and cell cycle kinetic parameters.

In Vivo Murine Tumor Models. The murine tumors used for in vivo evaluation are currently passaged in our laboratories and were obtained either from the National Cancer Institute (NCI), Bethesda, MD, or from Dr. Corbett, Wayne State University, Detroit, MI.

All tumors were maintained in the mouse strain of origin: DBA/2 mice, P388, L1210 leukemias; C57Bl/6 mice, B16 melanoma, Glasgow osteogenic sarcoma,[24] colon adenocarcinoma 38,[25] Lewis lung carcinoma, M5076 reticulum cell sarcoma; C3H/He mice, mammary adenocarcinoma 16/C.[26] These tumors were transplanted into the appropriate F1 hybrid or the strain of origin for therapy trials. Mice were bred at IFFA-CREDO (Lyon, France) from strains obtained from Jackson Laboratories, Bar Harbor, ME, or Charles River, France. Mice were over 18 g at the start of therapy. They were supplied food and water ad libidum.

Solid Tumors. Mammary adenocarcinoma 16/C, B16 melanoma, Glasgow osteogenic sarcoma, and Colon adenocarcinoma 38 were implanted subcutaneously (sc) with 30–60-mg tumor fragments by trocar on day 0. B16 melanoma was also implanted as an intraperitoneal 10% brei. Lewis lung carcinoma was implanted intramuscularly (im) with 10^6 cells. M5076 reticulum cell sarcoma was implanted ip with 10^6 cells.

Chemotherapy was started within 1–6 days after tumor implantation. Subcutaneously implanted tumors were measured with a caliper. Tumor weights were calculated from 2-dimensional measurements: tumor weight (mg) = $(l \times w^2)/2$, where l and w are the tumor length and width, respectively. Tumor growth inhibition (T/C value in percent) was used for the determination of activity where T and C are respectively the median tumor weight of the treated and the control groups. A T/C equal or less than 42% is considered significant antitumor activity by the NCI; a T/C value <10% is considered to indicate high antitumor activity and is the level used by NCI to justify further development. In the case in which survival was the end point for the therapeutic effectiveness (B16 melanoma ip brei, M5076 ip), the following criteria for activity was used: increase in life span, % ILS = 100 × [(median day of death of the treated group) − (median day of death of the control group)]/(median day of death of control group). Finally the percentage of metastasis inhibition was also used with the Lewis lung carcinoma: the surviving mice were killed on day 19, the lungs removed and fixed in Fekete solution, and the macroscopic lung metastasis was counted. The percentage metastasis inhibition was evaluated as follows: T/C %, where T and C are respectively the mean number of metastasis of the treated and the control groups.

Leukemias. Viable P388 (10^6) and L1210 (10^5) leukemia cells were inoculated intraperitoneally on day 0. Treatment was started on day 1 by the ip route. Animal mortality was checked daily. The antitumor activity was evaluated as follows: percent increase in host life span ILS % = 100 × [(median day of death of the treated group) − (median day of death of the control group)]/(median day of death of the control group). The NCI criteria for activity are moderate activity P388 ILS > 25%, L1210 ILS > 20% and significant activity P388 ILS > 75%, L1210 ILS > 50%.

Drug Treatment. Compound **24** was administered iv or ip. It was dissolved in 5% glucose in distilled water and injected as a solution under 0.2 mL iv and 0.5 mL ip.

Acknowledgment. This investigation was supported by a contract with INSERM-CNAMTS. The assistance of G. François (Laboratoire de Pharmacologie et de Toxicologie Fondamentales) is gratefully acknowledged.

Supplementary Material Available: Tables showing iv and ip toxicity of compound **24** in BD2F1 mice and percentage repartition of the MCF-7 cells and HBL human melanoma cancer cells into cell cycle (3 pages). Ordering information is given on any current masthead page.

(24) Glasgow, L. A.; Crane, J. L., Jr.; Kern, E. R. *J. Natl. Cancer Inst.* **1978**, *60*, 659.
(25) Corbett, T. H.; Griswold, D. P., Jr.; Roberts, B. J.; Peckham, J. C.; Schabel, F. M., Jr. *Cancer* **1977**, *40*, 2660.
(26) Corbett, T. H.; Griswold, D. P., Jr.; Roberts, B. J.; Peckham, J. C.; Schabel, F. M., Jr. *Cancer Treat. Rep.* **1978**, *62*, 1471.

On the Optimization of Hydrophobic and Hydrophilic Substituent Interactions of 2,4-Diamino-5-(substituted-benzyl)pyrimidines with Dihydrofolate Reductase

Cynthia Dias Selassie,*[†] Ren-Li Li,[‡] Martin Poe,[§] and Corwin Hansch[†]

Department of Chemistry, Pomona College, Claremont, California 91711, Department of Medicinal Chemistry, Beijing University, Beijing, China, and Merck Institute for Therapeutic Research, Rahway, New Jersey 07065-0900.
Received April 23, 1990

The inhibition constants (K_{iapp}) were obtained from the action of 68 2,4-diamino-5-(substituted-benzyl)pyrimidines on dihydrofolate reductase from an *Escherichia coli* strain MB 1428. Subsequently, these results were used to formulate appropriate quantitative structure–activity relationships (QSAR). Once again these equations emphasize the paramount importance of steric/dispersion factors in enhancing antibacterial potency. Hydrophobicity also plays a role, albeit a minor one. Comparisons with the QSAR obtained versus prokaryotic dihydrofolate reductase (DHFR) demonstrate subtle differences in binding behavior between meta and para substituents which may be effectively maximized in the design of more efficacious and selective antibacterial agents. The bacterial and avian QSAR equations can be used to calculate selectivity indices for trimethoprim, tetroxoprim, and two other specially designed 2,4-diamino-5-(substituted-benzyl)pyrimidines.

In the design of more efficacious drugs, one of the most serious problems is minimizing toxicity to the host while maximizing it toward the pathogen. Adrien Albert has succinctly defined this goal as "selective toxicity".[1] In the past, one hoped to make major advances by synthesizing and testing as many compounds as possible while always relying on the "intuition" of the medicinal chemist. Although this classical approach to drug research has not been deserted, the glimmerings of the promised land of "rational drug design" loom near the horizon. Elegant techniques for the isolation of enzymes and receptors and their subsequent cloning to obtain reasonable quantities of enzymes for routine work auger well for the future. In the case of antibacterial agents, one hopes to be able to detect significant differences between the human drug

[†] Pomona College.
[‡] Beijing University.
[§] Merck Institute for Therapeutic Research.

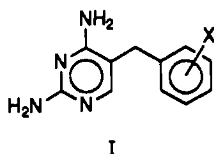
(1) Albert, A. *Selective Toxicity*, 7th ed.; Chapman and Hall: London, 1985; Chapter 4.

0022-2623/91/1834-0046$02.50/0 © 1991 American Chemical Society

receptor and that of the targeted pathogen. Excellent examples where valuable drugs have been designed from such a principle are dihydrofolate reductase (DHFR) inhibitors.[2] This enzyme, which is crucial for the biosynthesis of DNA, demonstrates considerable species variation, which opens up possibilities for the control of a variety of bacterial, protozoal, and fungal diseases.[3,4] In addition, since neoplastic tissue constitutes an altered differentiation state, it is more susceptible than normal tissue to the effects of antifolates. Studies from our laboratory and others have demonstrated that the interaction of inhibitors with DHFR is an excellent way to develop the critical methodology for drug design at the enzyme level.[5] Our efforts have focused on bacterial[6] and antitumor chemotherapy.[7]

There has been a huge ongoing effort to discover more effective antibacterial benzylpyrimidines.[5] However, there has been little concerted effort by a single laboratory using a standard test procedure to systematically explore the structure–activity relationships of this class of compounds on highly purified DHFR from several sources. Also, little effort has been made to establish methodology for the design of selective inhibitors utilizing QSAR (quantitative structure–activity relationships) and molecular graphics. Intense proprietary concern has meant that those most able to shed light on this problem have remained silent. In this report, we review and extend our comparison of the QSAR for DHFR from three different sources: *Lactobacillus casei*, *Escherichia coli*, and chicken liver.

