Molecular Electrostatic Potential Analysis for Enzymatic Substrates, Competitive Inhibitors, and Transition-State Inhibitors

Carey K. Bagdassarian,[†] Vern L. Schramm,*[,‡] and Steven D. Schwartz*[,†,‡]

Contribution from the Departments of Physiology and Biophysics and of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461

Received August 14, 1995[⊗]

Abstract: Recent advances in the application of kinetic isotope effects to enzyme-catalyzed reactions have provided reliable information for enzymatic transition state structures. A method is presented for quantifying the similarity of substrates and inhibitors with their enzyme-stabilized transition states. On the basis of transition-state stabilization theory for enzymatic reactions, molecules most similar to the transition state structure bind with greatest affinity. Molecular similarity measures are applied to compare substrates, competitive inhibitors, and transition state inhibitors with the transition state structures stabilized by the enzymes AMP deaminase, adenosine deaminase, and AMP nucleosidase. (*R*)- and (*S*)-Coformycin 5′-phosphate are inhibitors for AMP deaminase, with the *R*-species superior to its enantiomer. Formycin 5′-phosphate 4-aminopyrazolo[3,4-*d*]pyrimidine-1-ribonucleotide, and tubercidin 5′-phosphate inhibit AMP nucleosidase. The transition state for adenosine deaminase is analogous to that for AMP deaminase, allowing analysis of the tight-binding hydrate of purine ribonucleoside and of a weaker inhibitor, 1,6-dihydropurine ribonucleoside. The basis for ranking molecules for similarity to the transition state is the distribution of electrostatic potential at the molecular van der Waals surface. Spatial properties of a molecule are described through the topography of the surface, while the electrostatics capture ionic, hydrogen-bonding, and hydrophobic features. A test molecule is compared with the transition state by orienting the two species so that their van der Waals surfaces are maximally coincident. At this orientation, a single measure sensitive both to the electrostatic potential and its spatial distribution is used to rank the electronic similarity. For AMP deaminase, adenosine deaminase, and AMP nucleosidase, the transition state inhibitors are quantitatively more similar to the transition states than are the substrates. A strong correlation between the binding free energies and the similarity measures is found for most of the transition-state inhibitors in all three enzyme systems. This method is useful in the logical design of transition state inhibitors and may be applied to similarity searches of chemical libraries.

Introduction

The pathways of enzymatically catalyzed reactions involve geometrical and electrostatic distortion of substrate molecules into transition state configurations which lead to the reaction products. The binding constant for the transition state is greater than that for the substrate by a factor reflecting the rate enhancement over the uncatalyzed reaction, typically $10^{10}-10^{15}$. Because the transition state structure experiences tight binding to the catalytic site, chemically stable transition state analogs prove to be extremely effective enzyme inhibitors.[1] These mimics are bound in the catalytic site with interactions analogous to those used for transition-state stabilization. Thus, the best enzyme inhibitors are those most similar to the transition state structure.

The *de novo* design of transition state analogs relies on accurate models of enzyme-stabilized transition states. These have recently been made available by the application of quantitative kinetic isotope effect methods.[2] Inhibitor design can be augmented by a procedure for quantifying the similarity of any putative inhibitor to the target transition state structure. This paper reports on the development and the practical application of an intuitive and robust method to determine numerically the similarity of substrates and inhibitors to experimentally determined enzymatic transition state structures. The method promises to be useful for inhibitor design.

This comparison of stable molecules with transition state structures assumes that molecular shape and electrostatic potential distribution are the primary parameters for molecular similarity.[2] The initial step requires the spatial positioning of the mass of the compared molecules so that their physical surfaces are as coincident as possible. This is done by placing the center of the transition state at the origin of the coordinate system. The test molecule is likewise positioned and is rotated about the origin until the geometrical similarity with the transition state is maximized. Once this geometrically determined orientation is found, a second parameter, sensitive to *both* geometrical and electrostatic distributions, is used to score the resultant similarity.

The method is applied to the enzyme systems yeast AMP deaminase, mammalian adenosine deaminase, and bacterial AMP nucleosidase. Yeast AMP deaminase catalyzes the hydrolytic deamination of AMP or dAMP to ammonia and the respective deaminated nucleotide or deoxynucleotide. Similarities to the experimentally deduced transition states are determined for (*R*)-coformycin 5′-phosphate and its 2′-deoxy analog, both of which are transition state inhibitors, and for the substrates AMP and dAMP. The *S*-stereoisomer of 2′-deoxycoformycin 5′-phosphate, a weak inhibitor for AMP deaminase, is also compared to the transition state. Adenosine deaminase catalyzes the hydrolytic deamination of adenosine to inosine, allowing for study of its inhibitors 1,6-dihydropurine ribonucleoside, the 1,6-hydrated purine ribonucleoside, and (*R*)-coformycin and of the substrate adenosine. 1,6-Dihydropurine

* Address correspondence to these authors.
[†] Department of Physiology and Biophysics.
[‡] Department of Biochemistry.
[⊗] Abstract published in *Advance ACS Abstracts,* September 1, 1996.
(1) Wolfenden, R. *Acc. Chem. Res.* **1972**, *5*, 10−18.
(2) Schramm, V. L.; Horenstein, B. A.; Kline, P. C. *J. Biol. Chem.* **1994**, *269*, 18259−18262.

