Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry
Bioorganic & Medicinal Chemistry 15 (2007) 2346–2358

The effect of 5-substitution in the pyrimidine ring of dUMP on the interaction with thymidylate synthase: Molecular modeling and QSAR

Adam Jarmuła,[a,*] Piotr Cieplak,[b] Tadeusz M. Krygowski[c] and Wojciech Rode[a]

[a]*Nencki Institute of Experimental Biology, Polish Academy of Sciences, Pasteura 3, 02-093 Warsaw, Poland*
[b]*Burnham Institute for Medical Research, 10901 N. Torrey Pines Road, La Jolla, CA 92037, USA*
[c]*Chemistry Department, Warsaw University, Pasteura 1, 02-093 Warsaw, Poland*

Received 2 August 2006; revised 24 December 2006; accepted 17 January 2007
Available online 19 January 2007

Abstract—Thymidylate synthase (TS) is a target enzyme for a number of anticancer agents including the 5-fluorouracil metabolite, FdUMP. The present paper reports on molecular modeling studies of the effect of substitution at C(5) position in the pyrimidine ring of the TS substrate, dUMP, on the binding affinity for the enzyme. The results of molecular dynamics simulations show that the binding of C(5) analogues of dUMP to TS in the binary complexes does not undergo changes, unless a substituent with a large steric effect, such as the propyl group, is involved. On the other hand, apparent differences in the binding of the TS cofactor, resulting from varying substitution at dUMP C(5), are observed in the modeled structures of the ternary complexes of TS. These binding characteristics are supplemented with a classical QSAR model quantifying the relation between the affinity for TS and the substituent electronic and steric effects of C(5) analogues of dUMP. Based on the findings from the present work, the perspectives for finding promising new C(5) analogues of dUMP as potential agents targeted against TS are discussed.

1. Introduction

Thymidylate synthase (TS) (EC 2.1.1.45) is an important enzyme with a central role in DNA synthesis. TS catalyzes the conversion of 2′-deoxyuridine 5′-monophosphate (dUMP) and 5,10-methylene-5,6,7,8-tetrahydrofolate (mTHF) to 2′-deoxythymidine 5′-monophosphate (dTMP) and 7,8-dihydrofolate (DHF) via reductive methylation, in which mTHF serves as both methyl donor and reducing agent.[1–4] The reaction is a terminal step in the *de novo* biosynthetic pathway leading to dTMP (one of the four building blocks of DNA). Inhibition of TS blocks DNA synthesis and prevents cell proliferation. Therefore, targeting of TS for inactivation in TS-expressing cells, such as tumor cells, has grown to an important strategy in the development of drugs for chemotherapy.[5,6] Many compounds modeled after the substrate (dUMP) or cofactor (mTHF) have been tested as TS inhibitors and some are in clinical trial or on the market as drugs in anticancer, antiviral or antifungal chemotherapy.[7–11]

Inhibitory properties and mechanisms of action of many dUMP analogues, most of them substituted at the pyrimidine C(5) carbon (for the labeling of the dUMP molecule, see Fig. 1), have been extensively studied for last decades.[12–19] Affinity toward TS of dUMP C(5) analogues has been shown to span a very wide range, depending on the nature of particular C(5)-substituent. One of the most potent inhibitors in this group is 5-fluoro-dUMP (FdUMP),[20] an active metabolite of the anticancer drug, 5-fluorouracil.[21] Mechanism-based inhibition of TS by FdUMP involves a time-dependent formation of the ternary covalent complex between TS, FdUMP, and mTHF, upon which the reaction stops, as the fluorine substituent fails to dissociate from the pyrimidine ring, resulting in a slowly reversible

Abbreviations: DHF, 7,8-dihydrofolate; dTMP, 2′-deoxythymidine 5′-monophosphate; dUMP, 2′-deoxyuridine 5′-monophosphate; FdUMP, 5-fluoro-2′-deoxyuridine 5′-monophosphate; mTHF, 5,10-methylene-5,6,7,8-tetrahydrofolate; THF, 5,6,7,8-tetrahydrofolate; TS, thymidylate synthase.

Keywords: Binding affinity; dUMP analogues; Ligand binding; Molecular dynamics; Molecular modeling; QSAR; Thymidylate synthase (TS); TS inhibitors.

* Corresponding author. Tel.: +48 22 5892297; fax: +48 22 8225342; e-mail: a.jarmula@nencki.gov.pl

0968-0896/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2007.01.021

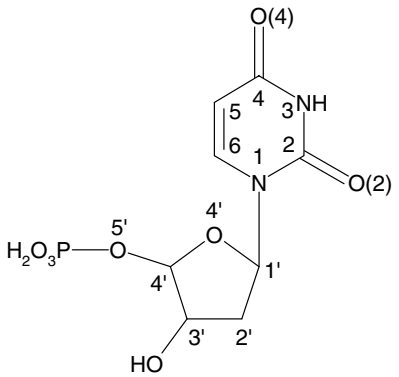

Figure 1. Schematic structure and labeling of the dUMP molecule.

inactivation of the enzyme. Other very strong mechanism-based inhibitors of TS, characterized without exception by the presence of strongly electron-withdrawing substituents, are 5-trifluoromethyl-dUMP,[20,22] 5-nitro-dUMP,[23–25] 5-formyl-dUMP,[20,26–28] and 5-ethynyl-dUMP.[18,29–31] On the other hand, weak inhibitory potencies against TS are usually characteristic for dUMP analogues with bulky substituents of either aliphatic or aromatic nature.[26,31–34]

The affinities for TS of dUMP C(5) analogues have been analyzed in relation to various parameters describing physicochemical properties of C(5) substituents in order to establish a valid quantitative structure–activity relationship (QSAR) model.[20,35] Those studies were only partially successful. They were able to distinguish between effects related to affinity (electronic and steric) and those that are not related to it, such as hydrophobicity. On the other hand, they failed to build models of a strong statistical significance due to low observation/descriptor ratios. In the meantime, the research focused on thymidylate synthase has been reinforced by the availability of three-dimensional structures from crystallography. The most prominent structural data in this field pertain to the binary TS–dUMP[36] and ternary TS–FdUMP–mTHF[37] crystallographic complexes. Other structural data related to dUMP or to the reaction product, dTMP, originate from the ternary TS–dUMP–5,6,7,8-tetrahydrofolate (THF)[38] and C146S TS–dTMP–DHF[39] complexes. However, the available structural data suffer from too much incompleteness to build a solid 3-D picture of how thymidylate synthase adjusts itself structurally to the substitutions at dUMP C(5) and how it relates to the affinity shown by C(5) analogues of dUMP.

5-Fluorouracil is not an optimal TS-inhibitory drug because it is inefficiently converted to FdUMP, with the rest of it catabolized to toxic metabolites. Tumors may also show intrinsic or acquired resistance to 5-fluorouracil.[40] The search for effective and specific inhibitors of TS continues, with one of the approaches focused on new substitutions at dUMP C(5). Therefore, a better understanding of the effect of C(5) substitution on the affinity for thymidylate synthase may help to strengthen the conceptual framework underlying this research. The present paper demonstrates results of molecular dynamics simulations for a series of binary complexes of TS with several C(5) analogues of dUMP. The binding behaviors of those analogues are derived from the trajectories and compared with each other. The average structures from the trajectories are superimposed on the structure of the ternary TS–dUMP–THF complex. This allows for a predictive assessment of the binding of the cofactor to each simulated complex. The computations are supplemented with a classical QSAR model relating the affinity for TS, shown by a series of C(5) analogues of dUMP, with physicochemical properties of the substituents at C(5).

2. Results and discussion

2.1. QSAR

To our knowledge, two QSAR models of the relationship between the activity of interaction with TS and physicochemical properties of the C(5) substituents in dUMP analogues have been published to date. Both models, by Wataya et al. ($r^2 = 0.908$, $s = 0.461$),[20] and by Jarmuła et al. ($r^2 = 0.997$, $s = 0.082$),[35] fall short in view of statistical relevance due to low observation numbers, each time of 9, and, with 3 descriptors used per model, unsatisfactory observation/descriptor ratios of 3. An obstacle for collecting more observations was due to both an insufficient reproducibility of the experimental activities from different laboratories and the lack of parameters for 'non-typical' substituents. Despite those problems, the models were helpful in recognizing primary substituent effects correlated with activity: electronic (represented by the resonance and inductive terms, σ^- and F in Wataya et al., and R_{new} and F_{new} in Jarmuła et al., respectively) and steric (represented by the molar refractivity M_R in Wataya et al., and by the Taft's E_s parameter in Jarmuła et al.).

