Linear Free Energy Relationship Studies of Enzyme Active Site Binding: Thymidylate Synthase

George Chang, Douglas Schwepler,† Charles J. Decedue,† Jan Balzarini,‡ Erik De Clercq,‡ and Mathias P. Mertes*

Department of Medicinal Chemistry, Biochemical Services Laboratory, University of Kansas, Lawrence, Kansas 66045, and Rega Institute for Medical Research, University of Leuven, B-3000 Leuven, Belgium. *Received April 17, 1987*

The requirements for active-site binding of thymidylate synthase from three sources, *Lactobacillus casei*, murine leukemia L1210, and human lymphoblast (Molt/4F), were investigated by analyzing the binding of a series of 5-(*p*-substituted phenyl)-2′-deoxyuridylates (N1-substituted 5-aryl-2,4-dioxopyrimidines) to the enzyme. Multiple regression analysis revealed that an increase in electron density of the heterocyclic ring and hydrophobic substituents enhance affinity. Correlations of biological results with spectral data indicated that higher electron densities at the oxygen atoms are responsible for increase in binding. These results support the presence of both a cationic binding site and a hydrophobic region. In addition, the results revealed an unusual reversal of electronic requirements for binding and catalysis. The formation of the binary complex is enhanced by electron-donating substituents, while the initial catalytic reaction, formation of the covalent ternary complex, is promoted and stabilized by electron-withdrawing substituents.

A quantitative description and analysis of the electronic properties of N1-substituted 2,4-dioxopyrimidines have applications in the biosynthesis, function, and metabolism of pyrimidine nucleotides and their polymeric products, RNA and DNA. A specific example is the biosynthesis of thymidine 5′-monophosphate, involving the enzyme thymidylate synthase (EC 2.1.1.45). This enzyme has been crystallized, and the general features of the catalytic sequence are known. Before the recent X-ray crystallographic data on the structure of thymidylate synthase were announced,[1] approaches to describing the chemical and physical properties of the active site were limited to qualitative methods and kinetic analysis. In an attempt to perform a more quantitative analysis, linear free energy relationship studies were pursued.

Two approaches were undertaken in this study. One involved correlation of inhibitory constants (K_i) with general physical parameters such as Hammett σ or Swain and Lupton F and R values,[2] while the other involved correlation with specific spectral data (NMR shifts and infrared carbonyl shifts). The advantages of the second approach include the ability to "dissect" the heterocyclic ring and examine the electron distribution at various regions. In contrast, the σ values only describe a general electronic effect, and thus a more detailed analysis of the electron distribution of the heterocycle is not possible. Furthermore, the spectroscopic data describe the present system, whereas the σ values are compiled from various different studies. From the relative importance of specific regions of the pyrimidine ring, as a function of active-site binding, the physical properties of the complementary surface in the active site may be deduced. The crux of this method lies in the fact that the spectroscopic data, as descriptors of electron distribution, should measure only inductive and resonance contributions. To this end, the spectroscopic data and σ values should be linearly correlated.

Previous studies of this nature have been hampered by complex interactions between the substituents and the heterocycle. As an example, Ellis and co-workers,[3] in studying the ^{13}C NMR of 5-substituted uracils, found no obvious correlation between substituent electronegativity and the chemical shifts at C5 and C6 of the pyrimidine ring. Instead, they considered the 5-substituted uracils as trisubstituted ethylenes and rationalized the data in terms of mesomeric acceptors and donors.

In an attempt to design an appropriate model system, the following rationale was taken. A series of models may be written in the form XGY, where X represents a substituent, Y represents the site at which the observed phenomenon takes place, and G is a skeletal group to which X and Y are attached (Table I). In the classical ionization constant study of substituted benzoic acids by Hammett, X is a set of substituents, G is a phenyl ring, and Y, the reactive site, is the carboxyl moiety. In an analogous manner, the entire 2,4-dioxopyrimidine ring can be considered as the reactive site Y, a phenyl ring as the skeletal group G, and a series of substituents X, attached on the phenyl ring.

The advantages in the proposed model system are as follows:

(1) The dissociation constant for the affinity of these models for the enzyme can be accurately measured because the insertion of the phenyl ring removes the substituents from the "reactive region" of the active site, thus avoiding reactions at the catalytic site of thymidylate synthase.

(2) The planarity of the two rings allows for the transmission of the electronic properties of the substituents to the pyrimidine ring.

(3) The insertion of the phenyl ring between the substituent and the heterocycle also should simplify the NMR studies. Shielding effects such as neighboring group anisotropic effect, steric effect, and heavy atom effect are not of concern.

(4) The presence of the phenyl ring should eliminate intramolecular interactions (i.e., hydrogen bonding) between the substituents and the carbonyl groups of the pyrimidine ring.

(5) Finally, the substituents can be chosen in the model to cover a wide range of σ and π values. A two-dimensional Craig plot of σ vs π is shown in Figure 1.

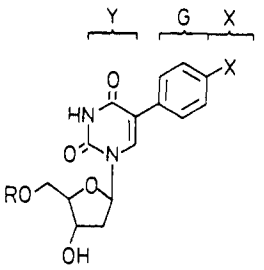
Quantitative and qualitative relationships between physical/spectral data and σ constants were observed in these models.[4] The results from this study revealed linear correlations between the biological data and the physical/spectral data. From the multiple regression analysis,

† Biochemical Services Laboratory.
‡ Rega Institute for Medical Research.

(1) (a) Tykarska, E.; Lebioda, L.; Bradshaw, T. P.; Dunlap, R. B. *J. Mol. Biol.* **1986**, *191*, 147-150. (b) Hardy, L. W.; Finer-Moore, J. S.; Montfort, W. R.; Jones, M. O.; Santi, D. V.; Stroud, R. M. *Science (Washington, D.C.)* **1987**, *235*, 448-455.
(2) Swain, C. G.; Lupton, E. C., Jr. *J. Am. Chem. Soc.* **1968**, *90*, 4328-4337.
(3) Ellis, P. D.; Dunlap, R. B.; Pollard, A. L.; Seidman, K.; Cardin, A. D. *J. Am. Chem. Soc.* **1973**, *95*, 4398-4403.
(4) Chang, G.; Mertes, M. P. *J. Org. Chem.* **1987**, *52*, 3625-3631.

