry* **1997**, *36*, 586-603.
13. Bolin, Jeffrey T.; Filman, David J.; Matthews, David A.; Hamlin, Ronald C.; Kraut, Joseph *J. Biol. Chem.* **1982**, *257*, 13650-13662.

*Supplemental Material 1**Molecular Modeling Pro***Start > Programs > ChemSW > Molecular Modeling Pro*****I. Draw and represent a molecule***

- Observe instructions in the blue message box (upside window) that give you information about the program.
- Construct carbon chain.
- Click **Add** and **C**, click on the screen and then one letter “C” will appear. Click on the “C” to get a green line, which represents an ethane molecule. Click on one of the bond ends to build the desired carbon single bond chain.
- Substitute a carbon atom by oxygen. Click **Change**, hit the “O” button to select oxygen and click on the desired carbon chain.
- Add unsaturated bond, e.g., carbonyl group. Click **Add** and hit **O** and **2** buttons to select an oxygen and double bond, respectively. Click on the desired carbon on the drawing window.
- Add (–OH). Click **Add**, **O** and **1**, and click on carbon atom to have an alcohol instead of methylene group.
- Delete, click **Delete**, click on the end atom to remove; double-click **Delete**: opens screen to indicate, one atom, one molecule, or all.
- Add hydrogens. Click the “H” button.
- Join molecules. Hit the **Connect** button on the upper right and click on the two ends of both molecules.
- Introduce groups. Click **Rings** and select a ring from the list. To add the ring in a molecule, click **Delete** to remove one hydrogen atom, hit **Connect**, and then click the two ends.

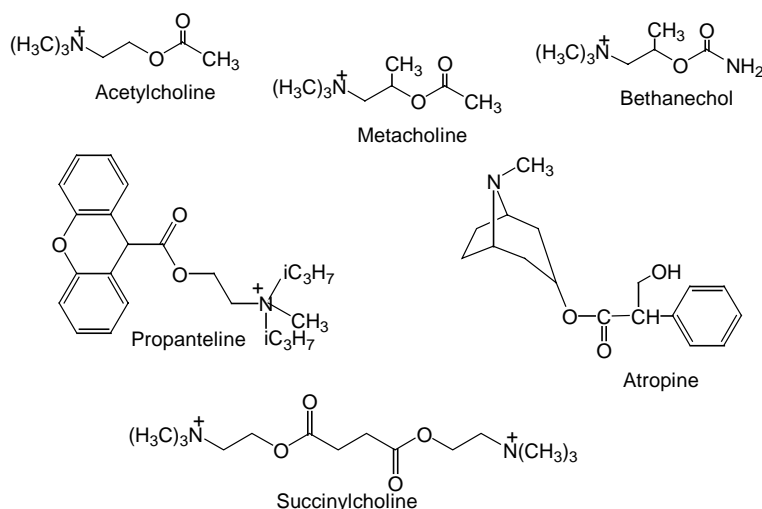
- Centralize the molecule. **Center**
- Increase or decrease size. **Scale**
- Construct two molecules in the same window. After construction of one, click **New mol** for drawing the second one.
- Hide hydrogens. **Display** (Main menu) > **Display hydrogens** (not activated)
- Save. **File** > **save** place extension (**macromodel file**-1 molecule, **MDL molfile** or **Reaction files**- 1, 2 or more molecules)
- Rotate molecules. **Rotate** > **X, Y, or Z-autorotate** or click mouse right button outside the figure.
- Move molecule. **Translate**, click in one of the atoms and drag it through the screen.
- Minimize energy. **Geometry** > **Minimize** (opens window, activate **Use ChemSite Amber**) click **Run**. Maximize window of **ChemSite**, click in **minimize Energy** (minimal energy, constant energy, or constant temperature). Save minimized molecule, close ChemSite and return to Molecular Modeling Pro.
- Orientation of planes X, Y. Start with minimized molecule, **Geometry** > **Place atom at origin** (click on original atom 1) **Geometry** > **Orient 1-2 atoms on x axis** (click on atoms 2 and 1) **Geometry** > **Orient 1-2 atoms in x-y plane** (click on atoms 3 and 2). Save minimized and oriented molecule on planes x and y.
- Superimpose molecules. Place molecules to be superimposed on the screen (minimized and oriented on x and y axis) and using tool **Translate** to place them where desired. To superimpose molecules in a systematic way, go to **Geometry** > **Dock (compare) molecule on molecule**, choose the atoms of both molecules by following the blue message box.