In the present paper, we present an improved QSAR model for a series of C(5) analogues of dUMP interacting with TS. Difficulties to enlarge the observation pool persisted and with the 11 compounds in the pool the presented model improved very slightly compared to its predecessors. However, the electronic inductive and resonance effects of the substituents are represented in the latter model by a single descriptor, resulting in a two-descriptor model (where the second descriptor is a steric one) and so improving the observation/descriptor ratio to an acceptable value of 5.5. Eq. 1 shows the best fit to a two-descriptor model (with the 95% confidence limits for each term in parentheses):

$$\log[1/K_i^{app}] = 4.04(\pm 0.52)\sigma_p^o - 2.28(\pm 0.37)M_R + 1.10(\pm 0.32) \quad (1)$$

$n = 11$, $r = 0.992$, $r^2 = 0.983$, $s = 0.216$, $P < 0.0001$, $q^2 = 0.971$, SPRESS = 0.282.

The model explains 98.3% of the total variance in the dependent variable and demonstrates a very good internal predictive ability of 97.1% (see r^2 and q^2, respectively). The P value is extremely low, confirming that the

model predicts the dependent variable in a statistically very significant manner. The three-descriptor models were also constructed but showed virtually no improvement in the percentage of the explained variance (results not shown). The physicochemical parameters and the binding affinity data for each compound in the series is listed in Table 1. The plot of the predicted versus observed $\log[1/K_i^{app}]$ is shown in Figure 3.

The most important descriptor in the present QSAR is an electronic sigma *para* constant, σ_p^o.[41] In a one-descriptor model it alone explains about 57% of the variance in the dependent variable, $\log[1/K_i^{app}]$ (one-term equations are not shown). Although in the QSAR the reaction center at C(6) is in the *ortho* position relative to the C(5)-substituent (Fig. 1), sigma *ortho* constants fitted poorly to developed models and a best fitting sigma *para* constant was used instead. It is well known that the electronic actions from *ortho* and *para* positions are alike, provided no substantial steric effects of the substituent are involved.[42] In general, *para* electronic constant can be considered as a term wherein the inductive and resonance effects manifest in an appropriate blend. In Eq. 1 positive values of σ_p^o result in an increase of $\log[1/K_i^{app}]$. It means that the electron-withdrawing effects (both inductive and resonance) of the C(5) substituents, giving positive contributions to σ_p^o, tend to raise the affinities of the C(5) analogues of dUMP for TS. On the other hand, the electron-donating effects give negative contributions to σ_p^o and thus act oppositely, tending to lower the affinity level. These results can be rationalized in view of the initial step in the mechanism of TS-catalyzed reaction being the nucleophilic addition of the active site thiolate cysteine (Cys 146) to the C(6) atom of dUMP.[4] The addition reaction will be the more likely to occur, the more electrophilic C(6) is, which renders a supportive role for electron-withdrawing substituents at C(5). In the current series of dUMP analogues, compounds with the most strongly electron-withdrawing substituents, 5-nitro-dUMP and 5-formyl-dUMP, are the most active ones (see the values of σ_p^o and $\log[1/K_i^{app}]$ in Table 1). The nitro and formyl groups are engaged in conjugation with the π-electron pair from the pyrimidine C(5)=C(6) bond. The latter effect shifts electron density away from C(6), working

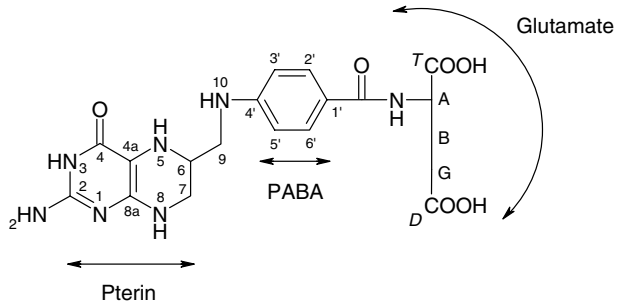

Figure 2. Schematic structure and labeling of the THF molecule.

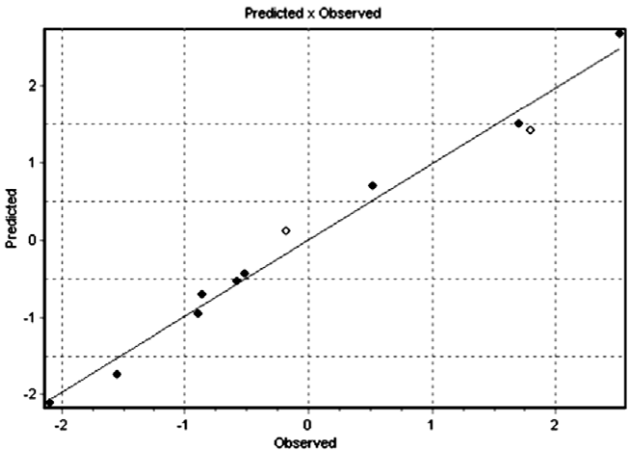

Figure 3. Observed $\log[1/K_i^{app}]$ values versus $\log[1/K_i^{app}]$ values predicted with QSAR Eq. 1. Filled or empty diamonds indicate residuals lower or higher than one standard deviation, respectively.

in the same direction as the electron-withdrawing inductive effect of these groups and thus consolidating the electrophilic character of C(6). This means that a cooperative inductive-resonance electron-withdrawing effect of substituent may grant C(5)-analogue of dUMP with an 'extra' binding affinity for TS. In comparison, the 5-fluoro and 5-chloro analogues of dUMP exhibit strong but relatively lower affinities for TS. In the context of electronic effects, the reason for this is that the resonance (electron-donating) effects of the halogen substituents are opposite to the inductive (electron-with-

Table 1. The physicochemical parameters and the binding affinity data used for deriving QSAR Eq. 1

Complex[a]	Substituent	σ_p^o	M_R	$\log[1/K_i^{app}]$[b]		
				Observed	Predicted	Residual
1	CH_3	−0.130	0.57	−0.86	−0.72	−0.14
2	C_2H_5	−0.127	1.03	−1.55	−1.76	0.21
3	C_3H_7	0.050	1.50	−2.09	−2.11	0.02
4	CHO	0.473	0.69	1.79	1.44	0.35
5	CH_2OH	0	0.72	−0.57	−0.54	−0.03
6	F	0.151	0.09	1.70	1.51	0.19
7	Cl	0.242	0.60	0.52	0.71	−0.19
8	Br	0.260	0.89	−0.18	0.12	−0.30
9	I	0.277	1.39	−0.89	−0.95	0.06
10	OH	−0.221	0.29	−0.51	−0.45	−0.06
11	NO_2	0.814	0.74	2.59	2.70	−0.11

[a] Complexes are numbered according to the identity of the substituent at dUMP. The numbering scheme is preserved throughout Tables 1–5.
[b] K_i^{app} values are in μM.

drawing) effects and partially counterbalance them, which is reflected in moderately positive values of σ_p^o for fluorine and chlorine (Table 1). Such imbalance between electronic effects is reversed in the case of the –OH substituent, where the electron-donating resonance effect outbalances the electron-withdrawing inductive effect (reflected in a negative value of σ_p^o; Table 1), resulting in a significantly lowered affinity of 5-hydroxy-dUMP for TS.

The second descriptor used in Eq. 1 is one of the most widely used steric parameters, molar refractivity, M_R,[43] which alone explains about 32% of the variance in $\log[1/K_i^{app}]$. In Eq. 1 large values of M_R result in a decrease of $\log[1/K_i^{app}]$, meaning that the steric effect tends to lower the affinity for TS in the series of the C(5) analogues of dUMP. This kind of influence is intuitively clear, considering the proximity of the reaction center at C(6) that should sensitize the binding behavior to changes in the bulk of the substituents at C(5). In the examined series, it is best illustrated by the example of the C(5)-halogen analogues, where the electronic effects of the substituents are of similar magnitude (as reflected by σ_p^o in Table 1), but differences in K_i^{app} of particular analogues are substantial and can only be explained due to the steric effect (see corresponding $\log[1/K_i^{app}]$ and M_R values in Table 1). Another illustration of the steric effect in the series are the lowest affinities for TS of the dUMP analogues substituted with the most bulky groups such as propyl, ethyl, and iodo.