0022-2623/88/1831-1141$01.50/0

Table I. Substituent Constants and Spectral Assignments for a Series of N1-Substituted 5-Aryl-2,4-dioxopyrimidines[4,7]



a: R = H
b: R = PO_3^{2-}

compd	X	π^a	σ^b	F^c	R^c	MR[d]	NMR[e] C6-H	NMR[e] N3-H	NMR[e] C6-C	IR:[f] C2=O
1	$N(CH_3)_2$	0.18	−0.83	0.10	−0.92	1.56	8.00	11.31	135.67	1697.6
2	NH_2	−1.23	−0.66	0.02	−0.68	0.54	7.94	11.26	135.32	1697.6
3	OH	−0.67	−0.37	0.29	−0.64	0.29	8.03	NA[g]	136.42	1699.5
4	OCH_3	−0.02	−0.27	0.26	−0.51	0.79	8.09	11.37	136.89	1713.0
5	$OCH_2C_6H_5$	1.66	−0.23	0.26	−0.51	3.22	8.10	11.42	136.95	1697.6
6	CH_3	0.56	−0.17	−0.04	−0.13	0.57	8.13	11.39	137.41	1711.1
7	H	0.00	0.00	0.00	0.00	0.10	8.18	11.43	137.94	1713.0
8	Br	0.86	0.23	0.44	−0.17	0.89	8.26	11.48	138.32	1711.1
9	$CONH_2$	−1.49	0.36	0.24	0.14	0.98	8.29	11.50	138.71	1699.5
10	CO_2CH_3	−0.01	0.45	0.33	0.15	1.29	8.40	11.53	139.34	1714.9
11	CF_3	0.88	0.54	0.38	0.19	0.50	8.39	NA[g]	139.36	1718.8
12	CN	−0.57	0.66	0.51	0.19	0.63	8.41	11.57	139.76	1718.8
13	NO_2	−0.28	0.78	0.67	0.16	0.74	8.51	11.61	140.40	1716.9

[a] π values represent hydrophobic interactions. [b] σ values represent electronic effect. [c] The terms F and R represent field-inductive and resonance effects, respectively. [d] MR values represent steric effects. The MR values were scaled by 0.1 to make them commensurable with the other parameters. [e] Chemical shifts are reported in parts per million and with respect to Me_4Si. [f] Carbonyl stretching bands are reported in wavenumbers (cm^{-1}). [g] Chemical shifts for these protons were not observed.

Table II. Inhibitory Effects of 5-Aryl-2′-deoxyuridylates on *L. casei*, Murine Leukemia (L1210), and Human Lymphoblast (Molt/4F) Thymidylate Synthase

		L. casei thymidylate synthase[a]			L1210 thymidylate synthase[b]		Molt/4F thymidylate synthase[c]	
compd	X	K_m, μM	K_i^{obs}, μM	$K_i^{calcd,d}$ μM	K_m, μM	K_i, μM	K_m, μM	K_i, μM
1b	$N(CH_3)_2$	4.50	1.41	0.62	1.28	48.0	1.51	66.3
2b	NH_2	4.47	1.13	1.40	1.27	11.3	1.40	15.4
3b	OH	4.80	7.62[e]		1.28	179.2[e]	1.54	134.1[e]
4b	OCH_3	5.79	1.21	1.51	1.27	33.0	1.48	39.4
5b	$OCH_2C_6H_5$	4.50	0.31	0.58	1.26	11.2	1.51	29.4
6b	CH_3	4.19	0.61	0.70	1.27	8.3	1.47	5.6
7b	H	4.00	0.96	1.20	1.28	12.2	1.34	6.8
8b	Br	4.00	1.81	1.98	1.27	29.2	1.51	21.7
9b	$CONH_2$	4.01	5.67	6.00	1.28	112.6	1.48	119.4
10b	CO_2CH_3	4.19	3.99	3.27	1.28	46.1	1.38	102.5
11b	CF_3	6.40	4.93	2.31	1.28	93.4	1.35	35.0
12b	CN	4.37	8.75	7.46	1.27	97.8	1.51	114.2
13b	NO_2	4.19	6.42	9.32	1.27	48.3	1.59	58.7

[a] Purified from amethopterin-resistant *L. casei*. [b] Partially purified from murine leukemia L1210 cells. [c] Partially purified from human cell culture (Molt/4F cells). [d] Values were calculated by using eq 5. [e] Values were not used in the derivation of eq 1–11.

it appears that higher electron densities in the heterocyclic ring enhance the formation of the initial noncovalent complex of enzyme model. In addition, the data support the presence of a cationic binding site and a hydrophobic region at the active site.

The K_i values for 13 N1-substituted 5-aryl-2,4-dioxopyrimidines, **1b**–**13b**, as assessed by their ability to inhibit thymidylate synthase catalyzed dTMP formation, are listed in Table II. Three different sources of enzyme were studied: *Lactobacillus casei*, murine leukemia L1210, and human lymphoblast (Molt/4F) thymidylate synthase. The inhibitory constants (K_i) for the *L. casei* enzyme in Table II were determined in a 30-s assay.