II. Exercises

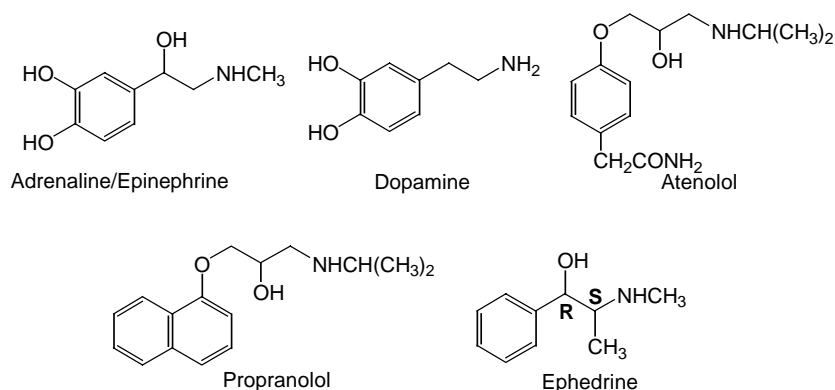
1. Calculation of parameters employed in the QSAR in groups of cholinergic (acetylcholine, metacholine, bethanechol, propanteline, atropine, and succinylcholine) and adrenergic drugs (adrenaline, dopamine, propranolol, atenolol, and ephedrine).

Separately, draw, minimize, align in a x-y plane, and save each molecule in a common directory in **My Documents** with **.dat** or **.mol** extension. After save, hit **Clear** button to draw other molecule.

Cholinergics:



Adrenergics:



File > Data base save > All macromodel files in directory (open screen directory C- indicate directory **My Documents** and click **OK**. Open option **Choose data base file name** indicate archive name, with extension **.csv** (save calculations in **Excel format**). The program

open a new window named **Select Properties to Store**, then choose desired parameters at **Select Physical Properties to Send to the Data Base** > Select: *interatomic distances, dihedral angle, molecular volume, Log P fragments, dipole moment, H bond acceptor, H bond donor*. **Done**. Follow the blue message box instructions in case of interatomic distance* and dihedral angle** parameters selections, otherwise click **ok**.

*For Interatomic distance:

cholinergic: click atoms between O (ester) and N (ammonium group), N–C–C–O.

adrenergic: click atoms between N (amino group) and the aromatic ring first carbon, N–C–C–C.

**For dihedral angle:

Cholinergics: between N–C–C–O bonds

adrenergic: N–C–C–C

Minimize Molecular Modeling Pro program, open data archive saved in Microsoft Excel (extension **.csv**). After organizing table data (containing QSAR) in Microsoft Excel, save with extension: working file of Microsoft Excel.

Analyze, compare data, and discuss the results.

2. Superimpose drugs, comparing acetylcholine to one of the following compounds:

bethanechol, propanteline, atropine, and succinylcholine; similarly, compare adrenaline with dopamine, propranolol, atenolol, and ephedrine.

Place two molecules from archive on the screen:

-File > Open (archive) > Open

- To place second molecule: **File > Open** (archive) > **Open** (when asked by option **Clear all molecules?** Click **No**)
- Separate molecules: **Geometry > Space molecules evenly**
- **Geometry > Dock (compare) molecule on molecule**
- Open screen: **number of atom pairs: 5, rotation angle: 2, translation increment: 0.1 > Done.**
- **** Observe blue message box** (*indicate comparative atoms, 4 for the first molecule, and 5 for the second*)
- Do not consider distance shown in the message box
- Rotate molecules (**Rrotate > X, Y, or Z- autorotate**)
- Click **Stop and Quit**
- Separate molecules (**Geometry > Space molecules evenly**)
- **Rotate > Rotation options > Rotate about molecule centers > Rotate all molecules > Done**

3. Analyses of the electrostatic density of cholinergic and adrenergic drugs (if preferred, place several molecules on the screen and work comparatively).

Display > (activate) labels

> (activate) **label hydrogens**

> (activate) **label carbons**

> (activate) **display hydrogens**

Display > Change display mode > activate: ball and stick, dot surface, display atom

labels, complete dot density: 3, activate color drawing > done (rotate molecule to analyze).

Try changes in menu **Format**

Vary parameters and diskette save the best visualization.