With use of 2,4-diamino-5-(substituted-benzyl)pyrimidines (I), QSAR for the inhibition of *L. casei* DHFR and chicken liver DHFR have been previously derived.[6,8]



I

In the eqs 1 and 2, K_i is the apparent inhibition constant, the subscripted π values represent the hydrophobic constants for substituents in the corresponding positions on the benzyl ring of I, and MR (scaled by 0.1) is the molar refractivity of substituents.[9] The prime associated with this term means that MR has been truncated at 0.79 regardless of the actual value of MR. For simultaneous substitution of the 3- and 4-positions, the maximum value is 2 × 0.79 (1.58). The minimum value is 2 × 0.1 (i.e. 2 × MR_H). Since MR is primarily a measure of substituent bulk, it appears that only part of the substituent contacts the enzyme.[4] The positive MR terms are presumed to model positive steric and/or dispersion effects, thus increasing the value of log $(1/K_i)$. The σ terms (Hammett constants) are of marginal value. The parameter β is an adjustable parameter of the bilinear model which helps to determine the optimum values of π and MR (π^0 and MR^0).[10] Each equation is based on 65 compounds (n), r is the correlation coefficient, and s is the standard deviation.

inhibition of *L. casei* DHFR by I

$$\log(1/K_i) = 1.24MR_4' + 0.52MR_3' + 0.42MR_5 - 0.13MR_5^2 + 0.46\pi_4 + 0.31\pi_3' - 0.92\log(\beta_4\cdot 10^{\pi_4} + 1) - 0.71\log(\beta_3\cdot 10^{\pi_3'} + 1) + 5.45 \quad (1)$$

$$n = 65,\ r = 0.894,\ s = 0.245,\ \log\beta_4 = -0.50,\ \log\beta_3 = -1.43,\ \pi_4^0 = 0.49,\ \pi_3^0 = 1.33,\ MR_5^0 = 1.66$$

inhibition of chicken liver DHFR by I

$$\log(1/K_i) = 0.39\pi_3 + 0.44\pi_4 - 0.75MR_5 + 0.44\sigma - 1.04\log(\beta_3\cdot 10^{\pi_3} + 1) + 0.37\pi_5 - 0.32\log(\beta_4\cdot 10^{\pi_4} + 1) + 4.70 \quad (2)$$

$$n = 65,\ r = 0.906,\ s = 0.207,\ \pi_3^0 = 2.45,\ \pi_4^0 \approx 3.00,\ \log\beta_4 = -0.18,\ \log\beta_3 = -2.69$$

It is evident that eqs 1 and 2 are complex expressions, and although they are based on a large number of compounds of quite varied structure, one may be uneasy with their content. It must be recognized that the active site of DHFR is complex and has a variegated surface and that inhibitor substitution has been made in the 3-, 4- and 5-positions of I with both small and large substituents (see Table I). Hence, eqs 1 and 2 are attempts to describe the "inlaying" of complex ligands into a complex, multifaceted cavity. Obviously steric effects are involved, but hydrophobic and electronic factors also come into play. A few moments reflection leads to the view that it is not the complexity of the correlation equations which is worrisome, but their simplicity which is striking. It is this simplicity of the QSAR which allows about 20% $[100(1 - r^2)]$ of the variance in log $(1/K_i)$ to go "unexplained".

Standing alone, QSAR 1 and 2 would not be easily convincing; however, when compared with the molecular graphics models based on the X-ray coordinates of the ternary complex of chicken liver DHFR–NADPH–trimethoprim (I, X = 3,4,5-$(OCH_3)_3$) and the ternary complex of methotrexate–NADPH–*L. casei* DHFR, the good agreement between the mathematical model and the graphics model confirms our beliefs in this approach to understanding complex enzyme–ligand interactions.[6] In fact, the many examples where QSAR and molecular graphic models (always based on X-ray crystallographic coordinates) are in good agreement encourages us in utilizing this double-pronged approach for verification in structure–activity studies.[11]

Equation 1, which accentuates the importance of terms in MR, indicates that the steric effect of substituents is paramount in the inhibition of the bacterial DHFR and that hydrophobic effects, except in the 3-position, are of modest importance. In the case of the vertebrate DHFR, the situation is just the reverse. The most important effect is hydrophobic, in positions 3 and 4; the large negative MR_5 term completely counterbalances the π_5 term. It is this difference which accounts for much, but not all, of the selectivity of the benzylpyrimidines for bacterial enzyme compared to vertebrate enzyme. The log $(1/K_i)$ for the

(2) Gready, J. E. *Adv. Pharmacol. Chemother.* **1980**, *17*, 37.
(3) Roth, B.; Rauckman, B. S.; Ferone, R.; Baccanari, D. P.; Champness, J. N.; Hyde, R. M. *J. Med. Chem.* **1987**, *30*, 348.
(4) Hitchings, G. H.; Smith, S. L. *Adv. Enzyme Regul.* **1980**, *18*, 349.
(5) Blaney, J. M.; Hansch, C.; Silipo, C.; Vittoria, A. *Chem. Rev.* **1984**, *84*, 333.
(6) Selassie, C. D.; Fang, Z-X.; Li, R.-L.; Hansch, C.; Debnath, G.; Klein, T. E.; Langridge, R.; Kaufman, B. T. *J. Med. Chem.* **1989**, *32*, 1895.
(7) Selassie, C. D.; Strong, C. D.; Hansch, C.; Delcamp, T. J.; Friesheim, J. H.; Khwaja, T. A. *Cancer Res.* **1986**, *46*, 744.
(8) Selassie, C. D.; Fang, Z-X.; Li, R. L.; Hansch, C.; Klein, T. E.; Langridge, R.; Kaufman, B. T. *J. Med. Chem.* **1986**, *29*, 621.
(9) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley Interscience: New York, 1979; Chapter 1.
(10) Kubinyi, H.; Kehrhahn, O. H. *Arzneim-Forsch.* **1978**, *28*, 598.
(11) Hansch, C.; Klein, T. E. *Acc. Chem. Res.* **1986**, *19*, 392.

Table I. Parameters Used in the Derivation of Eqs 5–8 for the Inhibition of *E. coli* DHFR by 2,4-Diamino-5-(substituted-benzyl)pyrimidines