2.2. Molecular dynamics: rmsds and rmsfs in the binary complexes

Table 2 presents the rmsds of the backbone atoms (N, C, CA, O) relative to either the energy-minimized (columns EQ w.r.t. MIN and AV w.r.t. MIN) or equilibrated (column AV w.r.t. EQ) structures of the investigated complexes. The values within each column are quite similar to each other, indicating no distinct differences between the global MD behaviors of each complex. The values in column AV w.r.t. EQ, representing the post-equilibration part of the simulations, are smaller than corresponding values representing the equilibration part (except for the TS–5-ethyl-dUMP complex) or the whole simulation (columns EQ w.r.t. MIN and AV w.r.t. MIN, respectively), meaning that the duration of equilibrations was sufficient to produce more stable trajectories thereafter. This is further confirmed by total potential and kinetic energies of the systems (not shown) converging to practically stable values after up to 120 ps of simulations (21 ps of gradual heating +99 ps of equilibration).

Table 2. The values of backbone rmsds (Å) in different simulation phases

Complex	Substituent	rmsd		
		EQ w.r.t. MIN[a]	AV w.r.t. EQ	AV w.r.t. MIN
1	CH_3	1.13	1.04	1.20
2	C_2H_5	1.12	1.15	1.23
3	C_3H_7	1.26	1.02	1.40
5	CH_2OH	1.15	0.83	1.15
6	F	1.21	0.84	1.27
7	Cl	1.07	0.97	1.15
8	Br	1.37	0.87	1.28
9	I	1.09	0.96	1.24
12	H	1.12	0.85	1.11

[a] EQ, MIN, and AV denote, respectively, the equilibrated, minimized and average structures.

Figure 4 shows the rmsfs per residue for the TS complexes with either unsubstituted or 5-substituted dUMP, the substitutions being either halogen (bromo) or alkyl (ethyl). Since we do not observe any significant differences between the data for these and other complexes in the series (not shown), these data can be perceived as representative for all simulations performed. The data show similar fluctuation profiles for each simulated system. The majority of the amino acids fluctuate to a moderate degree of 1.5–2.5 Å. The largest fluctuations indicated by peaks above approx. 2.5 Å (for the TS–dUMP and TS–5-ethyl-dUMP complexes) or 2 Å (for the TS–5-bromo-dUMP complex) correspond to loops belonging to surface regions of the protein. Among the most flexible areas are the C-terminal residues 257–264, the loop containing phosphate-binding Arg 21 and the region 215–222. The inner parts of the protein, including regions participating in binding and orienting the substrate, such as the central J helix (residues 174–192) containing Asn 177 hydrogen bonded to the pyrimidine ring of dUMP, fluctuate less intensively, showing overall good stabilities during the courses of simulations. The catalytic Cys 146 conforms to such a trend by undergoing regular fluctuations of low intensity.

2.3. Molecular dynamics: average structures of the binary complexes

The average structures resulting from MD simulations are characterized by (i) backbone (N, C, CA, O) and overall (all heavy atoms) rmsds relative to the average structure of thymidylate synthase in the complex with its native substrate, dUMP, and (ii) distances and other geometric parameters describing the binding positions of ligands with respect to the enzyme active site residues playing known mechanistic roles. Inspection of the rmsd values (Table 3) shows that the halogen substituents affect the backbone and overall structures less than substituents comprising at least one carbon unit (except that the overall rmsd for –I is slightly higher than for $–CH_3$). In the case of halogens, the largest rmsds correspond to the structures containing dUMP with either most bulky (iodo) or most polar (fluoro) substituents. In the rest of the series, the overall rmsds increase in parallel to increasing bulk of substituents, whereas the backbone rmsds increase in a slightly different order (compare M_R in Table 1 as a measure of substituent bulk). Both the backbone and overall rmsds clearly indicate the structure of the TS–5-propyl-dUMP complex, and the overall rmsds also, but less so, the structure of the TS–5-ethyl-dUMP complex, as two most different structures in the series with respect to the structure of the native complex, TS–dUMP. Characteristically, both 5-propyl-dUMP and 5-ethyl-dUMP are by far the least active

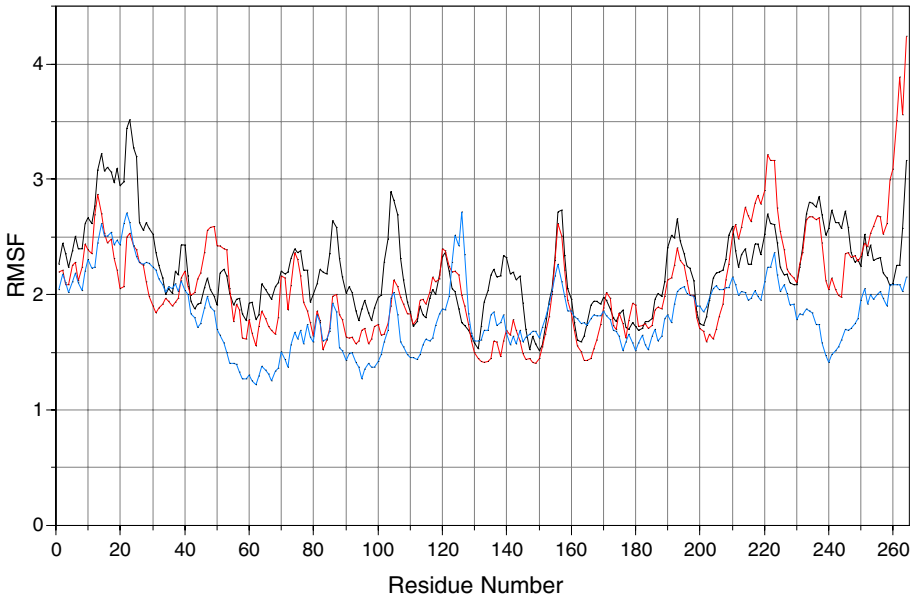

Figure 4. The backbone rmsfs per residue (Å) during simulations of the TS–dUMP (black), TS–5-ethyl-dUMP (red), and TS–5-bromo-dUMP (blue) complexes.

Table 3. The values of backbone and overall rmsds (Å) relative to the average structure of the TS–dUMP complex

Complex	Substituent	Backbone rmsd	Overall rmsd
1	CH_3	1.08	0.91
2	C_2H_5	1.23	1.31
3	C_3H_7	1.55	1.61
5	CH_2OH	1.25	1.09
6	F	0.71	0.86
7	Cl	0.59	0.73
8	Br	0.64	0.83
9	I	0.77	0.96

inhibitors of TS in the series, meaning that departures from the structure of the native complex have a strong reducing effect on the affinity for TS. Here, it should be noted that although the activity data in the present QSAR relate to the inhibition of mammalian (L1210) TS, corresponding data relating to inhibition of bacterial (*L. casei*, *E. coli*) TSs were also reported for the majority of the analogues from the present series.[13] A good qualitative agreement between both data sets (mammalian and bacterial; not shown) warrants discussion of the structural features of bacterial (*E. coli*) TS in correlation to activities obtained with mammalian (L1210) TS.

In the simulations, the emphasis was laid on modeling all studied complexes as non-covalent in order to compare the binding behaviors in a uniform sample and to allow a clear interpretation of the results. Non-covalent complexation is typical in the binary TS complexes, although some dUMP analogues such as 5-bromo-dUMP and 5-iodo-dUMP have been shown to form covalent complexes.[44] Taking no account of the latter effect should not affect the present results, which are aimed at gaining a picture of the conformational and structural changes in the enzyme ensuing from substitutions at C(5) of dUMP, rather than at analyzing the electronic and chemical changes that associate formations of covalent complexes.

The distances between the Cys 146 sulfur and pyrimidine C(6) carbon atoms (which later in the process of catalysis form a bond with each other) are similar in the series, except for the structure of the TS–5-propyl-dUMP complex, wherein this distance is 5.65 Å, compared to 3.27-3.40 Å elsewhere in the series and 3.16 Å in the crystal structure of the TS–dUMP complex (Table 4). The angle and distance conditions are satisfactory for the presence of hydrogen bonding between the pyrimidine N(3)–H and O(4) groups and Asn 177, as well as between the pyrimidine O(2) atom and Asp 169, in all average structures except that of the TS–5-propyl-dUMP complex, wherein O(2) and the amide nitrogen of Asp 169 are separated by a distance of 5.23 Å, at which hydrogen bond cannot be formed (Table 4). The N(3)–H–O–Asn 177 and O(4)–H–N–Asn 177 hydrogen bonds are very important for recognizing and orienting the molecule of substrate in the active site of the binary complex.[4,45,46] Our results show that this double H-bonding is strong and persists even if the substituent at C(5) exhibits a major steric effect, such as the propyl group.