Quantitative Structure–Activity Relationships (QSAR). Multiple regression analyses[5] of the K_i values for inhibition of *L. casei*, murine leukemia L1210, and human lymphoblast (Molt/4F) thymidylate synthase were performed in an attempt to study the physicochemical properties of the active site of these enzymes. Correlation of the K_i values from the *L. casei* study with Hammett σ and Hansch π values yielded eq 1 as the single-variable equation and eq 2 as the two-variable equation.

$$\log 1/K_i = -0.66\ (\pm 0.41)\sigma - 0.25\ (\pm 0.20) \quad (1)$$

$n = 12$, $r = 0.752$, $s = 0.314$, $F_{1,10} =$ 13.02 (enzyme *L. casei*)

$$\log 1/K_i = -0.64\ (\pm 0.32)\sigma + 0.23\ (\pm 0.18)\pi - 0.26\ (\pm 0.16) \quad (2)$$

$n = 12$, $r = 0.878$, $s = 0.240$, $F_{1,9} =$ 8.06 (enzyme *L. casei*)

The Hammett σ constants can be separated into resonance and inductive-field components. A stepwise de-

(5) Hansch, C.; Leo, A.; Unger, S. H.; Kim, K. H.; Nikaitani, D.; Lien, E. J. *J. Med. Chem.* **1973**, *16*, 1207–1216.

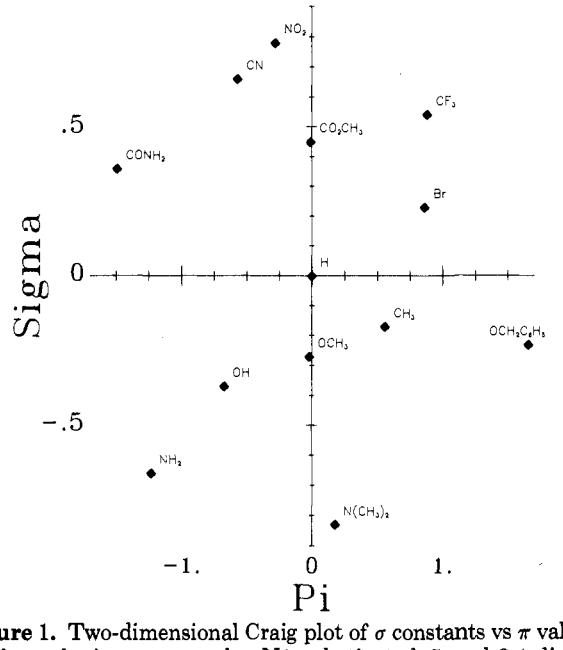


Figure 1. Two-dimensional Craig plot of σ constants vs π values of the substituents on the N1-substituted 5-aryl-2,4-dioxopyrimidines (1–13).

velopment of equations correlating inhibitory activity with Swain and Lupton F and R values[2] are shown in eq 3–5.

$$\log 1/K_i = -1.42\ (\pm 1.08)F + 0.07\ (\pm 0.36) \quad (3)$$

$n = 12$, $r = 0.680$, $s = 0.349$, $F_{1,10} =$ 8.61 (enzyme *L. casei*)

$$\log 1/K_i = -1.48\ (\pm 0.78)F + 0.28\ (\pm 0.19)\pi + 0.08\ (\pm 0.26) \quad (4)$$

$n = 12$, $r = 0.869$, $s = 0.249$, $F_{1,9} =$ 10.69 (enzyme *L. casei*)

$$\log 1/K_i = -1.13\ (\pm 0.82)F - 0.39\ (\pm 0.46)R + 0.25\ (\pm 0.17)\pi - 0.08\ (\pm 0.30) \quad (5)$$

$n = 12$, $r = 0.912$, $s = 0.218$, $F_{1,8} =$ 3.70 (enzyme *L. casei*)

The parenthesized values in these equations are the 95% confidence intervals; n represents the number of data points employed, r is the correlation coefficient, s is the standard deviation, and $F_{1,x}$ is the stepwise F statistic ($F_{1,10;\alpha=0.05} = 4.96$; $F_{1,9;\alpha=0.01} = 10.56$; $F_{1,8;\alpha=0.10} = 3.46$). To reduce the probability of chance correlation,[6] the number of data points ($n = 12$) permits the use of only three variables in an equation. Hence, evaluations of other parameters were made in comparison to eq 2. The addition of a steric term, MR, to eq 2 did not improve the correlation equation ($r = 0.878$), and the magnitude of the coefficient was small ($\rho = 0.01$). The use of an indicator variable, I, to differentiate H-donor and H-acceptor also failed to improve eq 2. Within the limits of the nucleotide samples tested, a linear relationship exists between affinity and substituent hydrophobicity; an improvement in correlation was not observed when a π^2 term was added. Other variables not listed were found insignificant in the correlation. The collinearity among significant variables is given in Table III. Ideally, the squared correlation coefficients should be less than 0.30.[7] The collinearities of F and R values with σ constants are expected since the former values represent the inductive and resonance components of the σ electronic effect. Calculated inhibitory constants for the *L. casei* enzyme by using eq 5 are listed in Table II.

Table III. Squared Correlation Matrix for Interrelationship of Variables in Eq 1–5

	σ	π	F	R
σ	1.00	0.00	0.59	0.88
π		1.00	0.00	0.02
F			1.00	0.26
R				1.00

Linear relationships were found between Hammett σ values and several spectral data of 5-aryl-2,4-dioxopyrimidine nucleosides.[4] Therefore, a series of equations correlating inhibitory constants with specific spectral data were developed. The regression analyses correlating K_i values and spectral data are shown in equations 6–9. The collinearity among these spectral data and σ values values is given in Table IV. As expected, high squared correlation coefficients are observed for the σ values and spectral data.

equation involving C6-H/NMR

$$\log 1/K_i = -1.99\ (\pm 1.10)\text{C6-H/NMR} - 0.21\ (\pm 0.20) \quad (6a)$$

$n = 12$, $r = 0.788$, $s = 0.293$, $F_{1,10} =$ 16.34 (enzyme *L.* casei)

$$\log 1/K_i = -1.95\ (\pm 0.75)\text{C6-H/NMR} + 0.24\ (\pm 0.15)\pi - 0.22\ (\pm 0.13) \quad (6b)$$