no.	X	log $(1/K_i)$ obsd[a]	log $(1/K_i)$ calcd[b]	$MR_{5'}$	$MR_{3'}$	MR_4	π_3
1	3,4,5-$(CH_2CH_3)_3$	7.82	8.15	0.79	0.79	1.03	0.86
2	3,5-$(OCH_3)_2$, 4-$OCH_2CH_2OCH_3$	8.35	8.14	0.79	0.79	1.93	0
3	3,4,5-$(OCH_3)_3$	8.08	7.90	0.79	0.79	0.79	0
4	3,5-$(OCH_3)_2$, 4-$N(CH_3)_2$	7.71	8.12	0.79	0.79	1.56	0
5	3,5-$(OCH_3)_2$, 4-Br	8.18	7.94	0.79	0.79	0.89	0
6	3,5-$(OCH_3)_2$, 4-SCH_3	8.07	8.09	0.79	0.79	1.38	0
7	3,5-$(OCH_3)_2$, 4-$C(CH_3){=}CH_2$	8.12	8.12	0.79	0.79	1.56	0
8	3,5-$(OCH_2CH_3)_2$, 4-pyrryl	7.66	8.27	0.79	0.79	1.95	0.38
9	3,5-$(OCH_3)_2$, 4-$O(CH_2)_7CH_3$	7.20	7.17	0.79	0.79	3.97	0
10	3,5-$(CH_2OH)_2$	6.31	6.33	0.72	0.72	0.10	−1.03
11	3,5-$(OCH_3)_2$	7.71	7.48	0.79	0.79	0.10	0
12	3,5-$(OCH_2CH_3)_2$	7.69	7.62	0.79	0.79	0.10	0.47
13	3-OC_2H_5, 5-OC_3H_7	7.69	7.62	0.79	0.79	0.10	1.05
14	3,5-$(OC_3H_7)_2$	7.41	7.62	0.79	0.79	0.10	1.05
15	3,5-$(CH_3)_2$	7.04	7.23	0.57	0.57	0.10	0.56
16	3,4-$(OH)_2$	6.46	6.52	0.29	0.10	0.29	0
17	3-NO_2, 4-$NHCOCH_3$	6.97	7.44	0.74	0.10	1.49	0
18	3,4-$(OCH_2CH_2OCH_3)_2$	7.22	7.52	0.79	0.10	1.93	0
19	3,4-OCH_2O	7.13	6.78	0.45	0.10	0.45	0
20	3,4-$(OCH_3)_2$	7.72	7.28	0.79	0.10	0.79	0
21	3-OH, 4-OCH_3	6.84	6.81	0.29	0.10	0.79	0
22	3-OCH_3, 4-OSO_2CH_3	7.94	7.52	0.79	0.10	1.70	0
23	3-OCH_3, 4-OH	7.54	7.00	0.79	0.10	0.29	0
24	3-OCH_3, 4-$OCH_2CH_2OCH_3$	7.77	7.52	0.79	0.10	1.93	0
25	3-OCH_3, 4-$OCH_2C_6H_5$	7.53	7.15	0.79	0.10	3.17	0
26	3-OSO_2CH_3, 4-OCH_3	7.80	7.28	0.79	0.10	0.79	0
27	3-$OCH_2C_6H_5$, 4-OCH_3	7.66	7.36	0.10	0.79	0.79	1.27
28	3-CF_3, 4-OCH_3	7.69	7.14	0.10	0.50	0.79	0.87
29	3-$O(CH_2)_7CH_3$, 4-OCH_3	7.16	6.91	0.10	0.79	0.879	3.69
30	3-OCH_2CH_3, 4-$OCH_2C_6H_5$	7.35	7.24	0.10	0.79	3.17	0.38
31	3-OCH_2CONH_2	6.57	6.86	0.79	0.10	0.10	0
32	3-CH_2OH	6.28	6.79	0.72	0.10	0.10	0
33	3-OSO_2CH_3	6.92	6.86	0.79	0.10	0.10	0
34	3-CH_2OCH_3	6.59	6.86	0.79	0.10	0.10	0
35	3-OH	6.47	6.39	0.29	0.10	0.10	0
36	3-$OCH_2CH_2OCH_3$	6.53	6.86	0.79	0.10	0.10	0
37	3-OCH_3	6.93	6.84	0.10	0.79	0.10	0.04
38	3-F	6.23	6.29	0.10	0.09	0.10	0.23
39	3-CH_3	6.70	6.78	0.10	0.57	0.10	0.52
40	3-Cl	6.65	6.81	0.10	0.60	0.10	0.67
41	3-Br	6.96	6.98	0.10	0.79	0.10	0.86
42	3-CF_3	7.02	6.72	0.10	0.50	0.10	0.88
43	3-$CH_2O(CH_2)_3CH_3$	6.55	6.94	0.10	0.79	0.10	1.30
44	3-I	7.23	6.96	0.10	0.79	0.10	1.12
45	3-$O(CH_2)_3CH_3$	6.82	6.90	0.10	0.79	0.10	1.55
46	3-$OCH_2C_6H_5$	6.99	6.90	0.10	0.79	0.10	1.56
47	3-$O(CH_2)_5CH_3$	6.86	6.70	0.10	0.79	0.10	2.63
48	3-$O(CH_2)_6CH_3$	6.39	6.58	0.10	0.79	0.10	3.23
49	3-$O(CH_2)_7CH_3$	6.25	6.47	0.10	0.79	0.10	3.79
50	4-NH_2	6.30	6.50	0.10	0.10	0.54	0
51	4-$NHCOCH_3$	6.89	6.84	0.10	0.10	1.49	0
52	4-OSO_2CH_3	6.60	6.87	0.10	0.10	1.70	0
53	4-OH	6.45	6.34	0.10	0.10	0.29	0
54	4-$OCH_2CH_2OCH_3$	6.40	6.87	0.10	0.10	0.93	0
55	4-NO_2	6.20	6.60	0.10	0.10	0.74	0
56	4-OCH_3	6.82	6.63	0.10	0.10	0.79	0
57	4-F	6.35	6.20	0.10	0.10	0.09	0
58	4-$N(CH_3)_2$	6.78	6.85	0.10	0.10	1.56	0
59	4-CH_3	6.48	6.51	0.10	0.10	0.57	0
60	4-Cl	6.45	6.53	0.10	0.10	0.60	0
61	4-Br	6.82	6.67	0.10	0.10	0.89	0
62	4-OCF_3	6.57	6.63	0.10	0.10	0.79	0
63	4-$O(CH_2)_3CH_3$	6.89	6.85	0.10	0.10	2.17	0
64	4-$OCH_2C_6H_5$	6.89	6.49	0.10	0.10	3.17	0
65	4-$O(CH_2)_5CH_3$	6.07	6.55	0.10	0.10	3.07	0
66	4-$O(CH_2)_6CH_3$	6.10	6.27	0.10	0.10	3.52	0
67	4-C_6H_5	6.93	6.77	0.10	0.10	2.54	0
68	4-H	6.18	6.21	0.10	0.10	0.10	0

[a] Observed in Tris buffer, pH 7.20. [b] Calculated with eq 8.

parent form of I (X = H) against chicken DHFR is 4.71 while for *L. casei* DHFR it is 5.20 (Δ log $(1/K_i)$ = 0.49). Hence, the intrinsic activity of the benzylpyrimidines is 3 times greater against bacterial DHFR. It is important to note that the clinically important drug trimethoprim is 800 times more potent against the bacterial DHFR than avian DHFR (log $(1/K_i)$ of *L. casei* DHFR = 6.88; log $(1/K_i)$ of chicken DHFR = 3.99; Δ log $(1/K_i)$ = 2.90).

While the above results provide a generally satisfying explanation of the selectivity of trimethoprim and other benzylpyrimidines, one requires more confirmation of this statistical approach to the understanding of drug-receptor interactions. Hence, it was decided to extend our earlier studies on the *E. coli* organism and its gene-amplified DHFR.[12]

In our first preliminary analysis of *E. coli* DHFR, we obtained the following QSAR for the inhibitory effects of I:[12]

$$\log (1/K_i) = 0.43\pi_{3,4,5} + 1.23\text{MR}_{3,5}' + 0.80\text{MR}_4' - 0.88(\beta\cdot10^{\pi_{3,4,5}} + 1) - 0.45\sigma_{R^-} + 5.81 \quad (3)$$

$$n = 43,\ r = 0.923,\ s = 0.263,\ \pi^0 = 0.64,\ \log \beta = -0.67$$

The parameters are as previously defined[12] except that σ_{R^-} is the summation of the "through" resonance contributions of the substituents on the phenyl ring. Because of the relatively few congeners, it was not possible to separately factor, with confidence, the substituents from the 3-, 4-, or 5-positions. The hydrophobic components and the molar refractivity values for the two meta positions were simply summed.