The sugar and phosphate moieties of the ligands maintain well hydrogen bonding to neighboring residues in the active site. Apart from the TS–5-propyl-dUMP complex which shows no such contacts, the sugar ring is always coordinated by the hydrogen bonds formed between the 3′-deoxyribose hydroxyl group and at least one residue, or both residues, of the His 207-Tyr 209 diad. A single H-bonding is present in only two complexes, TS–5-methyl-dUMP (dTMP) and TS–5-iodo-dUMP, both lacking an H-bond to the hydroxyl group of Tyr 209 (Table 4). However, the hydrogen bonds with

Table 4. Distances (Å) from dUMP C(6) to γS of catalytic Cys 146 and the presence (+) or absence (−) of hydrogen contacts to dUMP in the binary complexes

Complex	Substituent	C(6)–γS distance	Hydrogen contacts with				
			Asn 177	Asp 169	His 207	Tyr 209	Arg 21 and Arg 166[a]
1	CH_3	3.39	+	+	+	–	+
2	C_2H_5	3.27	+	+	+	+	+
3	C_3H_7	5.65	+	–	–	–	+
5	CH_2OH	3.29	+	+	+	+	+
6	F	3.29	+	+	+	+	+
7	Cl	3.29	+	+	+	+	+
8	Br	3.40	+	+	+	+	+
9	I	3.38	+	+	+	–	+
12	H	3.28	+	+	+	+	+
Crystal	H	3.16	+	+	+	+	+

[a] Simulations were performed for a single TS unit. Two arginines (Arg 126′ and Arg 127′) contributing from the other unit to the phosphate-binding site were not included.

His 207 and Tyr 209 are less important in orienting the nucleotide than the hydrogen bonds with Asn 177.[4,36] In fact, dTMP and 5-iodo-dUMP show similar orientation to other ligands in the series (apart from 5-propyl-dUMP), which all form H-bonding with Asn 177. In addition, a proper positioning of ligands in the active site is maintained by strong attractive electrostatic interactions, including hydrogen bonds, between the deprotonated phosphate moiety and two positively charged arginines, Arg 21 and Arg 166. The phosphate is anchored this way in all ligands in the series (Table 4).

The conformational arrangement of the ligand molecules themselves is similar in the series of average structures. In particular, the bases are in the *anti* orientation relative to the sugars in all ligands except the 5-propyl-dUMP molecule, where the base is oriented *syn* to the sugar. Pyrimidine nucleotides are known to significantly prefer the *anti* conformation,[47] the one found also in the crystal structure of the TS–dUMP complex. In the TS–5-propyl-dUMP complex structure, the propyl moiety is too big to be accommodated at the binding site without a substantial structural rearrangement. Therefore, the pyrimidine ring twists and adopts the *syn* orientation relative to the sugar, in order to dispose the propyl group between the indole rings of the two tryptophanes, Trp 80 and Trp 83, pushing them further away from each other by 1.1-1.6 Å and widening the angle between them relative to their positions in the crystal structure of the TS–dUMP complex (Fig. 5). This rearrangement does not disturb the hydrogen bonding of O(4) and N(3)-H with Asn 177, but locates the C(6) and O(2) atoms each about 2.2–2.3 Å further from the sulfur of Cys 146 and the amide nitrogen of Asp 169, respectively, as compared to the rest of average structures. It has a detrimental influence on the binding of the cofactor, as will be discussed in the next section of the paper.

2.4. Superimposing binary complexes with the ternary TS–dUMP–THF complex

The chemistry related to the catalytic reaction of thymidylate synthase depends on proper alignment of substrate and cofactor molecules in the ternary complex. The substrate, bound to TS and properly oriented, forms a binding surface for the cofactor. The cofactor binds as the second in order, following substrate binding, with the pterin ring stacked parallel to the pyrimidine ring.[48,49] In the present series of binary complex structures, all ligands except one are oriented in similar way as the native substrate, dUMP, in the crystal complex with TS. As discussed in detail earlier in the paper, an alteration to binding orientation of the nucleotide and some reorganization to the structure of the binding pocket were present in the TS–5-propyl-dUMP complex (Fig. 5). For the rest of average structures, the binding characteristics were largely similar, resulting in only minor differentiation between the architectures of the binding pockets of each complex. Nevertheless, some protein regions adjacent to the cofactor-binding site, such as especially the segment containing Trp 80 and Trp 83, showed structural differences to a relatively larger extent (not shown). In our superimposition approach, we proceed to assess with how much efficiency the cofactor would bind to each binary complex in the examined series. The latter is investigated by addressing two main questions: How the cofactor molecule aligns against each substrate analogue? How much stabilization it gains upon the binding via favorable hydrophobic contacts to the protein? Both issues will be successively discussed in the following paragraphs.

The alignment of the cofactor molecule against the molecules of each C(5)-analogue of dUMP, calculated using the superimposition approach, is illustrated by examples presented in Figure 6. The substrate analogues can be divided in two groups, depending on the occurrence, or lack of occurrence, of a deviation from parallel orientation between the pyrimidine (nucleotide) and pterin (cofactor) rings. In the first group, the orientation is completely (or almost completely) parallel, meaning a solid stabilization of the cofactor position in the ternary complex due to favorable stacking interactions between both rings. The first group is represented by six compounds, characterized by the presence at C(5) of hydrogen or substituents, such as fluoro, methyl, chloro, hydroxymethyl, and bromo, whose steric effects correspond to the M_R values in the range of up to 1 (see Table 1). In the first group, distances between the pyrimidine and pterin rings are similar, as illustrated by a distance separating the cofactor N(5) and nucleotide C(5)

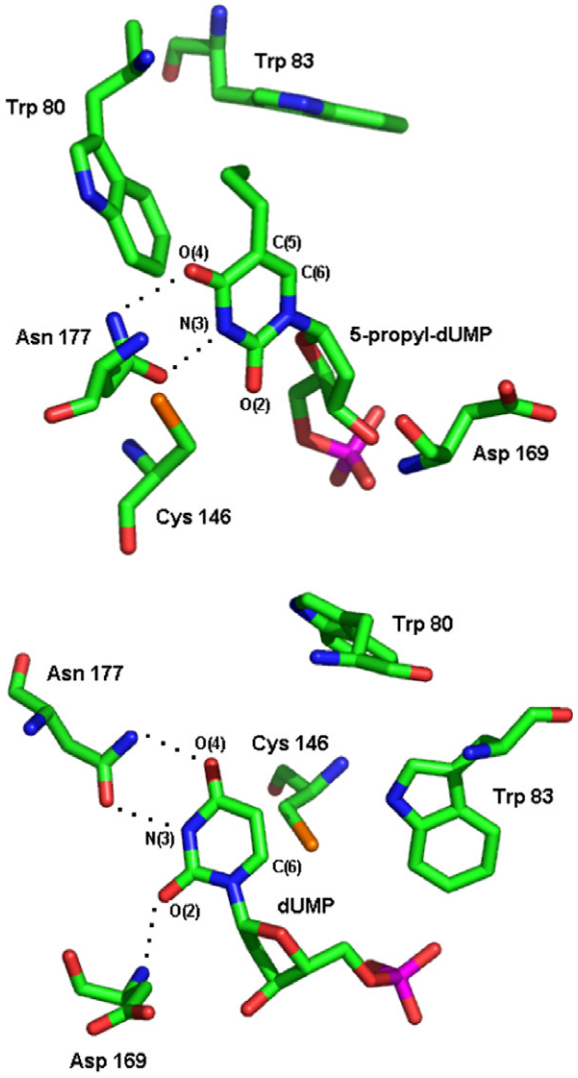

Figure 5. Substrate binding-pocket arrangements in the binary complexes, modeled, TS–5-propyl-dUMP (top) and crystal, TS–dUMP (bottom). Carbon atoms are green, oxygen atoms are red, nitrogen atoms are blue, phosphorus atom is violet, and sulfur atom is yellow. Both ligands preserve the hydrogen bonding with Asn 177 (dashed lines). The pyrimidine ring of 5-propyl-dUMP is oriented such that to allow the propyl group at C(5) to penetrate between the indole rings of Trp 80 and Trp 83. In comparison to dUMP, its O(2) atom lacks hydrogen bonding to Asp 169 and its C(6) atom—proper alignment for nucleophilic addition of Cys 146. Other binding-pocket residues and hydrogens are omitted for clarity.

atoms (which are bonded to each other through a methylene bridge upon formation of covalent bondage between components of the ternary complex), ranging, in individual superimpositions, from 3.10 Å to 3.36 Å. In the second group, there are the three remaining compounds that contain substituents such as ethyl, propyl, and iodo, whose steric effects correspond to the M_R values in the range of above 1 and so are stronger compared to the first group. According to our superimpositions, the THF molecule binds to the TS–5-propyl-dUMP complex differently than in the rest of models. Resulting from the 5-propyl-dUMP molecule being bound in an unusual orientation (see discussion in the previous section), the THF molecule finds a most adaptive location for binding, such that the pterin ring is twisted about 40° relative to the pyrimidine ring and looses part of the favorable stacking contacts to it (Fig. 6). The pyrimidine C(5) to pterin N(5) distance is 3.97 Å, with N(5) overlapping against nucleotide O(4) rather than C(5). In the case of the TS–5-ethyl-dUMP–THF and TS–5-iodo-dUMP–THF complex models, stabilization to the pterin ring is affected less severely, as alterations to the native orientation of this ring are less detrimental and involve its relatively small deviation from the parallel alignment with respect to the pyrimidine ring (Fig. 6). Compared to the first group, distances between pyrimidine C(5) and pterin N(5) are larger and equal to 3.60 Å and 3.85 Å in the models of the complexes containing 5-iodo-dUMP and 5-ethyl-dUMP, respectively. In addition, the overlapping between N(5) and C(5) is poorer.