4. Liposolubility and partial charge analysis of cholinergic and adrenergic drugs (compare both groups).

Format > Color molecule lipophilicity

Format > Size by charges

Supplemental Material 2

Protein Explorer (PE)

I. Introduction

Protein Explorer (PE) is a macromolecular 3D visualization program freely available on the Web, accessed through NETSCAPE, version 4.7, but it requires other programs like JAVASCRIPT and CHIME, also easily available on the Web. For the search on proteins or enzymes, whose X-ray crystallographic and NMR data have been included in the Protein Data Bank (PDB), the address www.rcsb.org should be used. The 3D structure of macromolecules can be visualized and explored in the program.

The macromolecules could be proteins, DNA, RNA, carbohydrates, and complexes of these, as for example, transcription regulatory proteins and DNA or enzymes and drugs. All the thousands of macromolecules in the data bank have a specific PDB file code (ID) with four digits, related to its atomic coordinates. PE will not show proteins that only have the linear amino acid sequence.

PE will only show macromolecules if structural data, specifically atomic coordinates, are available. Among the 80 000 human proteins, many are available in PDB only as fragments, and for others there are not atomic coordinates. Although many macromolecular structures are experimentally determined by X-ray crystallography, several trans-membrane proteins (e.g., receptors) are difficult to crystallize, and few examples are known.

After searching for the key word (enzyme, substrate, cofactor, and/or inhibitor) in the PDB homepage, it is important to save all IDs obtained. Specific macromolecules will be further analyzed individually in greater detail as the structural and conformational aspects and possible interaction with ligands, related to the HETERO group (substrate, inhibitor, cofactor, etc.).

Note that after clicking “explore” for a single protein, a new page is opened containing several data and a list of resources on a left-side lateral column. Between these two are very important: **View Structure** and **Download/Display File**. Search the **Medline** available link to connect to the original abstract or full article.

The second option, **Download/Display File** is necessary for the linear analysis of amino acid sequences, fragments, filaments, and existing bridges, but mainly for identification and discrimination of the HETERO group and amino acids present at the CATALYTIC SITE. The HETERO group in the molecule can be recognized by the abbreviations: HET, HETNAM, and FORMUL. In the line called “**site**”, amino acids present in the catalytic site and their respective sequence number may be identified. This information, essential for further stages in the studies of substrate and/or inhibitor interaction at the catalytic site, should be registered separately.

Not all proteins in the PDB have ligands related to substrate and /or inhibitor, and the preference should be for the ones having HETERO groups

View Structure leads to another window for 3D visualization with about 5 program options; many of them require additional programs available in the same page. Clicking in **Protein Explorer** opens a new page containing four windows. The water-solvated macromolecule will appear on the right side rotating in space. It is necessary to observe instructions **BUSY/READY** in the window showing the molecule. **BUSY** (red) indicates that commands are being processed, and **READY** (green) operation is completed. The sign **READY** must substitute **BUSY** before other commands are issued to preserve **PE**.

II. Operations

Observation:  means click.

Enter site www.rcsb.org (open page-**PROTEIN DATA BANK-PDB**): inform key word or **ID** code > **Find a structure** > open page **QUERY RESULT BROWSER** > select codes of interest > **Explore** > open page **STRUCTURE EXPLORER**. If the ID code is unknown, search **Search fields** > inform key words > **Search** > open page **QUERY RESULT BROWSER** > select codes of interest > **Explore** > open page **STRUCTURE EXPLORER**:

- Summary information**
- View structure**
- Download/display file**

In the page **STRUCTURE EXPLORER**, **View Structure** > **PROTEIN EXPLORER (needs Chime)** > opens page **The Protein Explorer Will Load** > **Start Protein Explorer 2 Beta (Umass Server)** > **Start Explorer Session** > opens a small instruction page > **OK** only after viewing the molecule and the button **BUSY (red)/READY (green)** is **READY** > open page **First View** with three windows > **Troogle Spining** to stop the molecule rotation and **Hide/Show Water** to hide water molecule and facilitate program operation, since these options occupy large segments of memory and slow down the procedure.

Page **First View** is not very resourceful, but contains fundamental molecule information. It allows return to the main page **PROTEIN EXPLORER** through option **Eric Martz, Forms for Recording Observations, Molecule Information, Help/Index/Glossary, One hour tour** tutorials, **Quit**, and the very important option **Explore More with Quick Views**.

In **First View** window, click the scroll bar to roll to the end and **↘Explore more with Quick Views >** opens page **Quick Views**, containing 4 windows and one command dialog box.

- *Window 1* (top, left): tools to be explored for changes in macromolecular visualization.
- *Window 2*: (middle, left): explaining messages for each tool used in Window 1
- *Window 3* (bottom, left): message box, containing results and information obtained during operations.
- *Window 4* (lateral, right): molecule represented with desired changes.