$n = 12$, $r = 0.920$, $s = 0.198$, $F_{1,9} =$ 12.96 (enzyme *L. casei*)

equation involving N3-H/NMR

$$\log 1/K_i = -3.12\ (\pm 2.10)\text{N3-H/NMR} - 0.23\ (\pm 0.22) \quad (7a)$$

$n = 12$, $r = 0.724$, $s = 0.329$, $F_{1,10} =$ 10.99 (enzyme *L. casei*)

$$\log 1/K_i = -3.20\ (\pm 1.47)\text{N3-H/NMR} + 0.27\ (\pm 0.17)\pi - 0.24\ (\pm 0.15) \quad (7b)$$

$n = 12$, $r = 0.900$, $s = 0.227$, $F_{1,9} =$ 12.06 (enzyme *L. casei*)

equation involving C6-C/NMR

$$\log 1/K_i = -0.21\ (\pm 0.13)\text{C6-C/NMR} - 0.29\ (\pm 0.20) \quad (8a)$$

$n = 12$, $r = 0.751$, $s = 0.315$, $F_{1,10} =$ 12.92 (enzyme *L. casei*)

$$\log 1/K_i = -0.21\ (\pm 0.10)\text{C6-C/NMR} + 0.24\ (\pm 0.18)\pi - 0.30\ (\pm 0.15) \quad (8b)$$

$n = 12$, $r = 0.890$, $s = 0.230$, $F_{1,9} =$ 9.77 (enzyme *L. casei*)

equation involving C2=O/IR

$$\log 1/K_i = -0.38\ (\pm 0.29)\text{C2=O/IR} - 0.37\ (\pm 0.25) \quad (9a)$$

$n = 11$, $r = 0.705$, $s = 0.338$, $F_{1,9} =$ 8.91 (enzyme *L. casei*)

$$\log 1/K_i = -0.36\ (\pm 0.27)\text{C2=O/IR} + 0.21\ (\pm 0.29)\pi - 0.40\ (\pm 0.23) \quad (9b)$$

$n = 11$, $r = 0.790$, $s = 0.310$, $F_{1,8} =$ 2.69 (enzyme *L. casei*)

(6) Topliss, C.; Costello, R. J. *J. Med. Chem.* **1972**, *15*, 1066–1068.
(7) Hansch, C.; Leo, A. *Substituent Constants For Correlation Analysis in Chemistry and Biology*; Wiley-Interscience: New York, 1979.

Table IV. Squared Correlation Matrix for Interrelationship of Variables in Eq 6–9

	σ	π	C6-H/NMR	N3-H/NMR	C6-C/NMR	C2=O/IR
σ	1.00	0.00	0.96	0.96	0.98	0.74
π		1.00	0.00	0.00	0.00	0.01
C6-H/NMR			1.00	0.97	0.98	0.70
N3-H/NMR				1.00	0.98	0.65
C6-C/NMR					1.00	0.75
C2=O/IR						1.00

For the *L. casei* thymidylate synthase study, eq 5 was the best three-variable correlation equation. The terms in this equation are justified by the *F* test, and the number of data points (four per variable) is sufficient for meaningful analysis. The K_i value for the phenol nucleotide **3b** was excluded from this and other regression equations, as it proved to be a deviate. The nucleotide **3b** was synthesized by catalytic reduction of the benzyloxy derivative **5b**. It is possible that decomposition of **3b** had occurred during the reduction reaction and/or purification.

Certain conclusions may be reached from eq 5 regarding the effects of the substituents on the inhibition of *L. casei* thymidylate synthase by 5-aryl-2′-deoxyuridylates. In terms of electronic effects, the negative coefficient associated with the variable *F* and *R* indicate that electronic withdrawal from the uracil heterocycle decreases inhibitory power. The large coefficient in the *F* term is an indication that the inductive contribution is the dominant electronic effect. The *R* term is not very significant since the addition of this term, though it increases the correlation coefficient of eq 5 to 0.912 and reduces the *F* statistics to the 90% *F*-distribution level. This is in accord with our study on the electronic properties of N1-substituted 5-aryl-2,4-dioxopyrimidines, which showed that the electronic distribution of the heterocycle is mainly controlled by inductive effects.[4] The positive coefficient of the third variable by eq 5, π suggests that lipophilic substituents enhance the binding of these analogues to the enzyme. The addition of a steric term, MR, to eq 2 did not improve the equation. Likewise, the use of MR, in place of the electronic term *R* in eq 5 resulted in a poorer correlation ($r = 0.899$) with no statistically significant improvement over eq 4. These results indicate that within this series of aryl nucleotides, interaction of the inhibitors with the enzyme appears to be independent of steric effects.

The high collinearity between Hammet σ values and spectral data of 5-aryl-2′-deoxyuridines permits the use of the latter to describe the electron distribution around the pyrimidine ring.[4] Due to the different methods of spectral analysis, the spectral data are not commensurable. Therefore, in the correlation equations involving spectral data, a comparison of the magnitude of the coefficients cannot be made. However, the signs of the coefficients are informative. The negative coefficient associated with the spectral data term in eq 6–8 indicate that nucleotides with chemical shifts of higher ppm values are less potent inhibitors. For the ^1H NMR of C6-H and N3-H and the ^{13}C NMR of C6-C, downfield shifts to higher ppm values reflect decreasing electron density. In essence, eq 6–8 indicate that decreasing electron density at C6-H, N3-H, and C6-C, respectively, deters binding.

Electron density values at the O2 and O4 oxygens, determined by direct measurement, are not available. However, the electron distribution at these oxygens can be inferred from infrared (IR) studies.[4] Multiple regression analysis of the K_i values with IR carbonyl bands of the C4=O gave no obvious correlation, thus it is not possible to determine the importance of the O4 oxygen in the binding of these analogues to the enzyme. A correlation of inhibitory constants with IR carbonyl shifts of C2=O yielded eq 9. Although the correlation coefficient of this equation ($r = 0.790$) is low, it is adequate to confidently assign a direction to the slope of the coefficient. A negative slope is associated with the carbonyl shift term, indicating that nucleotides with C2=O bands of lower frequency of absorption enhance the inhibitory power. The inductive effect reduces the length of the C=O bond, thus increasing its force constant and the frequency of absorption. Therefore, the less electron attracting substituent, resulting in greater electron density at O2, enhances the binding of these analogues to the enzyme. The positive π term in eq 6–9 suggests that hydrophobic substituents enhance binding. The consistance of the coefficient of the π term of eq 6–9 as well as eq 4 and 5 is an indication that the effect is real. The magnitude of the coefficient is low. It has recently been suggested by Recanatini and co-workers[8] that an uneven hydrophobic surface, resulting in subtle steric effects beyond the level of resolution of our study, may account for the low coefficient of the π term. The addition of a steric term or other variables to these equations was found to be statistically insignificant. Correlation with other spectral data also failed to produce statistically meaningful equations.