As known from crystallographic studies, apart from overlapping with the pyrimidine ring, cofactor position in the binding pocket is stabilized to a large extent by numerous hydrophobic contacts made by the pterin and PABA rings to side chains of some proximal amino acid residues, including the following sextet: Ile 79, Trp 80, Trp 83, Leu 143, Leu 172, and Phe 176.[3] To define and compare maps of contacts made to the cofactor in the models of each ternary complex in the series, we plotted hydrophobic interactions involving the THF molecule as determined by a distance cutoff of 4 Å using the *LIGPLOT* program.[50] The residues, which are engaged in interactions, are listed in Table 5. The most thorough and complete stabilization of cofactor position was observed in the model of the ternary complex involving FdUMP, where the five hydrophobic residues of Ile 79, Trp 80, Trp 83, Leu 172, and Phe 176, and the base of FdUMP, form together a symmetric shell around the cofactor rings. From this shell, hydrophobic contacts are made in a seemingly cooperative manner to all carbon atoms of the PABA and pterin rings (Fig. 7a). These contacts correspond very well (all present) to cofactor contacts in the crystal structure of the TS–FdUMP–mTHF complex (not shown). In the crystal structure, the mTHF molecule has an additional contact to Leu 143, which in our model is positioned about 0.5 Å further from the cofactor and slightly outside of the 4 Å range. Other structures, showing hydrophobic contacts of the cofactor molecule with as much as five residues in the binding pocket, include complexes with dUMP and dTMP. The combination of residues making contacts in the model of the TS–dTMP–THF complex is the same compared to the model of the TS–FdUMP–THF complex, whereas in the model of the native complex, TS–dUMP–THF, the stabilizing contacts are provided through another (but similar) combination of residues, including Trp 80, Trp 83, Leu 143, Leu172, and Phe 176. The contacts in the model reproduce quite well those present in the crystal structure of the TS–dUMP–THF complex, except that in crystal Leu 172 is shifted to a distance of above 4.5 Å away from the cofactor. In addition, in the crystal structure the side chain of Val 262 is closer and provides two contacts to the cofactor, one for the PABA and the other for the pterin ring.

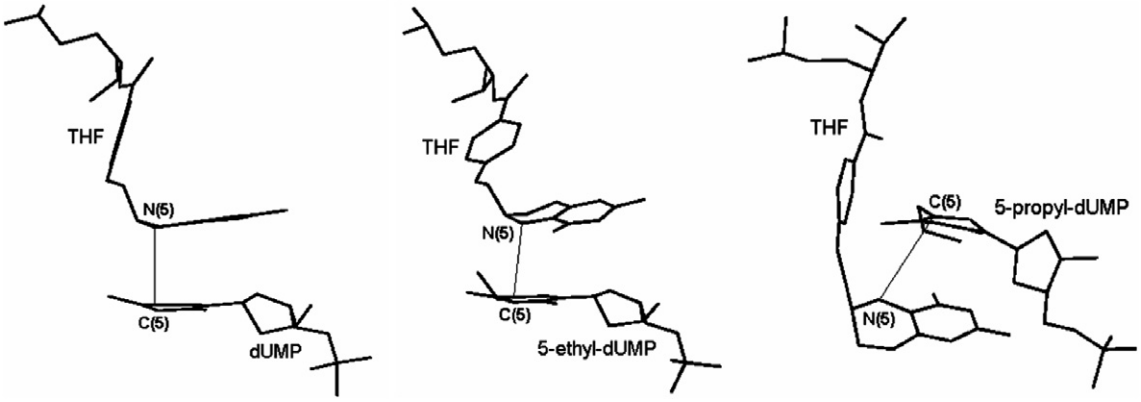

Figure 6. Three modes of alignment of the cofactor, THF, with respect to dUMP or its C(5)-analogue, present in the models of the ternary complexes: parallel (left), deviated from parallel (middle), and deviated from parallel as well as disturbed (right). The slopes of the solid lines connecting THF N(5) and dUMP C(5) are indicative of the degree of overlapping (left and middle), or of the lack of overlapping (right), of these atoms. The alignments shown are from the TS–dUMP–THF (left), TS–5-ethyl-dUMP–THF (middle), and TS–5-propyl-dUMP–THF (right) complexes. This figure and Figure 5 were prepared with the PyMOL program.[51]

Table 5. Hydrophobic contacts to the PABA and pterin rings of the cofactor as evaluated using 4 Å cutoff in the TS ternary complexes

Complex	Substituent	No. of residues contacting PABA and/or pterin rings	Residues contacting PABA and/or pterin rings
1	CH_3	5	Ile 79, Trp 80, Trp 83, Leu172, Phe 176
2	C_2H_5	3	Ile 79, Leu172, Phe 176
3	C_3H_7	5	Ile 79, Trp 80, Trp83, Leu143, Phe176
5	CH_2OH	4	Ile 79, Trp80, Trp83, Phe 176
6	F	5	Ile 79, Trp 80, Trp 83, Leu 172, Phe 176
Crystal	F	6	Ile 79, Trp 80, Trp 83, Leu 143, Leu 172, Phe 176
7	Cl	4	Ile 79, Trp 83, Leu 172, Phe 176
8	Br	4	Ile 79, Leu 172, Phe 176, Ala 263
9	I	3	Trp 80, Leu 172, Phe 176
12	H	5	Trp 80, Trp 83, Leu 143, Leu172, Phe 176
Crystal	H	5	Trp 80, Trp 83, Leu 143, Phe 176, Val 262

The reach of stabilizing contacts to the cofactor in the complexes containing dUMP and FdUMP looks very reasonable, considering that dUMP is the native substrate, and FdUMP—a strong mechanism-based TS inhibitor and one of the most active dUMP analogues in the series. On the other hand, dTMP has one of the weakest activities in the series, but in spite of that, in the model of the TS–dTMP–THF complex, the THF molecule appears both well positioned for catalysis and stabilized through contacts with protein to a level comparable to that in the models and crystal structures of the complexes containing dUMP and FdUMP. It was shown that the chemical steps of catalysis, occurring once the productive alignment of ligands and catalytic residues is reached, involve only minor conformational rearrangements.[3] Therefore, the binding of ligands in the ternary product complex, C146S TS–dTMP–DHF, at the point prior to complex dissociation, was found to closely resemble the binding in a tight ternary complex formed between TS, the substrate, dUMP, and a cofactor analogue, CB3717.[52] It was proposed that TS accomplishes the substrate/product discrimination through a water-mediated mechanism in which the enzyme uses a bound water molecule to disfavor binding of dTMP.[39]

Except of dTMP (and excluding the least active 5-propyl-dUMP for its different mode of binding), cofactor stabilization through hydrophobic contacts at the binding site shows a qualitative correlation with the activity of C(5) analogues of dUMP in the TS reaction (see Tables 5 and 1, respectively). The complex models of moderately active C(5) analogues, such as, in decreasing order of activity, 5-chloro-dUMP, 5-bromo-dUMP and 5-hydroxymethyl-dUMP, show cofactor contacts to four hydrophobic residues each. The complex models of the second and third least active C(5) analogues in the series, 5-ethyl-dUMP and 5-iodo-dUMP, respectively, in addition to a poor overlapping between the pterin (cofactor) and pyrimidine (C(5)-analogue) rings, show cofactor contacts to only three hydrophobic residues each. The plot of cofactor interactions in the model of the TS–5-iodo-dUMP–THF complex, indicating the lack of contacts to three carbon atoms of the pterin ring, is shown in Figure 7b.