The *command dialog box*, between *Windows 2 and 3*, is mainly used for PE experienced people.

In *Window 1*:

- ↘**Select**: the whole molecule or only parts may be selected.
- ↘**Display**: allows different format and representation of atoms in the molecule and selection of resources related to possible interactions in the macromolecule, such as contacts, surface.
- ↘**Color**: the whole molecule or parts may be colored in different ways.

In these operations, follow this order in the selection of menus, **Select > Display > Color**.

- ↘**Zoom+**: to increase the molecule size or **SHIFT + mouse left button**
- ↘**Zoom-**: to decrease molecule size or **SHIFT + mouse left button**
- ↘**Bkg**: to add or remove background black/white
- ↘**Water**: to show or hide water molecules
- ↘**Spin**: to rotate or stop rotation automatically
- ↘**Ligand**: to visualize or not the ligand related to the HETERO group

↖Center: to centralize the molecule. Clicking **↖Center** will open an optional window:

cancel or ok. If **↖Cancel** the whole molecule will be centralized; if **↖ok** one of the atoms can be selected to be centralized.

↖Stereo: to visualize or not pairs of molecule enantiomers

↖Slab: face cuts in the molecule to improve visualization

↖Slab + CTRL + mouse left button: permits visualization of the molecular interior; **CTRL**

+ mouse **right button** moves the molecule in the screen; mouse **left button** can rotate

molecule in the x - and y -axis and **SHIFT + mouse right button** rotate molecule in z -axis;

Observations: (i) remember that any utilized resource in *Window 1* will be explained in detail in *Window 2*; (ii) any atom selected or clicked in the *Window 4* molecule will be identified by its sequential number in the message box, *Window 3*.

Some of the options and links shown on the bottom of **Quick Views Window 1** are very important and should be mentioned:

↖Mol info: for more information on the molecules, to be used in subsequent visualizations

↖Reset view: to return to the **First View**, for the initial visualization of the molecule

↖Advanced Explorer: permits use of more advanced resources usually employed by people highly experienced in PE

III. Exercises

1. To localize residues or terminal atoms (C or N):

On page **Quick Views** indicate in **Select > ↖All; -Display > ↖Cartoon; Color > ↖N- >C Rainbow**; with the mouse click on the residues extremes and get the desired information on *Window 3*.

2. To determine protein composition:

On page **Quick Views**, indicate in **Select > ↘Protein; -Display > ↘Cartoon; -Color > Structure**. Observe information on *Windows 2 and 3* and type the command “**si**” (*command dialog box*)> enter. See the protein composition in *Window 3*.

3. To visualize hydrophilic and hydrophobic regions:

On page **Quick Views**, indicate in **Select ↘Protein; -Display > ↘Spacefill > -Color > ↘Polarity 2**

4. To visualize charge distribution on residues:

On page **Quick Views**, indicate in **Select ↘Protein; -Display > ↘Spacefill > -Color > ↘Polarity 5**

5. To visualize interactions between ligand and catalytic site residues:

There are two ways of doing this procedure: **5.1** and **5.2**

5.1. Usually employed when the catalytic site amino acids are described in the

Download/Display File:

On page **Quick Views > Mol Info >** using **options Header and Seq3D:**

↘Mol Info > opens page containing detailed information on the macromolecular structure and ligand; clicking **↘Header >** opens page **Structure Explorer: Download/Display File**, where **Site** (catalytic site that shows linear form of catalytic site amino acids, which chemically interact with the hetero-ligand) and **Hetero** (represents an abbreviation used to define the ligand interacting with the receptor by hydrogen bonds, van der Waals bonds or ionic bonds) should be assessed. Copy the catalytic site amino acids residues with respective numbers.

↖**Mol Info** > opens page, ↖**Seq3D** > opens page of **Seq3D**, containing 3 Windows:

Window 1': in **show clicked residue in** it is possible to select different molecule formats, including “**dots**” and “**accumulate selection**”.

Window 2': each mouse pointed amino acid in Window 3' will be identified in Window 2', that is, each letter representing the amino acid together with the linear sequence number.

Window 3': shows all the amino acid residues, identified through analytical methods and their corresponding linear position in the macromolecule.

With the help of the mouse, click in the catalytic site amino acids residues in Window 3', obtained from the previous step under **Mol Info** > **Header** options. The selected amino acids residues are discriminated in Window 2' and represented in the whole molecule in drawing *Window* (right side).