In a second assault on the problem, eq 4 was derived.[13]

$$\log (1/K_i) = 1.28\text{MR}_{3,5}' + 1.16\text{MR}_4' + 1.39\ \pi_{3,4,5} - 1.68 \log (\beta\cdot10^{\pi_{3,4,5}} + 1) - 0.09\sigma_{R^-} + 1.48I + 6.95 \quad (4)$$

$$n = 61,\ r = 0.907,\ s = 0.542,\ \pi^0{}_{3,4,5} = 0.07,\ \log \beta = 0.61$$

A shortcoming of eq 4 is that two different assay procedures were utilized so that it was essential to use an indicator variable (I) which is assigned the value of 1 for one method of testing[13] and 0 for another method of assay.[6] The coefficient of 1.48 with I shows that the analytical method used by Li and Poe produces $1/K_i$ values about 30 times as high (antilog of 1.48) as our methodology assuming eqs 4 and 3 to be similar. Equation 4 confirms our conclusion that σ_{R^-} is of marginal importance. Although there are many similarities between eqs 3 and 4, the differences are great enough, especially with regard to the coefficient of π, to warrant further investigation. The standard deviation is noticeably larger and r is significantly smaller for eq 4. Also, eq 4 is not based on the same treatment which we have recently found to be applicable to *L. casei* DHFR.[6] Thus, it is important to make a more careful evaluation of eq 3 using measured partition coefficients and data from a single test methodology. The diversity in the substitution positions also makes it critical to readdress the QSAR in terms of the separate contributions of each position.

In order to determine all of the K_i values by using the same procedure and to include new variations of I, we now extend eq 3 by means of 25 new data points and achieve a finer resolution of substituent effects which we believe will be of more value in the design of more selective DHFR inhibitors.

Experimental Section

The partition coefficients, substituent constants, and the experimental method for the determination of the apparent K_i have been previously reported.[6] Three congeners, trimethoprim [X = 3,4,5-$(OCH_3)_3$], tetroxoprim [X = 3,5-$(OCH_3)_2$, 4-$OCH_2CH_2OCH_3$] and a 3,5-$(OCH_3)_2$ derivative, were retested to reduce error bars that were present in the data obtained from the earlier enzymic analysis. The results for tetroxoprim remained the same but the other two values are slightly altered.

(12) Hansch, C.; Li, R.-L.; Blaney, J. M.; Langridge, R. *J. Med. Chem.* **1982**, *25*, 777.
(13) Li, R.-L.; Poe, M. *J. Med. Chem.* **1988**, *31*, 366.

Table II. Squared Correlation Matrix

	π_3'	MR_3'	MR_4	MR_5'
π_3'	1	0.27	0.08	0.09
MR_3'		1	0.03	0.02
MR_4			1	0.00
MR_5'				1

Results

Equations 5–8 derived from the data in Table I show the development of the QSAR for *E. coli* DHFR for benzylpyrimidines I.

$$\log (1/K_i) = 1.13\ (\pm0.35)\text{MR}_5' + 6.57\ (\pm0.18) \quad (5)$$

$$n = 68,\ r = 0.618,\ s = 0.477,\ F_{1,66} = 40.7$$

$$\log (1/K_i) = 1.04\ (\pm0.31)\text{MR}_5' + 0.69\ (\pm0.31)\text{MR}_3' + 6.33\ (\pm0.19) \quad (6)$$

$$n = 68,\ r = 0.725,\ s = 0.421,\ F_{1,65} = 19.8$$

$$\log (1/K_i) = 0.92\ (\pm0.25)\text{MR}_5' + 0.90\ (\pm0.25)\text{MR}_3' + 0.81\ (\pm0.25)\text{MR}_4 - 0.22\ (\pm0.07)\text{MR}_4^2 + 5.95\ (\pm0.19) \quad (7)$$

$$n = 68,\ r = 0.850,\ s = 0.327,\ F_{2,63} = 44.6,\ \text{MR}_4^0 = 1.87\ (\pm0.19)$$

$$\log (1/K_i) = 0.95\ (\pm0.24)\text{MR}_5' + 0.89\ (\pm0.27)\text{MR}_3' + 0.80\ (\pm0.22)\text{MR}_4 - 0.21\ (\pm0.07)\text{MR}_4^2 + 1.58\ (\pm0.73)\pi_3' - 1.77\ (\pm0.80) \log (\beta_3\cdot10^{\pi_3'} + 1) + 6.65\ (\pm0.36) \quad (8)$$

$$n = 68,\ r = 0.890,\ s = 0.290,\ F_{3,60} = 19.8,\ \text{MR}_4^0 = 1.85\ (\pm0.20),\ \pi_3^0 = 0.73\ (\pm1.06),\ \log \beta_3 = 0.175$$

The squared correlation matrix is outlined in Table II. The parameters of eqs 5–8 have the same connotation as for eq 1. The figures in parentheses are for the construction of the 95% confidence limits and the F values establish the significance of the terms in each equation compared to the preceding one. The above correlation matrix shows that there is reasonable orthogonality between the variables of eq 8 except for π_3 and MR_3, which show some collinearity. The eigenvalues, with the fraction of the variance accounted for by each in parentheses, are 17.0 (0.43), 1.11 (0.28), 0.86 (0.22), and 0.33 (0.08). This would justify the use of the four variables.

As with *L. casei* DHFR, the most important parameter is MR, whose values for the 3- and 5-positions are truncated at 0.79. Equation 7, having only MR terms, accounts for 72% of the variance in $\log (1/K_i)$ while eq 8 accounts for 79%. Thus the π terms account for only 7% of the variance. Although the overall bulk of the substituents is critical, it is the bulk of the initial atoms of the substituent adjacent to the phenyl ring that is most crucial.

Attempts to use bilinear terms or parabolic terms for MR_3 and MR_5 were not statistically as successful as using the truncated MR′ values. Most of the activity enhancement of both meta substituents comes from a positive steric effect of the first one or two atoms attached to the benzene ring. An optimum bulk for para substituents of 1.85 can be established and an optimum for π_3 of 0.73 is found via the bilinear equation. This optimum may be compromised because of the slight collinearity between MR_3' and π_3' especially since both terms are modeling the same enzymic site.

The *E. coli* QSAR is considerably simpler than that of the *L. casei* DHFR (six terms vs eight). Although the quality of fit is slightly less significant, hydrophobic factors appear to be even less important than for the *L. casei* system, as no hydrophobic contribution appears to be

rendered by para substituents. It must be noted, however, that the π_4 contribution to inhibition with *L. casei* DHFR is minimal because of the small coefficient with this term. Additionally, the low value of 0.49 for the hydrophobic component in the para position means that the maximum positive contribution in this position would be 0.23 (0.46 × 0.49), which is within the standard deviation of the regression eq 1. In the case of 3-substituents, the optimum hydrophobicity for *E. coli* is considerably smaller than that of *L. casei*—0.73 vs 1.33, suggesting a smaller binding region for *E. coli* DHFR. The MR_3 term plays a slightly different role with the two bacterial enzymes. In the case of *E. coli* DHFR, a truncated value of 0.79 works best while with *L. casei* DHFR an optimum value of 1.66 is seen.

The most potent compounds in Table I are those with 3,5-dimethoxy substitution. With these two substituents, the best 4-substituted derivative is 4-$OCH_2CH_2OCH_3$, having a MR_4 value of 1.95, which is very close to the predicted ideal value of 1.85. The 4-(dimethylamino) and 4-pyrryl-3,5-dimethoxy analogues are both poorly fit, being less active than expected even though their MR_4 values are near ideal. These two amino compounds would be constrained at more or less right angles to the phenyl ring and it may be this projection above and below the ring plane which is deleterious. The simple 4-(dimethylamino) derivative is well predicted. The 3,5-dimethoxy-4-isopropenyl derivative, despite its structural similarity to the 3,5-dimethoxy-4-(dimethylamino) analogue, is well predicted and is one of the most active compounds in Table I. This may be related to its hydrophobicity.