In the model of the TS–5-propyl-dUMP–THF complex, positions of both ligands are strongly disadvantageous in the context of catalysis. The THF molecule contacts five hydrophobic residues (Table 5), suggesting that it is well stabilized at the binding site. Positioned as it is, however, the cofactor molecule is not correctly disposed to bind to C(5) of the pyrimidine ring of the nucleotide (Fig. 6). Moreover, the initiation of catalysis by the nucleophilic addition of Cys 146 to pyrimidine C(6) appears very problematic in this model, providing that

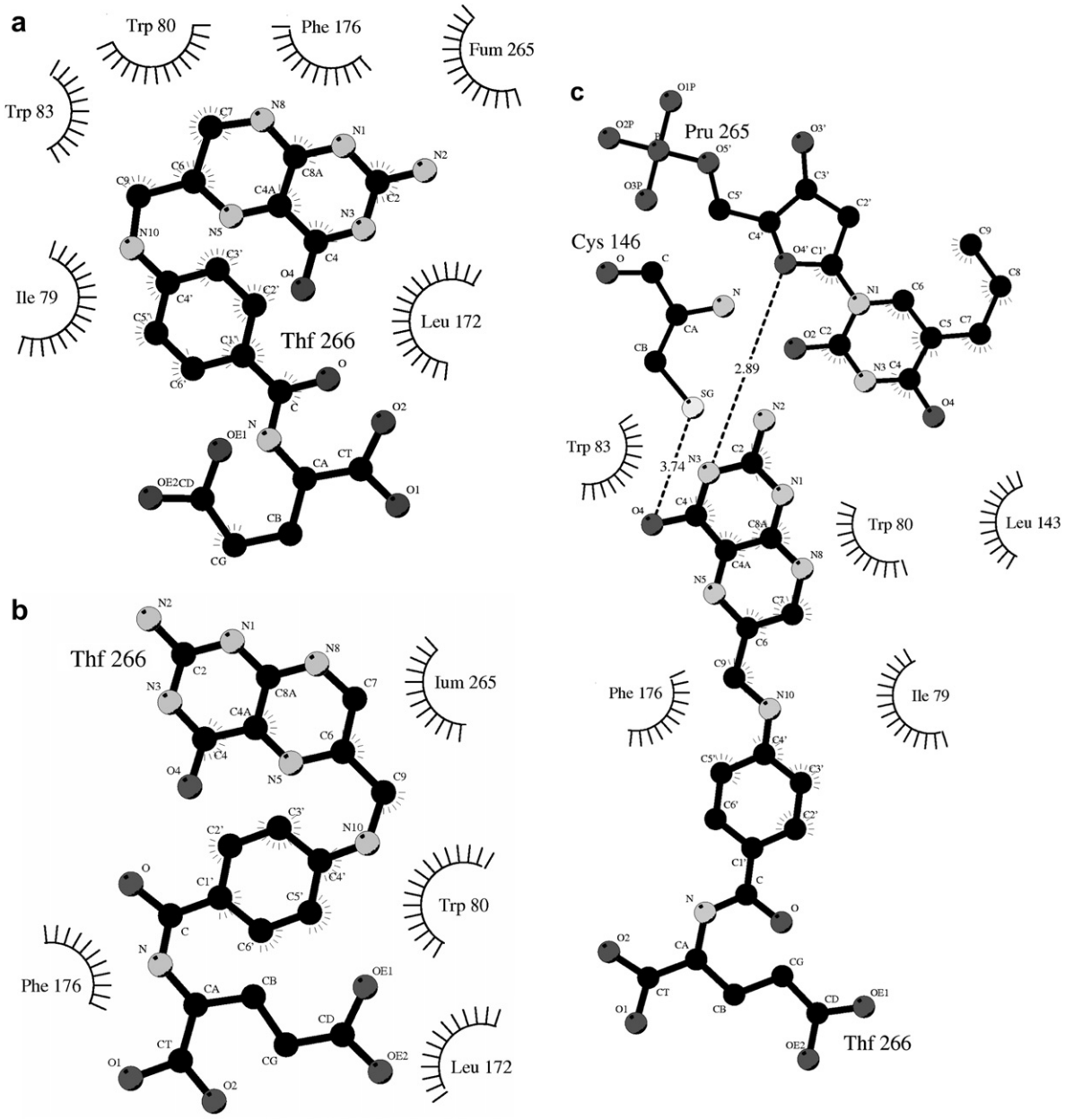

Figure 7. LIGPLOT drawings of THF–nucleotide and THF–protein contacts in the modeled ternary complexes, (a) TS–5-fluoro-dUMP–THF, (b) TS–5-iodo-dUMP–THF and (c) TS–5-propyl-dUMP–THF. Hydrophobic contacts are indicated by spoked arcs. Hydrogen contacts to THF in (c) are depicted with dashed lines. Fum 265, Ium 265 and Pru 265 denote, respectively, 5-fluoro-dUMP, 5-iodo-dUMP, and 5-propyl-dUMP. Hydrogens are omitted for clarity.

Cys 146 is not only unusually distant from C(6) (*cf.* Table 4) but also misoriented. Its sulfur atom is located at a hydrogen bonding distance to O(4) of the cofactor pterin ring, rather than facing the pyrimidine ring (Fig. 7c).

3. Conclusions

Some of the C(5) analogues of dUMP are potent inhibitors of thymidylate synthase. Their potency increases with electron-withdrawing character of the C(5) substituent and when their steric effect is smaller. Here, we present a 2D-QSAR model quantifying this dependence in a statistically relevant manner. We also present the results of 3D-molecular modeling, where we study the effect of C(5) substitution on the binding to TS of both dUMP analogues (in the binary complex) and the cofactor (in the ternary complex). We observe that the presence of C(5) substituents does not impair the binding of dUMP analogues in the binary complex, except of 5-propyl-dUMP, which contains the largest substituent of all studied here analogues. The consequences of varying substitution at C(5) become much more apparent upon the binding of the cofactor to the complexes formed between TS and each dUMP analogue. In the model of the TS–5-propyl-dUMP–THF complex, both the nucleotide and the cofactor are bound in a highly unfavorable manner from the catalysis point of view, which offers a rationale for a lack of activity of 5-propyl-dUMP as a mechanism-based inhibitor of TS.[31] In

the rest of the models of ternary complexes, the binding of the cofactor is either (i) virtually unaltered relative to the TS–dUMP–THF complex, or (ii) altered through a decrease in the level of stabilization of the THF molecule at the binding site, or (iii) altered as indicated in (ii) as well as through a change in the orientation of the THF molecule. In general, the cases (i), (ii) and (iii) correspond to the most, medium, and least active C(5) analogues of dUMP in that group, respectively.

Our results show that relatively small changes in the ligand-binding characteristics may result in considerable differences in the affinity of C(5) analogues of dUMP to TS. In the case of substituent comprising three carbon units (propyl), the impact of the steric effect is detrimental to the binding to such extent that it dramatically lowers the affinity of 5-propyl-dUMP. These data suggest that correspondingly unfavorable binding characteristics would occur in other cases where a substituent at C(5) shows similar or stronger steric influence than the propyl group. The experimental data that conform to such reasoning, such as low affinities for TS shown by dUMP analogues with large substituents at C(5), are frequent.[26,31–34] On the other hand, among the C(5) analogues of dUMP characterized by the presence of substituents with lower steric effects, quite a number showed good inhibitory activities against TS,[18,20–31] but none, apart from FdUMP, was as yet successful in terms of anti-TS chemotherapy. In view of this, and given our results, we think that the odds for success in searching for new C(5) analogues of dUMP that would turn out to be good mechanism-based inhibitors against TS and future chemotherapeutic agents are, in general, not high. One of the premises for such prediction is that a desired combination of substituent size and electronic effects, that is, a significantly lower steric effect compared to the propyl group (necessary in view of the results of molecular modeling) and a considerable electron-withdrawing influence, respectively, will not be frequent in the potential pool of C(5) analogues with no tested affinities for TS. Nevertheless, we can think of a few new candidate analogues, containing at C(5) substituents characterized by promising combination of effects, such as especially nitroso (–NO), nitrate ($-ONO_2$), and fluorosulfonyl ($-SO_2F$). Their predicted affinities reflected by $\log[1/K_i^{app}]$ calculated from QSAR Eq. 1 are 3.59, 1.99, and 2.79, respectively, all the values pointing to the possibility of a very strong inhibition (compare with Table 1). Those potential dUMP analogues and strong TS inhibitors might thus be worth future investigation.