↖**Slab** + **CTRL** + mouse **left button** for improved visualization of interactions.

↖**Zoom** + and mouse rotate molecule for improved observation.

Clicking on macromolecule in *Window 4*, page **Seq3D** is automatically minimized in the screen's lower bar, but it can be reactivated if necessary.

On the page **Quick Views**, change format and color of the ligand, protein, or both of them, following the order **Select** > **Display** > **Color**.

5.2. When the catalytic site amino acids in the **Download/Display File** is not described:

On page **Quick Views**, indicate in *Window 1*: **Select** > ↖**Ligand**; **Display** > ↖**Contact**;

Color > **do not modify color**. Observe in *Window 2* of **Quick Views** the **Build Contact**

Surface options: () **Step by Step (for beginners)**; () **Step by Step (automatically)**; () **All**

at Once (for experts). Click in the first option ↖**Step by Step (for beginners)** > ↖**Show**

Contacts. There will appear 8 items to be selected in *Window 2*, click in all of them and

observe the molecule alterations presented on *Window 4*. **Zoom+** the molecule to improve

visualization. After clicking item 8, a new page appears in *Window 2*, named **Controls for Contact Surface**, where it is possible to see different ways of representing both the catalytic site and the ligand, during this operations observe the instructions to understand the meaning of the colors and formats. Clicking in an atom, it will be identified in *Window 3*.

Obs.: If **Contacts (Controls for Contact Surface)** is not visualized in *Window 2*, click in **Restore Contacts Help & Controls**.

The option **Select: Ligand, Display: Contacts** generates a contact surface between macromolecules and atoms of ligand. The spheres are atoms of amino acid residues sufficiently close to the ligand (3.5 to 4.5 Å, noncovalent-bonded, but linked by chemical interactions, such as hydrogen bonds). The sticks show more distant atomic residues, up to 7 Å of the contact surface. Interactions **cation-pi, salt bridges** (ionic bonds), and **H bond** (hydrogen bonds) in the menu **Display at Quick Views** page are shown for intramolecular interactions in the protein itself; thus, it is not possible to see these interactions directly with the ligand. These can be only inferred from the distances observed in the contact surface.

The surface of ligand atoms, previously selected, is colored (gray, white, and magenta) by the distance to neighboring atoms. Dark areas are very distant from any of the atoms of the macromolecule contact surface. Light or white areas are sufficiently close for van der Waals hydrophobic interactions. Pink or magenta areas are sufficiently close for hydrogen bonding. Oxygen and nitrogen favorable for hydrogen bonding are shown as spheres. Covalent bonds between close atoms are shown as sticks for any atom in the neighborhood of 7 Å. Spheres and sticks are colored according to the chemical elements (CPK, that is, red for oxygen, blue for nitrogen, and gray for carbon).

OBS.: If the catalytic site residues, obtained during procedure **5.1**, are known, it is possible to distinguish these amino acids in the representation of contact surface, obtained in exercise **5.2**. With the structure represented by contact surface in exercise **5.2**, open or

reactivate **Seq3D** (via **Mol Info**). In **show clicked residue in ▾Dots▾Accumulated selection**, and further **▾**in the catalytic amino acid residues, observing their identification in Window 2' of **Seq3D** and the changes of the molecular structure in **Quick View Window 4**.

6. To measure inter- and intra-atomic distances and dihedral angles, and identify with abbreviations the residues and ligands shown in Quick View Window 4

In the **Quick View** page, in **Select > ▾All, in Display > ▾Clicks**. In *Window 2*, select desired option, as, for example, to represent the *distance* in the structure (*Window 4*) or the message box (*Window 3*). Click on the two desired atoms to have the distance in Ångstroms. To change the option **▾ Stop** in *Window 2*. Select the option in *Window 2*, e.g., **> ▾Display labels on atoms**.

Ivone Carvalho*
carronal@usp.br

High standards in safety measures should be maintained in all work carried out in Medicinal Chemistry Laboratories.

The handling of electrical instruments, heating elements, glass materials, dissolvents and other inflammable materials does not present a problem if the supervisor's instructions are carefully followed.

This document has been supervised by Prof. Ivone Carvalho (carronal@usp.br) who has informed that no special risk (regarding toxicity, inflammability, explosions), outside of the standard risks pertaining to a Medicinal Chemistry laboratory exist when performing this exercise.

If your exercise involves any "special" risks, please inform the editor.