S0002-7863(95)02781-8 CCC: $12.00

ribonucleoside, a competitive inhibitor, is less potent than coformycin while the hydrate is even more tightly bound. Though both AMP deaminase and adenosine deaminase catalyze analogous reactions, these two enzymes have different binding constants for analogous inhibitors and are treated as two separate systems. AMP nucleosidase hydrolyzes the N-ribosidic bond of AMP, yielding adenine and ribose 5′-phosphate. The transition state for AMP nucleosidase is compared to the transition state inhibitor formycin 5′-phosphate, to 4-aminopyrazolo[3,4-*d*]pyrimidine-1-ribonucleotide, to tubercidin 5′-phosphate, and to the substrate AMP. The binding affinities for these four molecules decrease in the order listed. The K_i values for the inhibitors of these three enzymes are detailed below, with a discussion of the relevant structural and electrostatic features. In all three enzymatic examples, the similarities to the transition states for transition state inhibitors and substrates are correlated with the binding free energies.

Two molecules are similar if they share close geometries, created through atomic arrangement, and chemical features such as hydrophobicity, nucleophilicity, electrophilicity, and hydrogen-bond-donor−acceptor sites. These features can therefore be used to group like molecules.[3]

A well-known family of molecular similarity measures is based upon comparison of the electron density or electrostatic potential that is generated by the molecules of interest. The comparison may include all space within the van der Waals surfaces of the molecules or may be restricted to all regions outside the molecular cores. For a comprehensive review of the use of molecular electrostatic potentials for the comparison and description of stable molecules, see ref 4 and the work of Carbó *et al*.,[5] Bowen-Jenkins *et al*.,[6] Hodgkin and Richards,[7,8] Petke,[9] and Richard.[10]

In contrast, the chemical descriptor featured in the present work is the molecular electrostatic potential calculated at the van der Waals *surface* of the molecule. For molecules described by quantum mechanics, all points in space have some electron density and a reasonable cutoff is required to define a geometry. This surface is usually defined as the 95% surface, i.e., 95% of the electron density is contained within it. The electrostatic potential at this van der Waals surface of the target molecule provides a reasonable description of the hydrogen-bond-donor and -acceptor sites an enzyme will encounter. The electrostatic potential at any point on the van der Waals surface is defined by all electronic and nuclear charge interactions with a positive unit test charge placed at that point. Thus, this single descriptor contains information about the entire electronic makeup of the molecule and defines regions of nucleophilicity and electrophilicity on the molecular surface.[11,12] Hydrogen-bonding sites and hydrophobic regions are also revealed, along with the pockets used in macromolecular interactions (protein−protein, nucleic acid−ligand, antibody−antigen).[13,14] An assumption central to this work is that transition state analogs share the same mechanisms for energy stabilization in the catalytic site as do the transition state structures themselves. It has been proposed, for example, that serine proteases lower the transition state energy through interactions largely of electrostatic origin,[15] an observation made via analysis of electrostatic potential surfaces.

Other investigators have used surface properties for molecular comparison in a variety of ways. For example, Chau and Dean[16,17] have projected electrostatic potentials onto a sphere enclosing a molecule. Gasteiger *et al*.[18] reduced, using a neural network, molecular electrostatic potential distributions to two-dimensional maps. Most recently, discrete points defining the van der Waals surface of a molecule have been collapsed onto a single autocorrelation vector:[19] ensuing similarity comparisons are via these vectors for a set of molecules. Lipophilicity/hydrophobicity patterns can be displayed on surfaces,[20] and Mezey and co-workers have developed group-theoretic methods to study electrostatic potential distributions on van der Waals surfaces.[21−23] The generation of electrostatic potentials remains a computational challenge.[24] (For different theories of molecular comparison, see ref 25−27.) In the approach described here, molecular properties are directly compared without mapping of molecular surfaces to less complex structures. As will become clear, a region on one molecule will be scored for similarity to a *neighborhood* around the analogous region of the other molecule. As such, the present scheme provides a more global similarity comparison. The method is simple to implement and the similarity measures are sensitive simultaneously to electrostatic and geometrical features. Most importantly, the formalism is used to quantify for the first time the similarity of substrates and transition state inhibitors to experimentally defined enzyme-stabilized transition state structures.[2] Application of this method is of significance for enzymatic transition state theory, inhibitor design, and for computational screening of chemical libraries.