Another perspective on the inhibition of the three types of DHFR can be realized by comparing the highly potent, clinically utilized trimethoprim (TMP) with its parent benzylpyrimidine.

	E. coli DHFR	*L. casei* DHFR	chicken
TMP	8.08	6.88	3.98
X = H	6.18	5.20	4.71
	1.90	1.68	−0.72

Although TMP is much more potent against the *E. coli* enzyme, comparison with the parent congener (X = H) indicates that substituents enhance its activity slightly more than for *L. casei* DHFR. The methoxy groups actually have a detrimental effect in the case of the chicken DHFR. The parent benzyl pyrimidine and TMP are about 10 and 15 times more effective, respectively, against *E. coli* than *L. casei* enzyme.

The more detailed analysis of eq 8 enhances our understanding of the hydrophobic interactions at play. It is clear from both graphics and the QSAR that only 3-substituents have a significant selectivity index. The π coefficient of eq 8 resembles that of eq 4 more closely than that of eq 3. The artifactual "through resonance" term which has almost vanished in eq 4 has, with the better selection of substituents, disappeared from eq 8. And, for the first time, we can see from eq 8 that 5-substituents do have a significant role to play.

Discussion

Deriving QSAR for the large number of complex benzylpyrimidines binding to the convoluted DHFR active site is a serious challenge. The QSAR for the three types of DHFR of eqs 1, 2, and 8 are satisfying and are of considerable value in the design of new congeners, even though they do not account for about 20% of the variance in log $(1/K_i)$. While sharper correlations for enzymic QSAR have been reported, none of these encompass such a complicated receptor or such a wide variety of substituents.[11] It is apparent that obtaining a statistically improved equation pertaining to a heterogenous data set containing unsymmetrical and various sized substituents would be difficult without resorting to the use of an inordinate number of variables. This is particularly true when a flexible substituent containing heteroatoms binds in hydrophilic space created by heteroatoms in the enzyme. The degree of flexibility of the enzyme creates additional complications, although molecular dynamics calculations could prove valuable in this area. Added to these constraints is some uncertainty in the parameter values as well as in the crystallographic structure of the enzyme and the structure of the enzyme in solution.

At present there is no agreed upon definition of "hydrophobicity". Acceptance of an operational definition such as octanol/water partition coefficients as a standard grants no assurance that octanol is the best solvent. As in the case of σ, where different types of structures and reactions have necessitated the use of several variations of σ, it may be necessary to employ more than one solvent system to model all of the different types of biological systems. Leahy et al. have recently discussed this problem and proposed a new solvent—propylene glycol dipelargonate.[14] However recent results confirm that octanol/water partition coefficients remain an appropriate operational definition of hydrophobic character at the enzymic level.[14-16]

The electronic properties of substituents are much better defined and more extensively studied than hydrophobic properties, but even here problems may arise since σ constants fluctuate in value, depending on the nature of the surrounding solvent. This is particularly valid for substituents which can strongly interact with the solvent (such as NH_2 or an OH with water). Hence, if the substituent is engulfed in the receptor, its electronic effect may vary considerably from that in the solvent.

A most serious drawback is that compounded by steric interactions. It remains difficult to formulate a general set of steric parameters of any reasonable size which can be used to quantitatively describe all of the steric interactions of a complex ligand with an intricate receptor. Thus, it is envisioned that if electronic and hydrophobic properties of substituents can be delineated with some confidence, then what remains will be steric effects sometimes confounded by hydrogen bonding and dipolar interactions. It is readily apparent that there is no neat deterministic solution of the SAR problem even with purified enzymes, and the problem is undoubtedly magnified with cellular systems or whole animals.

Nevertheless, we are pleased with the rate that QSAR has progressed and optimistic about the prospects for further advances. The combined use of molecular graphics and QSAR has demonstrated that there is good correspondence between what one deduces from the terms in a QSAR and what one sees from a graphics analysis of the ligands fitting a model of the receptor.[5,11]

The parameter MR needs some discussion at this point. It was first introduced into SAR studies in 1945 by Pauling and Pressman, who considered it as a model for the polarizability of substituents.[17] They indicated that MR is related to London dispersion forces as follows: $MR = 4\pi N\alpha/3$, where polarizability is represented by α, N is Avogadro's number, and π is 3.14. Experimentally MR is

(14) Leahy, D. E.; Taylor, P. J.; Wait, A. R. *Quant. Struct. Act. Relat.* **1989**, *8*, 17.
(15) Sakurai, T.; Margolin, A. L.; Russell, A. J.; Klibanov, A. M. *J. Am. Chem. Soc.* **1988**, *110*, 7236.
(16) Matsumura, M.; Becktel, W. J.; Mattews, B. W. *Nature* **1988**, *334*, 406.

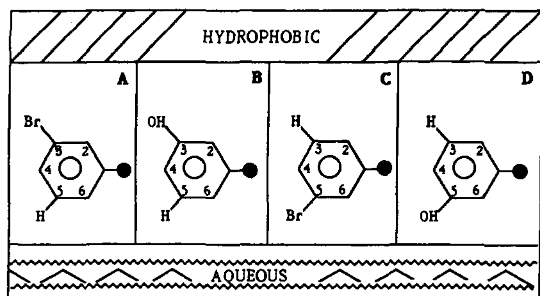


Figure 1. Different hydrophobic/hydrophilic binding modes of (substituted-benzyl)diaminopyrimidines.

normally obtained via the Lorentz–Lorenz equations, MR = $[(n^2 - 1)/(n^2 + 2)](MW/d)$, where n is the index of refraction, MW is the molecular weight, and d represents density. Many studies in the early part of this century showed that MR is a highly additive, constitutive property of organic compounds so that it constitutes a readily calculable value. Since n varies minimally for most organic compounds, MR as obtained experimentally is primarily a measure of volume (MW/d).[9] Thus, in recent years it has been viewed as a measure of bulk and as a "rough and ready" steric parameter. Interpretation of the role of MR in a correlation equation depends on the coefficient with this term. A negative coefficient is less ambivalent—it is assumed to be indicative of a negative steric effect. Since MR contains no significant information on the geometry of the substituent, it cannot be expected to graphically elucidate steric problems. It is used to model intermolecular effects between ligand and receptor.

A more complex problem involves interpretation of a positive coefficient with MR. The idea that it represents dispersion forces has lost credence since it has been recognized that a ligand or a substituent partitioning from an aqueous environment to a protein surface exchanges one set of dispersion forces for another with little or no overall change in free energy. Thus, the importance attached to dispersion forces in drug–receptor interactions has diminished.

Assuming little collinearity between π and MR and that correlation is with MR, not π, implies contact with a polar receptor surface. This is somewhat oversimplified since a bulk effect in hydrophobic space could be correlated with MR under some conditions (3-space in the case of eq 8). Recent QSAR–molecular graphic studies of enzyme–ligand interactions with known X-ray crystallographic coordinates of the enzyme provide insight and support for the type of parameterization used in eq 1 and 8.[11] In a number of such studies it has been found that a phenyl ring by 180° rotation can place meta substituents in two different environments as illustrated in Figure 1. Here we consider two different substituents—a hydrophobic bromo and a hydrophilic hydroxy group.