Multiple regression analysis of the K_i values for inhibition of murine leukemia (L1210) and human lymphoblast (Molt/4F) thymidylate synthase yielded eq 10 and 11. Unlike the regression equations obtained from the *L. casei* study, the correlation coefficient and *F* statistic values of these equations are too low to place much confidence in the correlation.

best two-variable equation for L1210 thymidylate synthase

$$\log 1/K_i = -1.21\ (\pm 0.95)F + 0.16\ (\pm 0.23)\pi - 1.20\ (\pm 0.32) \quad (10)$$

$n = 12$, $r = 0.728$, $s = 0.308$, $F_{1,9} =$ 2.48 (enzyme L1210)

best two-variable equation for Molt/4F thymidylate synthase

$$\log 1/K_i = -1.32\ (\pm 1.08)F - 0.18\ (\pm 0.26)\pi - 1.20\ (\pm 0.36) \quad (11)$$

$n = 12$, $r = 0.719$, $s = 0.3440$, $F_{1,9} =$ 2.41 (enzyme Molt/4F)

The electronic term *F* in the equations for the murine leukemia L1210 and human lymphoblast (Molt/4F) thymidylate synthase suggests that electron-donating substituents enhance binding. On comparison to the *L. casei* equations, a smaller dependence of lipophilicity is observed for both of these enzymes. Attempts to correlate these biological activities to the spectral data resulted in poor correlations.

Discussion

In a preliminary QSAR study on the inhibition of *L casei* thymidylate synthase by 5-substituted 2′-deoxyuridylates,

(8) Recanatini, M.; Klein, T.; Yang, C.-Z.; McClarin, J.; Langridge, R.; Hansch, C. *Mol. Pharmacol.* **1986**, *29*, 436–446.

Wataya and co-workers[9] concluded that electron withdrawal from the heterocyclic ring increases affinity for the enzyme. This electronic effect would increase the acidity of the N3-H of the pyrimidines and thus enhance their ability to act as hydrogen-bond donors. However, due to a large span of K_i values (>10^3), differences of the pK_a values of the N3-H cannot be the only significant factor in binding of these analogues. Ellis and co-workers,[3] in studying the ^{13}C NMR of 5-substituted uracils, suggested that electron withdrawal from the uracil ring would lower the π-electron density at the C6-position, thus making that carbon of the inhibitors more susceptible to nucleophilic attack. A steric parameter was included in the preliminary QSAR study, indicating that in the series tested the bulkier 5-substituents are detrimental to binding. The lack of a hydrophobic term, log P or π, suggests that hydrophobic effects do not play an important role in affinity of their analogues for the enzyme–cofactor complex.[9]

In this QSAR study on the inhibition of *L casei* thymidylate synthase by 5-aryl-2′-deoxyuridylates, strikingly different results were obtained. The K_i values ranged from 0.31 to 8.75 μM, which amounts to a difference in binding energy of ~2 kcal/mol. From eq 5, it is apparent that electron-donating groups enhance binding.

Equation 9 suggests that increasing electron density at the O2 oxygen atom enhances binding. The lack of correlation between inhibitory constants and the IR shifts of the C4 carbonyl limits any conclusions made concerning the importance of O4 on the binding of these analogues to the enzyme. However, it is not unreasonable to assume that the electronic effect of the substituents on the O2 oxygen affect the O4 oxygen in a similar manner.

Although the nature of the intermolecular interaction between the oxygen and its corresponding binding region of the enzyme is not known, hydrogen bonding interactions, or electrostatic interactions of the oxygen with a positively charged group of the enzyme, are possibilities. It should be noted that one or more arginine residues have been implicated in the active site of *L. casei* thymidylate synthase.[1,10,11] Fersht and co-workers have recently shown via protein engineering techniques that specific hydrogen bonding energy between uncharged parts of ligands and proteins is 0.5–1.5 kcal/mol.[12a] In another study, Bartlett and Marlowe have shown that the intrinsic binding energy arising from a highly specific hydrogen bonding interaction is 4.0 kcal/mol.[12b] In our study, a difference in binding energy of ~2 kcal/mol was observed, which favors a combination of hydrogen bonding and weak intermolecular electrostatic interactions. This difference in binding energy may be too large to be described solely by changes in hydrogen-bonding interaction. It should be added that the pK_a values of the compounds tested in the present study were around 9, thus at the pH of the assay (pH 6.8), these compounds remained uncharged in the bulk solvent.

The equations correlating K_i values and NMR data indicate that the ability of the N3 proton of the pyrimidines to act as hydrogen-bond donors is not important in the differential binding of the nucleotides examined. Furthermore, a decrease in the electron density at C6 also does not enhance binding. However, electron distribution at the C6 would play a role in the second step of the enzymatic reaction: covalent bond formation. At this stage, a lower electron density at C6 of the inhibitors would make that carbon more susceptible to nucleophilic attack from the enzyme, leading to a covalent complex. A recent study on the active site of thymidylate synthase, utilizing electron spin resonance, gave evidence that the effect of the electron-withdrawing substituents on binding is minimal for the binary complex formation but is instead exerted during the formation of the ternary complex.[13] The other important factor in inhibition of thymidine 5′-monophosphate formation by the aryl nucleotides examined is the lipophilicity of the substituents; increasing lipophilicity enhances binding. The size of the substituents examined does not appear to interfer with the affinity of the analogues for the enzyme.