Electrostatic potentials calculated at the van der Waals surfaces have been used effectively in the *qualitative* study of the substrates, transition states, and transition state-like inhibitors of the enzymes AMP deaminase,[28] AMP nucleosidase,[29] nucleoside hydrolase,[30] and purine nucleoside phosphorylase.[31] Of these, AMP deaminase and AMP nucleosidase provide two test cases for the quantitative measures of molecular similarity. These enzymes bind their respective transition state analogs, (*R*)-coformycin 5′-phosphate and formycin 5′-phosphate, 3×10^7 and 2600 times tighter than they do their common substrate

(3) Perelson, A. S.; Oster, G. F. *J. Theor. Biol.* **1979**, *81*, 645−670.
(4) Tasi, G.; Palinko, I. *Topics Curr. Chem.* **1995**, *174*, 45−71.
(5) Carbó, R.; Leyda, L.; Arnau, M. *Int. J. Quantum Chem.* **1980**, *17*, 1185−1189.
(6) Bowen-Jenkins, P. E.; Cooper, D. L.; Richards, W. G. *J. Phys. Chem.* **1985**, *89*, 2195−2197.
(7) Hodgkin, E. E.; Richards, W. G. *Int. J. Quantum Chem.: Quantum Biol. Symp.* **1987**, *14*, 105−110.
(8) Richards, W. G.; Hodgkin, E. E. *Chem. Br.* **1988**, *24*, 1141−1144.
(9) Petke, J. D. *J. Comput. Chem.* **1993**, *14*, 928−933.
(10) Richard, A. M. *J. Comput. Chem.* **1991**, *12*, 959−969.
(11) Sjoberg, P.; Politzer, P. *J. Phys. Chem.* **1990**, *94*, 3959−3961.
(12) Politzer, P., Truhlar, D. G., Eds. *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Plenum Press: New York, 1981.
(13) Weiner, P. K.; Langridge, R.; Blaney, J. M.; Schaefer, R.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 3754−3758.
(14) Novotny, J.; Sharp, K. *Prog. Biophys. Molec. Biol.* **1992**, *58*, 203−224.
(15) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629−3637.
(16) Chau, P.-L.; Dean, P. M. *J. Mol. Graph.* **1987**, *5*, 97−100.
(17) Dean, P. M.; Chau, P.-L. *J. Mol. Graph.* **1987**, *5*, 152−158.
(18) Gasteiger, J.; Li, X.; Rudolph, C.; Sadowski, J.; Zupan, J. *J. Am. Chem. Soc.* **1994**, *116*, 4608−4620.
(19) Wagener, M.; Sadowski, J.; Gasteiger, J. *J. Am. Chem. Soc.* **1995**, *117*, 7769−7775.
(20) Heiden, W.; Moeckel, G.; Brickmann, J. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 503−514.
(21) Mezey, P. G. *Int. J. Quantum. Chem.: Quant. Biol. Symp.* **1987**, *14*, 127−132.
(22) Mezey, P. G. *Shape in Chemistry: An Introduction to Molecular Shape and Topology*; VCH Publishers: New York, 1993.
(23) Arteca, G. A.; Hernandez-Laguna, A.; Randez, J. J.; Smeyers, Y. G.; Mezey, P. G. *J. Comput. Chem.* **1991**, *12*, 705−716.
(24) Alkorta, I.; Villar, H. O.; Arteca, G. A. *J. Comput. Chem.* **1993**, *14*, 530−540.
(25) Johnson, M. A.; Maggiora, G. M., Eds. *Concepts and Applications of Molecular Similarity*; Wiley-Interscience: New York, 1990.
(26) Dean, P. M. *BioEssays* **1994**, *16*, 683−687.
(27) Hopfinger, A. J. *J. Am. Chem. Soc.* **1980**, *102*, 7196−7206.
(28) Kline, P. C.; Schramm, V. L. *J. Biol. Chem.* **1994**, *269*, 22385−22390.
(29) Ehrlich, J. I.; Schramm, V. L. *Biochemistry* **1994**, *33*, 8890−8896.
(30) (a) Horenstein, B. A.; Schramm, V. L. *Biochemistry* **1993**, *32*, 7089−7097. (b) Horenstein, B. A.; Schramm, V. L. *Biochemistry* **1993**, *32*, 9917−9925.
(31) Kline, P. C.; Schramm, V. L. *Biochemistry* **1995**, *34*, 1153−1162.