Example A is a normally expected hydrophobic–hydrophobic contact represented by π of which many examples have now been confirmed by QSAR–graphics studies, including the present.[11] However, for hydrophilic substituents it has been determined that when given the option, polar substituents appear to orient away from the hydrophobic surface (model B) into the aqueous/polar phase (model D).[11] Hence, dispersion forces do not hold the hydroxy group onto the hydrophobic surface in B. Consequently using its π value of −0.67 in a QSAR analysis would not yield a good correlation since OH would prefer to reside in the aqueous/polar phase. Instead, π of 0 for the meta hydrogen should be utilized as demonstrated in model D. This approach has been employed in eq 1 and 8, similar to what has been done with other enzymes.[11] The same predicament could conceivably befall ortho substituents, but not para substituents. This cannot occur unless rotation of about 180° around the bond occurs—anchoring the ligand through some functional group to the receptor is not significantly constrained energywise. Thus, as we probe deeper into the study of receptor binding, more thought must be given to the possible orientation of substituents when assigning them suitable parameters. In the case of two hydrophobic meta substituents, one assigns a normal π constant to the more hydrophobic of the two substituents and the appropriate hydrophobic constant to the other substituent. The lesson from doing QSAR is that one cannot simply "machine load" parameters, press the regression-analysis button, and get the answer. Trial and error studies must be done before an appropriate approximate model can be formulated.

The C and D type situations are presumed to be modeled by MR where polar space on the receptor is involved. These types of interactions have been documented[11,18] in a few examples where the structure of the enzyme is known. At present there does not appear to be an interaction of type B where a polar substituent positively enhances binding via dispersion forces. If it does occur, it is not discernible beyond the detection limits of QSAR at present.

It is now apparent from QSAR–graphics analysis that a positive coefficient with MR terms can be deduced for cases such as C and D. We are reluctant to ascribe these positive coefficients as being due primarily to dispersion forces, although these may be involved. The results in hand[11,18] suggest the following. In several instances polar substituents appear to help position ligands so that they are better able to participate in catalysis.[11,18] This has been viewed as a "buttressing" effect. It is well-known from studies of simple organic reactions that such effects can be extremely important in affecting reactions rates.[19] Another mechanism, which might be especially important with inhibitors, would be distortion of the orientation of critical amino acid residues involved in catalysis with the normal substrate and/or cofactor. This latter mechanism could be operative in the case of the DHFR inhibitors in the present study.

In the parameterization of eq 8, the best correlation results occur when the substituents are analyzed on the basis of the following assumptions: (1) A hydrophobic substituent in the meta position is assigned its π value at the 3-position according to model A. (2) A hydrophilic substituent in the meta position rotates 180° such that it is placed in the 5-position and assigned its negative π value. The corresponding H is thus placed in the 3-position and accordingly assigned a value of 0. See model D. (3) Hydrophobic substituents in both meta positions warrant the placement of the more hydrophobic substituent in the 3-position and the less hydrophobic substituent in the 5-position. They are then assigned their appropriate hydrophobic π values. (4) Likewise, hydrophilic substituents in both meta positions lead to placement of the more hydrophilic substituent in position 5 and the less hydrophilic substituent in position 3.

Compound 10 in Table I (3,5-$(CH_2OH)_2$) is an excellent test of these assumptions. Since the substituents are equivalent, it does not matter where they are assigned. One of them is thus forced into hydrophobic 3-space and

(17) Pauling, L.; Pressman, D. *J. Am. Chem. Soc.* **1945**, *67*, 1083.
(18) Compadre, C. M.; Hansch, C.; Klein, T. E.; Langridge, R. *Biochim. Biophys. Acta.*, in press.
(19) Menger, F. M. *Acc. Chem. Res.* **1985**, *18*, 128.

the other one is aligned in hydrophilic 5-space. All in all, it is very well predicted by eqs 1, 2, and 8. Although its steric properties are close to the "magic" OCH_3 group, it is much more hydrophilic and correspondingly much less effective. This reduces its activity versus the two bacterial enzymes since one hydroxymethyl group is forced into a hydrophobic region. However, the effect is much greater with the chicken DHFR, where two hydrophilic groups are forced into hydrophobic space. To date, this is the weakest benzylpyrimidine type inhibitor that we have found versus chicken liver DHFR. It is also the least effective benzylpyrimidine-type growth inhibitor against both methotrexate sensitive and resistant leukemia cells.

With *E. coli* DHFR, a single hydrophobic meta substituent appears to bind in hydrophobic space and produces a positive steric effect as well as a hydrophobic effect. Thus, it is best represented by a π term and a truncated MR term. Polar meta substituents which appear to orient away from hydrophobic 3-space are parameterized only by the truncated MR. The coefficients in eq 8 suggest that the MR effect is about the same with either 3- or 5-substitution. However, since there is some collinearity between π_3' and MR_3', separation of steric and hydrophobic effects in 3-space may not be sharply delineated. The increase in inhibitory activity parallels the increase in hydrophobicity of meta substituents up to the optimum value of 0.73 and then the more hydrophobic substituents have essentially no effect, i.e. the slope of the right hand side of the bilinear model is only −0.19 (1.58−1.77). Considering the confidence limits, this is not significantly different from zero.

A particularly interesting aspect of this analysis concerns the role of the methoxy moieties since TMP contains three such groups and many researchers have attributed some exceptional character to it. Analyzing the methoxy groups via MR′ along with other meta substituents demonstrates that there is no "special aura" associated with these groups. The methoxy-containing compounds are no better or worse fit than other substituents of their size and hydrophobicity.

The same method of analysis has been employed for both bacterial enzymes, but the avian DHFR is not treated in this manner. The coefficients in eq 2 for π_3, π_4, and π_5 are essentially the same, suggesting the same quality of hydrophobic interaction for small groups in these positions. For larger groups in the 3- and 4-positions, optimum values (π_0) are found for 3- and 4-substituents, although the amount of space available differs for each type (compare π_0 values). No optimum could be defined for π_5. The situation here is more complicated. 5-space on the enzyme is restricted because of the presence of Tyr-31. However, it has been shown that the side chain of this group moves rather easily, opening up a rather large hydrophobic region behind it. Thus, the steric effect seems to be accounted for by MR_5. Given the option of binding in 3- or 5-space, a single meta substituent appears to bind in 3-space regardless of whether or not it is hydrophobic or hydrophilic. While hydrophilic substituents binding in 3-space would have a negative effect on log $(1/K_i)$, they appear to favor this binding mode rather than displace Tyr-31.

The QSAR's of eqs 1, 2, and 8 provide some broad semblance of receptor–ligand fit and are of great value in the design of more selective drugs. Since the QSAR for human and chicken DHFR are quite similar, chicken DHFR can be used as a surrogate for the human enzyme.[8] In comparing the bacterial and prokaryotic enzymes, it is evident that steric effects modeled by MR are absolutely crucial for increasing inhibitory potency of bacterial DHFR, while the hydrophobic properties of substituents primarily govern potency toward vertebrate DHFR. This provides an important clue for designing more selective drugs. Polar substituents with ideal MR values will increase bacterial potency, but the binding of these substituents will be diminished at the hydrophobic surfaces of the vertebrate enzyme. One could hardly ask for a better situation to design selective inhibitors. This simple fact was only illuminated rather recently via QSAR despite the fact that an enormous effort has been made during the last 30 years by a number of large drug companies and many academic investigators to understand the selectivity of the benzylpyrimidines, particularly TMP.[6]

The inherent difficulty in dealing with the complexity of the above QSAR and the number of parameters needed to describe the situation is recognized, but the bilinear and parabolic terms provide key clues in the design of new congeners and focus one's attention on the most likely optimum size (hydrophobic as well as steric) so that more refined substituents can be designed for maximum effect.