Recently, several newly developed approaches related to the TS-targeted anticancer research have been reported, such as the enzyme-catalyzed therapeutic activation (ECTA)[11,53,54] or approaches devising miscellaneous FdUMP or 5-fluorouracil pro-drugs, including the one utilizing FdUMP[N] oligodeoxynucleotides (ODNs).[55] While those approaches are beyond the scope of this paper, they address the deficiencies in the treatment by 5-fluorouracil (see Introduction) and thus are worth to note here as promising strategies for the research utilizing dUMP analogues as novel chemotherapeutic agents targeted against TS.

4. Methods

4.1. QSAR

Binding affinities to thymidylate synthase of a series of dUMP analogues, C(5)-substituted with $-CH_3$, $-C_2H_5$, $-C_3H_7$, –CHO, $-CH_2OH$, –OH, $-NO_2$, –F, –Cl, –Br, and –I, determined as K_i^{app} values describing inhibition of the reaction of L1210 TS-catalyzed tritium release from [C(5)–^{3}H]dUMP in the presence of the cofactor, result from previous studies performed in this laboratory,[31,56] with the exception of the K_i^{app} values for 5-nitro-dUMP and 5-formyl-dUMP that were collected from the literature.[13] No other C(5) analogues of dUMP, for which both K_i values for inhibition of L1210 TS and established substituent parameters would exist, were available.

A classical QSAR was developed with the *BuildQSAR* program[57] using the multiple linear regression analysis. The dependent variable was $\log[1/K_i^{app}]$. The databank of physicochemical descriptors (independent variables), supplied with *BuildQSAR*, was supplemented with additional parameters from the vdWaterbeemd data set[41] and then loaded to the program. The model pre-selection was performed using the genetic algorithm method as implemented in *BuildQSAR*. Since the observation number n was only 11, we opted for two-descriptor models, in order to keep the observation/descriptor ratio higher than 5:1, as is necessary to consider a QSAR model statistically relevant.[58] A few pre-selected models, characterized by highest correlation coefficients r (or goodnesses of fit r^2) and lowest standard deviations s, were further analyzed, with non-collinearity and the significance level P serving as additional selectors. The best QSAR equation, selected in accordance with those criteria, was inspected and validated using the 'leave-one-out' technique (results not shown). An internal estimate of the predicting ability of the final QSAR is given by the goodness of prediction q^2, based on the standard error of predictions SPRESS.

The descriptors used in the QSAR equation are the molar refractivity M_R[43] and the σ_p^o electronic parameter.[41] These descriptors are discussed in detail, along with their application, elsewhere.[59,60] Briefly, M_R is a measure of overall bulk and polarizability; it can be used for a substituent or for the whole molecule. The σ_p^o parameter is a Hammett *para* substituent constant reported in the unified sigma-zero scale by Sjöström and Wold (SW). The SW scale is the result of statistical processing of data for a large number of different reaction series, thus being of a general nature, that is, not referred to a particular system.

4.2. Molecular dynamics (MD)

Simulations were initiated from the crystallographic structure of the binary complex of *Escherichia coli* TS with dUMP (reference code 1bid)[36] from the Protein Data Bank. Crystallographic waters were discarded. Hydrogen atom positions were generated with the *LEAP* module of the AMBER 6.0 molecular modeling

software package.[61] The phosphate group of dUMP was kept fully deprotonated. Its charge of −2 was counterbalanced by two +1 charges placed on the arginine residues (Arg 21 and Arg 166) hydrogen-bonded to the phosphate. The following groups, –H, $–CH_3$, $–C_2H_5$, $–C_3H_7$, $–CH_2OH$, –F, –Cl, –Br, and –I, were inserted at C(5) of dUMP, resulting in the 9 systems for MD simulations. Each system was solvated in a cubic box of TIP3P water molecules[62] extended up to 9 Å away from the solute, which resulted in 28574–28607 molecules per simulation (dependent on the system). Eleven sodium counterions were added per system to neutralize the net charge of the protein.

All systems were energy-minimized using the mixed steepest-descent and conjugate-gradient minimization procedures, and then gradually heated to 300 K in three 7 ps intervals (0 K → 100 K, 100 K → 200 K, and 200 K → 300 K). Next, MD equilibrations at 300 K for 99 ps, followed by data production runs for 1000 ps, were carried out.

All simulations were performed using the SANDER module of AMBER. The Cornell et al.[63] all-atom force field parameters and charges were used. Simulations were performed in the NpT ensemble with a constant temperature of 300 K and a constant pressure of 1 atm. The temperature and pressure coupling parameters were both 1 ps. Periodic boundary conditions and an atom-based cutoff of 8 Å for non-bonded van der Waals interactions were applied. The long-range electrostatic interactions were calculated using the particle-mesh Ewald (PME) procedure.[64,65] A 1 fs time step and the SHAKE algorithm[66] to constrain bonds to hydrogen atoms were used. A total of 1000 snapshots were saved during each data production run, one snapshot per each 1 ps of molecular dynamics simulation.

Missing force field parameters for the dUMP analogues were obtained using available crystallographic data and by performing additional *ab initio* RHF/6-31G* geometry optimizations using the Gaussian98 program.[67] The partial atomic charges for the dUMP analogues were derived from electrostatic potentials, calculated at the RHF/6-31G* theory level, using restrained electrostatic potential (RESP) fitting method.[68,69] The RESP charges, the additional force field parameters as well as force field atom type assignment for the dUMP analogues used in this work are available from the authors upon request.

4.3. Analysis and post-processing of the data from MD trajectories

A basic statistical analysis of the MD trajectories was performed with the *ptraj* module of AMBER. Root-mean-square-deviations (rmsds) were calculated to assess the conformational stability during the simulation courses. The conformational flexibility was analyzed in terms of root-mean-square-fluctuations (rmsfs). Computing the average structures from simulations was performed in two steps. First, water and counterions were stripped and the positions of remaining atoms in the structure were time-averaged over the 1000 ps of the data production run. Second, the average coordinates were briefly minimized with *SANDER* to ensure a reasonable local geometry. Once computed, each average structure was (i) analyzed using a visual inspection, and (ii) superimposed by the least-squares fitting to the crystal structure of the ternary complex of *E. coli* TS with dUMP and THF (PDB reference code: 1kzi).[38] Each superimposition was performed over skeletal atoms of the pyrimidine ring using the *DS Visualizer* program.[70] The cofactor-binding site residues in the average structure from MD simulation and the molecule of THF from the crystal structure were optimized (using a fast, Dreiding-like force field as implemented in *DS Visualizer*) to allow a correlated reorientation and relieve minor steric clashes. The superimpositions were then inspected with emphasis on the binding of THF against each C(5)-analogue of dUMP. Here it should be noted that the superimposition approach used in this work capitalizes on a proven similarity between the binding modes of dUMP in the binary and ternary TS complexes,[71] as well as on a nearly identical binding to *E. coli* TS of the molecules of THF (Fig. 2) and the cofactor mTHF[38] (differing from THF by the presence of a five-membered ring closed via a methylene carbon bonded to N(10) and N(5)).

Acknowledgment

Part of computations was carried out at the Interdisciplinary Centre for Mathematical and Computational Modelling of Warsaw University (computational Grant to A.J.).

References and notes

1. Santi, D. V.; Danenberg, P. V. Folates in pyrimidine nucleotide biosynthesis. In *Folates and Pteridines*; Blakely, R. L., Benkovic, S. J., Eds.; Wiley: New York, 1984; Vol. 1, pp 345–398.
2. Carreras, C. W.; Santi, D. V. *Annu. Rev. Biochem.* **1995**, *64*, 721–762.
3. Stroud, R. M.; Finer-Moore, J. S. *Biochemistry* **2003**, *42*, 239–247.
4. Finer-Moore, J. S.; Santi, D. V.; Stroud, R. M. *Biochemistry* **2003**, *42*, 248–256.
5. Papamichael, D. *Stem Cells* **2000**, *18*, 166–175.
6. Chu, E.; Callender, M. A.; Farrell, M. P.; Schmitz, J. C. *Cancer Chemother. Pharmacol.* **2003**, *52*, S80–S89.
7. Grem, J. L. *Invest. New Drugs* **2000**, *18*, 299–313.
8. Beale, P.; Clark, S. Tomudex. Clinical Development. In *Anticancer Drug Development*; Jackman, A. L., Ed.; Humana Press, Inc: Totowa, NJ, 1999; pp 167–181.
9. Bavetsias, V.; Jackman, A. L. *Curr. Med. Chem.* **1998**, *5*, 265–288.
10. Jackman, A. L.; Theti, D. S.; Gibbs, D. D. *Adv. Drug Deliv. Rev.* **2004**, *56*, 1111–1125.
11. Sergeeva, M. V.; Cathers, B. E. *Biochem. Pharmacol.* **2003**, *65*, 823–831.
12. De Clercq, E.; Balzarini, J.; Torrence, P. F.; Mertes, M. P.; Schmidt, C. L.; Shugar, D.; Barr, P. J.; Jones, A. S.; Verhelst, G.; Walker, R. T. *Mol. Pharmacol.* **1981**, *19*, 321–330.