The presence of a cationic binding site on the enzyme has been postulated. Kalman and Bardos[14] determined a K_i value of 4×10^{-8} M for 5-mercapto-2′-deoxyuridylate (MUdRP). The pK_a for the ionization of the N3 proton of 5-mercaptopyrimidines[15] is 10.5–10.6, while the sulfhydryl group (pK_a 5.34) is essentially ionized at physiological pH.[14] The high affinity of this compound for the enzyme could not be explained in terms of the acidity of the N3 proton. Rather, it was postulated that the thiolate anion of MUdRP interacts with a cationic binding site on the enzyme.[14]

The interaction of 5-nitro-2′-deoxyuridylate (NO_2dUMP) with thymidylate synthase also suggests the presence of a cationic binding site.[16,17] A K_i value of 0.5×10^{-6} M has been determined for NO_2dUMP.[16] The pK_a of the N3 proton (6.5–6.6)[16] is acidic and capable of hydrogen bonding. However, at the pH of the enzyme assay (pH 6.8), the majority of NO_2dUMP would be ionized. Wataya and co-workers[17] proposed that the anionic form of NO_2dUMP is responsible for the high affinity of this compound. As the anion, the electron density is highly localized at the oxygen atoms and would be stabilized by positively charged amino acid residues of the enzyme.

5-Fluoro-2′-deoxyuridylate (FdUMP) has been shown to have low affinity for thymidylate synthase in the absence of cofactor.[18,19] The N3 proton of FdUMP has a pK_a of 7.8, thus at the pH of the enzyme assay (pH 7.4), FdUMP would only be partially ionized. Furthermore, fluorine is a strong inductive substituent (σ_I = 0.52).[7] Under these conditions, high localization of electron density at the O2 and O4 oxygen atoms would not be expected.

It is of interest to compare these results to the predicted computer graphics model of the complex of substrate and enzyme formulated by docking the substrate with the structure of thymidylate synthase as recently reported from X-ray crystallographic studies.[1b] The primary sites of active-site binding in the reported model involve the conserved amino acid residues Arg^{218} and $Arg^{179'}$ (from the other monomer of the active dimeric form), and the side-chain amides of Asn^{229} and Gln^{217}. The former protonated residues form an electrostatic bond to the 5′-

(9) Wataya, Y.; Santi, D. V.; Hansch, C. *J. Med. Chem.* **1977**, *20*, 1469–1473.
(10) Cipollo, K. L.; Dunlap, R. B. *Biochemistry* **1979**, *18*, 5537–5541.
(11) Belfort, M.; Maley, G. F.; Maley, F. *Arch. Biochem. Biophys.* **1980**, *204*, 340–349.
(12) (a) Fersht, A. R.; Shi, J.-P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M. Y.; Winter, G. *Nature (London)* **1985**, *314*, 235–238. (b) Bartlett, P. A.; Marlowe, C. K. *Science (Washington, D.C.)* **1987**, *235*, 569–571.
(13) Devanesan, P. D.; Bobst, A. M. *J. Med. Chem.* **1986**, *29*, 1237–1242.
(14) Kalman, T. I.; Bardos, T. J. *Mol. Pharmacol.* **1970**, *6*, 621–630.
(15) Bardos, T. J.; Kalman, T. I. *J. Pharm. Sci.* **1966**, *55*, 606–610.
(16) Maggiora, L.; Chang, C. T.-C.; Torrence, P. F.; Mertes, M. P. *J. Am. Chem. Soc.* **1981**, *103*, 3192–3198.
(17) Wataya, Y.; Matsuda, A.; Santi, D. V. *J. Biol. Chem.* **1980**, *255*, 5538–5566.
(18) Galivan, J. H.; Maley, G. F.; Maley, F. *Biochemistry* **1976**, *15*, 356–362.
(19) Santi, D. V. *J. Med. Chem.* **1980**, *23*, 103–111.

phosphate of the substrate, and the latter potentially could form a hydrogen bond to the N3-C4O portion of the pyrimidine ring. From the results of our studies it would appear that the N3-H is not involved in hydrogen bonding to the enzyme. Rather, the difference in binding energy within our series of compounds favors a combination of hydrogen bonding and electrostatic interactions between the high electron density site on the O2 oxygen and a positively charged site on the enzyme. This would also explain the high inhibitory constants noted for the ionic 5-nitro and the 5-mercapto derivatives of the substrate. It was noted from the X-ray crystal structure that the positively charged guanidinium group of $Arg^{178'}$ is located such that it could rotate to within 7 Å of the catalytic thiol, which may also put the charge in a position to form a hydrogen/electrostatic bond with one of the electron-rich carbonyl oxygens of the pyrimidine ring. Since it is recognized that a significant conformational change precludes cofactor binding, it is not unreasonable to postulate that substrate binding also induces a change in the tertiary structure that repositions a guanidinium cation of an arginine residue.

Multiple regression analysis of the K_i values for the inhibition of murine leukemia L1210 and human lymphoblast (Molt/4F) thymidylate synthase yielded correlation equations with electronic terms similar to those obtained from the *L. casei* thymidylate synthase study, suggesting that the electronic requirements for the binding of 5-aryl-2′-deoxyuridylates to these enzymes are essentially the same. However, the active-site binding of the enzymes are different. The K_i values for the inhibition of L1210 (8–113 μM) and Molt/4F (6–119 μM) thymidylate synthase are generally 10–20 times larger than those of *L. casei* (0.3–9 μM). Conversely, a lower K_m value is observed for the former two enzymes (K_m = 1.3 and 1.5 μM, respectively) than for the latter (K_m = 4.5 μM). In addition, the regression analysis equation for both of these enzymes suggest that there is a lower dependence of lipophilicity compared to the *L. casei* system. However, it should be noted that the *F* statistic values of these two equations are too low to place much confidence in the correlations.