The new results with *E. coli* DHFR are qualitatively like those found for *L. casei* DHFR and using data from a single test system yields a more coherent QSAR (lower standard deviation than eq 4). The problem of "ring flipping" appears to occur with both bacterial DHFR–benzylpyrimidine interactions and is in line with what has been found with other enzymes.[11] This potential behavior of meta substituents (which have similar elements of binding symmetry) has very important consequences for drug design. It is likely that many drug receptors have a cleftlike character in which an aromatic ring could bind such that hydrophobic meta substituents (up to a limited size) would enhance such binding, but hydrophilic substituents could orient toward the aqueous phase and subsequently show no effect. When such cases can be uncovered via QSAR and/or molecular graphics, there is another possible benefit to be derived. It has been pointed out that hydrophobic drugs enter the central nervous system (CNS) more readily than hydrophilic drugs.[20] Moreover, in a given series of compounds, toxicity often increases with hydrophobicity,[21] and hydrophobic compounds tend to induce cytochrome P-450, which results in their increased metabolism.[22,23] Hence, as a general principle one should strive to make drugs as hydrophilic as possible, commensurate with efficacy. In knowing from QSAR studies of isolated receptors that one has a situation similar to that of eq 1 and 8, one can design ligands with an optimum hydrophobic substituent in one meta position and a hydrophilic substituent in the other meta position such that the overall hydrophobicity (log *P*) of the drug is as low as possible. This would maximize hydrophobic interaction with the receptor and minimize untoward hydrophobic effects such as CNS penetration, nonspecific toxicity, and interactions with the P-450 enzymes.

Designing and synthesizing new congeners and checking their activity against a developing QSAR can provide insight not available from graphic studies of ligands and receptors. Almost all of the compounds in Table I were designed before the availability of graphic capabilities so that the QSAR was of significant assistance in understanding how the ligands fit to the receptor. The reverse process, deducing the QSAR from the graphics, is not yet

(20) Hansch, C.; Björkroth, J. P.; Leo, A. *J. Pharm. Sci.* **1987**, *76*, 663.
(21) Hansch, C.; Kim, D.; Leo, A.; Novellino, E.; Silipo, C.; Vittoria, A. *CRC Crit. Rev.* **1989**, *19*, 185.
(22) Hansch, C. *Drug Metab. Rev.* **1972**, *1*, 1.
(23) Sargent, N.; Upshall, D.; Bridges, *J. Biochem. Pharmacol.* **1982**, *31*, 1309.

Table III. Selectivity Indices of 2,4-Diamino-5-(substituted-benzyl)pyrimidines Designed via QSAR 2 and 8

no.	X	log $(1/K_i)$ *E. coli*	log $(1/K_i)$ chicken	SI *E. coli*	SI *L. casei*	log *P*
1	3,4,5-$(OCH_3)_3$	8.08[a]	3.98[a]	4.10[d]	2.90	−1.55[e]
2	3,5-$(OCH_3)_2$, 4-$OCH_2CH_2OCH_3$	8.35[a]	3.64[a]	4.71[d]	2.62	−1.81[e]
3	3-C_2H_5, 4,5-$(CH_2OH)_2$	7.96[b]	3.58[c]	4.28[d]	3.13	−2.23[f]
4	3-C_2H_5, 4-CH_2OH, 5-$NHCOCH_3$	8.02[b]	3.15[c]	4.87[d]	3.36	−2.17[f]

[a] Observed values. [b] Calculated with eq 8. [c] Calculated with eq 2. [d] log $(1/K_i)$ of *E. coli* − log $(1/K_i)$ of chicken. [e] Measured in 0.1 N HCl/octanol. [f] Calculated values.

possible. There is no way to forsee the *positive* attributes of the steric effect of small polar substituents on bacterial DHFR, an effect the basis of which we still do not comprehend. Although we have some idea about what the magnitude of a coefficient with a π term might be, from our graphics capabilities our ideas are very rudimentary.[11] However, study of graphics alone provides no idea of the electronic effect of substituents normally modeled by Hammett constants; although, recently molecular orbital calculations are showing promise for this purpose.[24,25] Graphics analysis without data and without QSAR is difficult to interpret.

The results from this study as well as our earlier ones demonstrate that in formulating a QSAR major geometrical features of a receptor and the shape of the evolving QSAR with respect to all reasonable possibilities must be considered. The simplest case is that of substituent interaction with a surface as shown in Figure 1. This can evolve in complexity as one attaches other surfaces until, with five surfaces, a box (or pocket) is obtained. Any of these may be encountered in nature and the limits of a simple surface (one side), a cleft surface (three sides), a cleft with a bottom surface (4 sides), or a pocket (5 sides) will profoundly affect the shape of the final QSAR. To attain a satisfactory QSAR, a "prepared mind" must be in a "loop". Clues to the shape of the ultimate QSAR emerge from a study of the residuals from the elementary QSAR. In the analysis of residuals it is important to use a regression program which has an option for listing the difference between calculated and observed activity (residuals) in increasing or decreasing order. It is here that one can see trends of certain classes of substituents which do not fit the correct mode. For example, polar meta substituent may be badly fit because they have been "forced" into hydrophobic space. Or bulky substituents may be badly fit, revealing that steric effects are substantial.

The bilinear model (used with MR, π, or other suitable parameters) can be extremely helpful in exploring the limits of the receptor in two ways; when the slope of the right side of the bilinear part of the QSAR is essentially zero one knows that at the break points larger substituents are extending beyond the enzyme surface into aqueous space. On the other hand, if the slope of the right side of the bilinear model is steep, a steric effect is probably involved. At some point, the constraints of a hydrophobic pocket will begin to limit the optimum benefit of the hydrophobic interaction of large substituents. Sometimes, as in the present study, this problem can be addressed by the combination of π and MR terms.

The results in this report clearly confirm our finding with *L. casei* DHFR that the phenyl ring has two possible binding modes which can result from 180° rotation. Our QSAR demonstrates that such rotation, which is not restricted energywise by resonance interaction of the phenyl ring with the pharmacophoric function, allows hydrophobic substituents to contact hydrophobic space and hydrophilic substituents to contact hydrophilic space to maximize binding potential. While this optimization of hydrophobic and hydrophilic contact is most readily seen with meta-substituted phenyl rings, there are undoubtly other types of symmetry elements in ligands which afford ambivalent binding possibilities. Thus, in developing a QSAR from a complex set congeners, one needs to consider early on in the study elements of symmetry not precluded by internal energy restrictions within the ligand which may allow polar and nonpolar elements to bind in more than one way.

As pointed out previously, the two QSAR can be used to maximize selectivity in the design of new congeners in an ongoing research program. Table III compares selectivity indices (SI) for the clinically important trimethoprim and tetroxoprim with two congeners previously considered for *L. casei* DHFR. Congener 4 is calculated to be as potent as TMP but is projected to be 6 times as selective. In the case of *L. casei*, it is expected to be 3 times as selective as TMP. In these examples only substituents which have been carefully studied (Table I) have been used in the design of the new congeners. That is, no attempt has been made to extend beyond spanned substituent space. Many other possibilities can readily be designed via eqs 2 and 8.

Following traditional drug-modification methodology, one is unlikely to make congeners containing a hydrophobic substituent in one meta position and a hydrophilic substituent in the other meta position, since such compounds are usually much more difficult to synthesize. While the predictions obtained with these or other such QSAR cannot be expected to be perfect (each QSAR is "80% correct"), the results are likely to be considerably better than intuitively guided synthesis. Moreover, as the synthesis and testing program progresses, this should firm up the QSAR for still better projections. Equations 1, 2, and 8 are quite complex so that it is easy to see why it has been so difficult to devise new congeners more selective than TMP.

Clearly the parameterization of a set of congeners substituted in several different ways is not a routine matter but requires considerable thought and depends on experience for success. The final arbitrator is not statistics, important as they are, but the overall feeling of the researcher that the model be viable. It should encompass knowledge about receptor binding, the underlying physical-organic chemistry, and past QSAR models. It is imperative that the new QSAR model conform with other known and appropriate QSAR at that point in time. Thus, while our QSAR for *L. casei* DHFR seemed reasonable with respect to the above criteria, it becomes far more viable when examined in light of the present results with *E. coli* DHFR.

Registry No. 1, 36821-85-9; 2, 36821-85-9; 3, 738-70-5; 4, 56066-63-8; 5, 56518-41-3; 6, 68902-57-8; 7, 69194-91-8; 8,

(24) Shusterman, A. J.; Johnson, A. S.; Hansch, C. *Int. J. Quant. Chem.* **1989**, *36*, 19.