13. Balzarini, J.; De Clercq, E.; Mertes, M. P.; Shugar, D.; Torrence, P. F. *Biochem. Pharmacol.* **1982**, *31*, 3673–3682.
14. Balzarini, J.; De Clercq, E. *Biochim. Biophys. Acta* **1984**, *785*, 36–45.
15. Jarmuła, A.; Anulewicz, R.; Leś, A.; Cyrański, M. K.; Adamowicz, L.; Bretner, M.; Felczak, K.; Kulikowski, T.; Krygowski, T. M.; Rode, W. *Biochim. Biophys. Acta* **1998**, *1382*, 277–286.
16. Jarmuła, A.; Cyrański, M. K.; Leś, A.; Krygowski, T. M.; Rode, W. *Pol. J. Chem.* **1998**, *72*, 1958–1962.
17. Kalman, T. I.; Nie, Z. *Nucleosides Nucleotides* **2000**, *19*, 357–369.
18. Kalman, T. I.; Nie, Z.; Kamat, A. *Nucleosides Nucleotides Nucleic Acid* **2001**, *20*, 869–871.
19. Saxl, R. L.; Reston, J.; Nie, Z.; Kalman, T. I.; Maley, F. *Biochemistry* **2003**, *42*, 4544–4551.
20. Wataya, Y.; Santi, D. V.; Hansch, C. *J. Med. Chem.* **1977**, *20*, 1469–1473.
21. Parker, W. B.; Cheng, Y. C. *Pharmacol. Ther.* **1990**, *48*, 381–395.
22. Eckstein, J. W.; Foster, P. G.; Finer-Moore, J.; Wataya, Y.; Santi, D. V. *Biochemistry* **1994**, *33*, 15086–15094.
23. Matsuda, A.; Wataya, Y.; Santi, D. V. *Biochem. Biophys. Res. Commun.* **1978**, *84*, 654–659.
24. Wataya, Y.; Matsuda, A.; Santi, D. V. *J. Biol. Chem.* **1980**, *255*, 5538–5544.
25. Mertes, M. P.; Chang, C. T.; De Clercq, E.; Huang, G. F.; Torrence, P. F. *Biochem. Biophys. Res. Commun.* **1978**, *84*, 1054–1059.
26. Kampf, A.; Barfknecht, R. L.; Shaffer, P. J.; Osaki, S.; Mertes, M. P. *J. Med. Chem.* **1976**, *19*, 903–908.
27. Santi, D. V.; Sakai, T. T. *Biochem. Biophys. Res. Commun.* **1972**, *46*, 1320–1325.
28. Santi, D. V.; Sakai, T. T. *Biochem. Biophys. Res. Commun.* **1971**, *42*, 813–817.
29. Barr, P. J.; Nolan, P. A.; Santi, D. V.; Robins, M. J. *J. Med. Chem.* **1981**, *24*, 1385–1388.
30. Barr, P. J.; Robins, M. J.; Santi, D. V. *Biochemistry* **1983**, *22*, 1696–1703.
31. Rode, W.; Kulikowski, T.; Kędzierska, B.; Jastreboff, M.; Shugar, D. *Biochem. Pharmacol.* **1984**, *33*, 2699–2705.
32. Edelman, M. S.; Barfknecht, R. L.; Huet-Rose, R.; Boguslawski, S.; Mertes, M. P. *J. Med. Chem.* **1977**, *20*, 669–673.
33. Barfknecht, R. L.; Huet-Rose, R. A.; Kampf, A.; Mertes, M. P. *J. Am. Chem. Soc.* **1976**, *98*, 5041–5043.
34. Al-Razzak, L. A.; Schwepler, D.; Decedue, C. J.; Balzarini, J.; De Clercq, E. *J. Med. Chem.* **1987**, *30*, 409–419.
35. Jarmuła, A.; Leś, A.; Rode, W. *Bioorg. Chem.* **2000**, *28*, 156–162.
36. Stout, T. J.; Sage, C. R.; Stroud, R. M. *Structure* **1998**, *6*, 839–848.
37. Hyatt, D. C.; Maley, F.; Montfort, W. R. *Biochemistry* **1997**, *36*, 4585–4594.
38. Fritz, T. A.; Liu, L.; Finer-Moore, J. S.; Stroud, R. M. *Biochemistry* **2002**, *41*, 7021–7025.
39. Fauman, E. B.; Rutenber, E. E.; Maley, G. F.; Maley, F.; Stroud, R. M. *Biochemistry* **1994**, *33*, 1502–1511.
40. Copur, S.; Aiba, K.; Drake, J. C.; Allegra, C. J.; Chu, E. *Biochem. Pharmacol.* **1995**, *49*, 1419–1426.
41. van de Waterbeemd, H. http://www.ndsu.nodak.edu/qsar_soc/resource/datasets/vdw1.htm.
42. Charton, M. *Progr. Phys. Org. Chem.* **1971**, *8*, 235–317.
43. Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR: Hydrophobic, Electronic, and Steric Constants*; American Chemical Society: Washington, 1995.
44. Garrett, C.; Wataya, Y.; Santi, D. V. *Biochemistry* **1979**, *13*, 2798–2804.
45. Liu, L.; Santi, D. V. *Biochemistry* **1993**, *32*, 9263–9267.
46. Finer-Moore, J. S.; Liu, L.; Schafmeister, C. E.; Birdsall, D. L.; Mau, T.; Santi, D. V.; Stroud, R. M. *Biochemistry* **1996**, *35*, 5125–5136.
47. Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, 1984.
48. Finer-Moore, J.; Montfort, W. R.; Stroud, R. M. *Biochemistry* **1990**, *29*, 6977–6986.
49. Matthews, D. A.; Villafranca, J. E.; Janson, C. A.; Smith, W. W.; Welsh, K.; Freer, S. *J. Mol. Biol.* **1990**, *214*, 923–936.
50. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Eng.* **1995**, *8*, 127–134.
51. DeLano, W. L. The PyMOL Molecular Graphics System. DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org.
52. Montfort, W. R.; Perry, K. M.; Fauman, E. B.; Finer-Moore, J. S.; Maley, G. F.; Hardy, L.; Maley, F.; Stroud, R. M. *Biochemistry* **1990**, *29*, 6964–6977.
53. Sergeeva, O. A.; Khambatta, H. G.; Cathers, B. E.; Sergeeva, M. V. *Biochem. Biophys. Res. Commun.* **2003**, *307*, 297–300.
54. Costi, M. P.; Ferrari, S.; Venturelli, A.; Calo, S.; Tondi, D.; Barlocco, D. *Curr. Med. Chem.* **2005**, *12*, 2241–2258.
55. Gmeiner, W. H. *Curr. Med. Chem.* **2005**, *12*, 191–202.
56. Felczak, K.; Miazga, A.; Poznański, J.; Bretner, M.; Kulikowski, T.; Dzik, J. M.; Gołos, B.; Zieliński, Z.; Cieśla, J.; Rode, W. *J. Med. Chem.* **2000**, *43*, 4647–4656.
57. Oliveira, D. B.; Gaudio, A. C. *Quant. Struct.-Act. Relat.* **2000**, *19*, 599–601.
58. Cronin, M. T. D.; Schultz, T. W. *J. Mol. Struct. (THEOCHEM)* **2003**, *622*, 39–51.
59. Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, 1995.
60. Sjöström, M.; Wold, S. *Chem. Scr.* **1976**, *9*, 200–210.
61. Case, D. A.; Pearlman, D. A.; Caldwell J. W.; Cheatham, T. E.; Ross, W. S., III; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P.A. AMBER 6, 1999 University of California, San Francisco.
62. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
63. 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*, 5179–5197.
64. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
65. York, D. M.; Darden, T. A.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345–8348.
66. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
67. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.;

Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian 98 (Revision A.1), Gaussian, Inc., Pittsburgh, PA, 1998.
68. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
69. Cieplak, P.; Bayly, C. I.; Cornell, W. D.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1357–1377.
70. Discovery Studio, Accelrys, Inc., San Diego, CA, 2005. http://www.accelrys.com/dstudio/.
71. Finer-Moore, J. S.; Liu, L.; Birdsall, D. L.; Brem, R.; Apfeld, J.; Santi, D. V.; Stroud, R. M. *J. Mol. Biol.* **1998**, *276*, 113–129.