Quantitative structure–activity relationships for quinazolines as inhibitors of thymidylate synthase from different sources have been studied.[20] Multivariate statistics were used to develop the correlation equations. The inhibition of bacterial thymidylate synthase (*L. casei*) by quinazolines yielded correlation equations that were significantly different from those for the same enzyme from mammalian sources (mouse leukemia L1210S).[20] This finding, which suggests that the two enzymes are dissimilar, supports the results of the present study.

In the future design and synthesis of inhibitors of thymidylate synthase, it may be worthwhile to consider the electron distribution of the pyrimidine ring. As suggested in the present study, an increase in electron density at the oxygen(s) of the heterocycle enhances the initial binding to give the noncovalent complex. However, previous studies have shown that a decrease in electron density at the C6 position would enhance covalent complex formation in the second step of the enzymatic reaction.[3,13] Clearly, opposing electronic requirements exist: high charge density for the first step and low charge density for the second step in the catalytic reaction. Two approaches to this electronic dichotomy are presented. One can envision a mechanism-based inhibitor with an electron-donating group at the C5 position of the pyrimidine ring, which would enhance the formation of the noncovalent complex. After this initial association, the C5 substituent would be oxidized to give an electron-withdrawing substituent, which would enhance formation of the covalent enzyme–inhibitor complex. Such studies are currently in progress in these laboratories in the synthesis of C5 substituted hydroquinone derivatives of the substrate. In the ternary complex of reduced cofactor (H_2folate), the hydroquinone inhibitor, and enzyme, the inhibitor could be oxidized to the corresponding quinone, which is known to rapidly inactivate thymidylate synthase.[21]

A second possibility would be to utilize a highly acidic inhibitor, which at physiological pH would result in an ionic pyrimidine ring. The presence of electron-withdrawing groups at the C5 position would increase the acidity of the heterocycle. However, the electronic effects of the substituent must be sufficient to lower the pK_a of the N3 proton below the pH of the enzyme assay (pH 6.8–7.4). Under these conditions a greater proportion of the inhibitor would be ionized, resulting in high localization of electron density at the oxygen atoms of the pyrimidine ring. This would satisfy the electronic requirements for binding to a cationic region in the active site of the enzyme. The presence of the electron-withdrawing group also would decrease the electron density at the C6 carbon, thus satisfying the electronic requirements for covalent complex formation and assist in delocalization of the resulting charge. These two factors have been demonstrated with model reactions and in enzyme inhibition studies with 5-nitro-2′-deoxyuridine 5′-phosphate, which has a net dissociation constant of 10^{-10} partitioned as 10^{-7} for the initial noncovalent complex and 10^{-3} for cleavage of the covalent bond formed in the second step of the reaction.[16]

Experimental Section

Electrophoresis experiments were performed on a Hormuth Pherograph Model MINI 68. 5′-Nucleotidase and 3′-nucleotidase were purchased from Sigma Chemical Co. Ion-exchange chromatography was performed with DEAE-Sephadex A25 (stationary phase) in a K15 column (length, 300 mm; i.d., 15 mm); both were products of Pharmacia.

The preparation of the nucleosides used in this study was reported.[4] Tetrahydrofolic acid (H_4folate) was purchased from Sigma Chemical Co. in 65% purity in 50-mg ampules. A stock solution was prepared by mixing the crude H_4folate with a suspension containing 2-mercaptoethanol (0.5 mL of 10 M solution) and sodium bicarbonate (15 mg). The resulting mixture was deoxygenated by bubbling nitrogen through the suspension. After centrifugation the concentration of H_4folate was determined in the supernatant by spectrophotometric analysis based on the λ_{max} at 296 nm (ϵ 25 000). The substrate [5-^{3}H]-2′-deoxyuridine 5′-phosphate at a specific activity of 21 Ci/mmol was purchased from Moravek Biochemicals, Industry, CA, and diluted with cold substrate purchased from Sigma Chemical Co. St. Louis, MO, to give a specific activity of 500 μCi/μmol. The studies with the murine leukemia (L1210) thymidylate synthase were performed with partially purified enzyme.[22]

Thymidylate synthase as a crude extract from methotrexate-resistant *L. casei* was purchased from the New England Enzyme Center as a dialyzed ammonium sulfate concentrate. The enzyme was purified and crystallized according to the method of Maley and co-workers[23] to give, after activation by dialysis for 24 h at 4 °C against 0.1 M potassium phosphate buffer at pH 6.8 containing 50 mM 2-mercaptoethanol, a preparation with a specific

(20) Chen, B. K.; Horvath, C.; Bertino, J. R. *J. Med. Chem.* **1979**, *22*, 483–491.

(21) Al-Razzak, L. A.; Schwepler, D.; Decedue, C. J.; Balzarini, J.; DeClercq, E.; Mertes, M. P. *J. Med. Chem.* **1987**, *30*, 409–419.

(22) Balzarini, J.; DeClercq, E.; Mertes, M. P.; Shugar, D.; Torrence, P. F. *Biochem. Pharmacol.* **1982**, *31*, 3673–3682.

(23) Galivan, J. H.; Maley, G. F.; Maley, F. *Biochemistry* **1975**, *14*, 3338–3344.

activity of 2.6–3.2 μM of product mg^{-1} min^{-1} based on the radioisotope assay. The enzyme concentration was determined by absorbance at 278 nm by using a molar absorptivity of 1.05×10^6 and a dimer molecular weight of 70 000.

General Procedure for the Phosphorylation of 5-Aryl-2′-deoxyuridine 1b, 4b–9b, 11b–13b. The nucleoside[4] (0.05–0.10 mmol) was dissolved in 0.5 mL of freshly distilled trimethyl phosphate and cooled to 0 °C. To this stirred solution was added 2 equiv of freshly distilled phosphorus oxychloride. After the mixture was stirred for 2 h at 0 °C, a second aliquot of phosphorus oxychloride (4 equiv) was added, followed by a third aliquot (2 equiv) after 6 h. The resulting mixture was allowed to stir at 0 °C for 8 h. The reaction was neutralized with 2.0 M triethylammonium bicarbonate (pH 7.8) and concentrated under vacuum. The residue was dissolved in water and resolved on DEAE-Sephadex A-25 with a linear gradient composed of 500 mL each of 0.0–0.6 M triethylammonium bicarbonate (pH 7.8). The isolated yields ranged from 25% to 81%. The products were identified with electrophoresis, UV, and ^{31}P NMR analyses.