(25) Compadre, R. L.; Debnath, A. K.; Shusterman, A. J.; Hansch, C. *Environ. Mol. Mutagen.* **1990**, *15*, 44.

73090-70-7; **9**, 78025-72-6; **10**, 77113-54-3; **11**, 20344-69-8; **12**, 100515-03-5; **13**, 107698-02-2; **14**, 107698-00-0; **15**, 100515-04-6; **16**, 71525-05-8; **17**, 69945-54-6; **18**, 73356-41-9; **19**, 13932-40-6; **20**, 5355-16-8; **21**, 83166-76-1; **22**, 107698-03-3; **23**, 73356-40-8; **24**, 107697-99-4; **25**, 83158-06-9; **26**, 107698-04-4; **27**, 78233-99-5; **28**, 50823-96-6; **29**, 98612-09-0; **30**, 98612-08-9; **31**, 80407-58-5; **32**, 77113-56-5; **33**, 77113-58-7; **34**, 77113-57-6; **35**, 77113-55-4; **36**, 80416-29-1; **37**, 59481-28-6; **38**, 69945-57-9; **39**, 69945-56-8; **40**, 69945-58-0; **41**, 69945-59-1; **42**, 50823-94-4; **43**, 77113-61-2; **44**, 30077-60-2; **45**, 77113-63-4; **46**, 69945-60-4; **47**, 77113-62-3; **48**, 80407-62-1; **49**, 77113-60-1; **50**, 69945-50-2; **51**, 69945-53-5; **52**, 107698-01-1; **53**, 30077-67-9; **54**, 80407-59-6; **55**, 69945-52-4; **56**, 20285-70-5; **57**, 836-06-6; **58**, 69945-51-3; **59**, 46726-70-9; **60**, 18588-43-7; **61**, 69945-55-7; **62**, 49561-94-6; **63**, 77113-59-8; **64**, 49873-11-2; **65**, 80407-61-0; **66**, 80407-60-9; **67**, 93317-64-7; **68**, 7319-45-1; DHFR, 9002-03-3.

Absolute Structure–Cytotoxic Activity Relationships of Steganacin Congeners and Analogues[1]

Kiyoshi Tomioka,† Tsuneo Ishiguro,† Hidemichi Mizuguchi,† Nobuyasu Komeshima,† Kenji Koga,*,† Shigeru Tsukagoshi,‡ Takashi Tsuruo,‡ Tazuko Tashiro,‡ Seiichi Tanida,§ and Toyokazu Kishi§

Faculty of Pharmaceutical Sciences, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan, Cancer Chemotherapy Center, Cancer Institute, Kamiikebukuro, Toshima-ku, Tokyo 170, Japan, and Central Research Divisions, Takeda Chemical Industries, Ltd., 17-85 Jusohonmachi 2-chome, Yodogawaku, Osaka 532, Japan. *Received January 19, 1990*

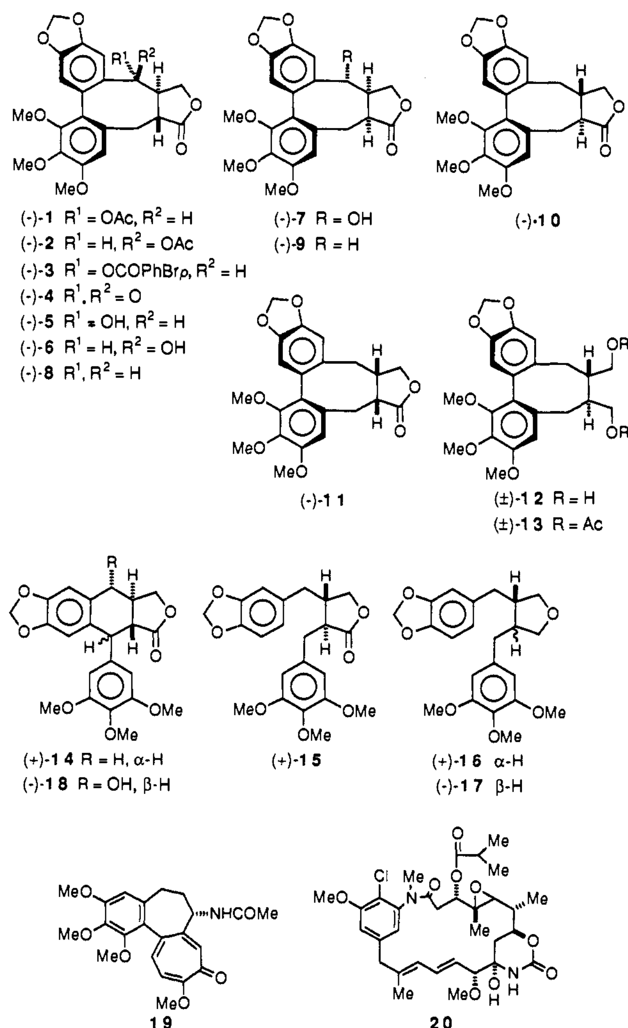
The cytotoxic activities of optically pure and racemic steganacin congeners and analogues against KB cells in culture and the inhibitor activity of cilia regeneration in *Tetrahymena* were studied with regard to absolute and relative configurations. The stereochemical requirements of dibenzocyclooctadiene lignan lactones for activity were clarified.

Steganacin [(−)-1], a naturally occurring dibenzocyclooctadiene lignan lactone, has been reported to show significant antitumor activity in vivo against P388 leukemia in mice and in vitro against cells derived from a human carcinoma of the nasopharnyx (KB).[2] It has been suggested that steganacin, like other spindle poisons, such as the ansamitocins (maytansine),[3] colchicine,[4] and podophyllotoxin,[5] exerts its antimitotic activity through an effect on spindle microtubules.[6]

Absolute structure–activity relationships of chiral compounds have been focus in recent medicinal chemistry. We have been involved in recent years with the asymmetric total synthesis of lignans and have found that the absolute configuration of natural (−)-steganacin had to be corrected and drawn as (−)-1, contrary to the antipodal structure proposed by Kupchan.[7] Since all of the possible optically pure enantiomers of steganacin congeners and analogues can be prepared by using the asymmetric synthesis we developed, the study of structure–activity relationships in enantiomers allow us to obtain greater insight into the structural requirements for antitumor activity. We report here that the correct absolute configuration around the pivotal bond and that the orientation of the lactone carbonyl are critical for expression of the antitumor activity of dibenzocyclooctadiene lignan lactones.

The dibenzocyclooctadiene compounds (±)-, (−)-, and (+)-steganacin (**1**),[7] episteganacin (**2**),[7] stegane (**8**),[7,8] picrostegane (**9**),[8] isostegane (**10**),[8] isopicrostegane (**11**),[8] (±)- and (−)-steganol (**5**),[7] (−)-[(*p*-bromobenzoyl)oxy]stegane (**3**),[7] (−)-steganone (**4**),[7] (−)-episteganol (**6**),[7] and (−)-picrosteganol (**7**)[9] (Chart I) were prepared as described previously. Isostegane derivatives (±)-**12**, **13**,[8] (±)- and (−)-isodeoxypodophyllotoxin (**14**),[10] (±)-, (−)-, and (+)-deoxypodorhizone (**15**),[7] (−)- and (+)-burseran (**16**), and (−)-**17**,[11] were also prepared as described. Steganacin derivatives (−)-**21**–**25** (Chart II) were prepared from (−)-**5**,[7] respectively, as described in the Experimental Section.

Chart I



Natural (−)-podophyllotoxin (**18**) and (−)-colchicine (**19**) were purchased from Aldrich Chemical Co. and purified

† University of Tokyo.
‡ Cancer Institute.
§ Takeda Chemical Industries Ltd.

 © 1991 American Chemical Society