5-(*p*-Aminophenyl)-2′-deoxyuridine 5′-Phosphate (2b). 5-(*p*-Nitrophenyl)-2′-deoxyuridine 5′-phosphate (**13b**, 0.04 mmol) was dissolved in 10 mL of water. To this was added 5 drops of concentrated HCl and 10 mg of 5% palladium on carbon. After being shaken under 40 psi of hydrogen for 5 h, the mixture was made basic with methanolic ammonia and filtered through Celite. The filtrate was concentrated and purified on DEAE-Sephadex A-25 with a linear gradient of triethylammonium bicarbonate (500 mL each of 0.0–0.6 M, pH 7.8) to afford compound **2b**.

5-(*p*-Hydroxyphenyl)-2′-deoxyuridine 5′-Phosphate (3b). 5-[*p*-(Benzyloxy)phenyl]-2′-deoxyuridine 5′-phosphate (**5b**, 0.035 mmol) was dissolved in 10 mL of water. To this was added 10 mg of 5% palladium on carbon. After being shaken under 45 psi of hydrogen for 10 h, the mixture was filtered through Celite. The filtrate was concentrated and purified on DEAE-Sephadex A-25 with a linear gradient of triethylammonium bicarbonate (500 mL each of 0.0–0.6 M, pH 7.8) to afford compound **3b**.

5-[*p*-(Methoxycarbonyl)phenyl]-2′-deoxyuridine 5′-Phosphate (10b). Compound **10b** was obtained via the general procedure for the phosphorylation of 5-aryl-2′-deoxyuridine, with the exception that the reaction was neutralized with anhydrous triethylamine.

***L. casei* Thymidylate Synthase Enzyme Assay.** The enzyme was assayed by using the radioisotope assay described.[24] The solution (0.1 mL) contained 35 mM mercaptoethanol, 0.22 mM *dl*-tetrahydrofolic acid, 6.75 mM formaldehyde, 1.2 mM sodium bicarbonate, 3 mM magnesium chloride, 0.12 mM EDTA, 6 mM potassium phosphate buffer, pH 6.8, and 5 μL of the diluted enzyme solution, substrate, and when indicated, inhibitor. Control reactions lacked the cofactor, tetrahydrofolic acid. The substrate 2′-deoxy-[5-^{3}H]uridine 5′-phosphate was used at a specific activity of 500 μCi/μmol. For maximum velocity studies for active enzyme determination, the substrate concentration was 50 μM. The assays were started by the addition of the enzyme to the complete mixture and then incubated at 30 °C. Incubation was stopped at 30 s by the addition of 50 μL of 20% trichloroacetic acid. A 20% aqueous suspension of charcoal (0.25 mL) was added, and the solution was vortexed and allowed to stand for 15 min. The suspension was filtered through a cotton-plugged Pasteur pipet, and 0.1 mL of the filtrate was counted in Beckman Ready Solve HP scintillation fluid. Counting efficiency was 45%; control samples lacking the cofactor were found to have less than 5% of the respective sample counts.

Inhibitory constants (K_i) were determined by duplicate assays by using in each two or more concentrations of inhibitor together with five concentrations of substrate. The K_i values were determined by a slope replot of the values obtained in double-reciprocal plots of velocity vs substrate concentrations.

Murine Leukemia L1210 and Human Lymphoblast (Molt/4F) Thymidylate Synthase Enzyme Assay. The enzyme was prepared and partially purified according to the procedure reported by Balzarini et al.[22] The assay mixture contained 0.26 mM tetrahydrofolic acid, 5.0 mM formaldehyde, 15 mM 2-mercaptoethanol, 0.1 M sodium fluoride, 1.2–16 μM [5-^{3}H]dUMP of 0.25 Ci, and an appropriate amount of the inhibitor in a total volume of 30 μL in 0.05 M potassium phosphate buffer at pH 7.5. The reaction was started by addition of 10 μL of the enzyme, and the resulting mixture was incubated at 37 °C for 30 min. The reaction was stopped by addition of 160 μL of a charcoal suspension (100 mg/mL in 2% trichloroacetic acid). After centrifugation for 10 min at 1000*g*, 0.1 mL of the supernatant was assayed for radioactivity in a toluene-base scintillant.

Acknowledgment. This work was supported by Grant CA 7522 from the National Cancer Institute of the National Institutes of Health, Grant 3.0040.83 from the Belgian Fonds voor Geneeskundig Wetenschappelijk Onderzoek, and Grant 85/90-79 from the Belgian Geconcerteerde Onderzoeksacties. We gratefully acknowledge Dr. Corwin Hansch for constructive criticisms and comments.

Registry No. **1a**, 108664-86-4; **1b**, 113997-13-0; **2a**, 108664-87-5; **2b**, 113997-14-1; **3a**, 108664-88-6; **3b**, 113997-15-2; **4a**, 89647-11-0; **4b**, 113997-16-3; **5a**, 108664-89-7; **5b**, 113997-17-4; **6a**, 92510-80-0; **6b**, 113997-18-5; **7a**, 76756-28-0; **7b**, 76756-31-5; **8a**, 108664-90-0; **8b**, 113997-19-6; **9a**, 108664-91-1; **9b**, 113997-20-9; **10a**, 108664-92-2; **10b**, 113997-21-0; **11a**, 108664-93-3; **11b**, 113997-22-1; **12a**, 92524-53-3; **12b**, 113997-23-2; **13a**, 108664-94-4; **13b**, 113997-24-3; thymidylate synthetase, 9031-61-2.

(24) Maggiora, L.; Chang, C. T.-C.; Hasson, M. E.; Bigge, C. F.; Mertes, M. P. *J. Med. Chem.* **1983**, *26*, 1028–1036.