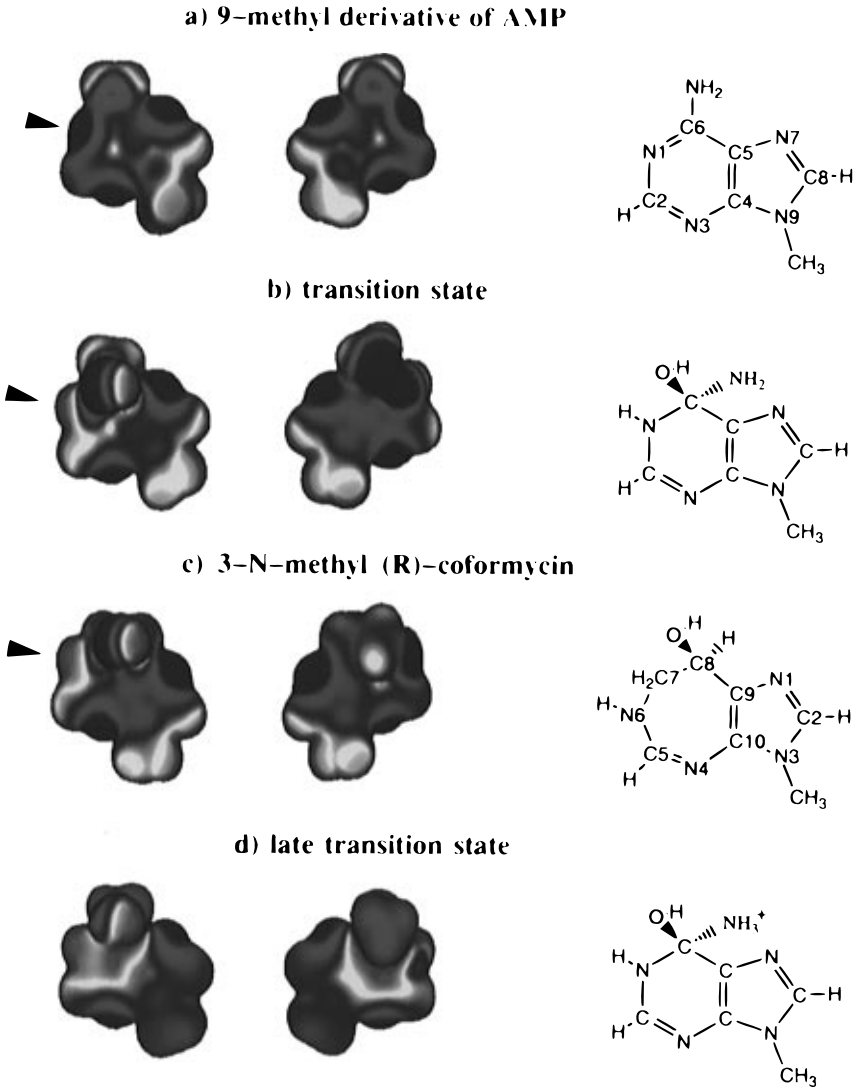


Figure 1. Molecular electrostatic potential surfaces for model substrate, transition states, and transition state inhibitor of the AMP deaminase reaction. All molecules substitute a methyl in place of β-D-ribose 5′-phosphate, which is constant and unchanged in the AMP deaminase reaction. The color images on the left are front and rear views of substrate (a), transition state (b), transition state inhibitor (c), and a late transition state characterized by the presence of the NH_3^+ leaving group (d). (*S*)-Coformycin is not shown but is the mirror image of the *R*-form. The arrows point to the N-1 locus on the methyl derivative of AMP, on the transition state, and to the N-6 locus on the inhibitor. The color spectrum from red to blue is in the direction of decreasing relative positive charge. The structures on the right correspond to the orientations of the left-most molecular electrostatic surfaces.

AMP. The implication is that the inhibitors resemble more closely the appropriate transition states than do the substrates, the criterion of similarity being a comparison of the molecular electrostatic potential surfaces. A visual confirmation was offered in the above references. Features of AMP deaminase and AMP nucleosidase provide two of the few enzyme systems where experimental evidence is available, from kinetic isotope effects, to establish the nature of the transition states. Two additional inhibitors for the AMP nucleosidase system, 4-aminopyrazolo[3,4-*d*]pyrimidine-1-ribonucleotide and tubercidin 5′-phosphate, will be considered. Adenosine deaminase is also included in this study, with its inhibitors (*R*)-coformycin, 1,6-dihydropurine ribonucleoside, and 1,6-hydrated purine ribonucleoside. The measures of molecular similarity described here are broadly applicable since both stable molecules and transition state constructs are readily accommodated. Although this work emphasizes the comparison of substrates, transition states, and known transition state inhibitors, it will be apparent that the algorithm should be equally applicable for analyzing proposed inhibitors for a given enzyme-stabilized transition state.

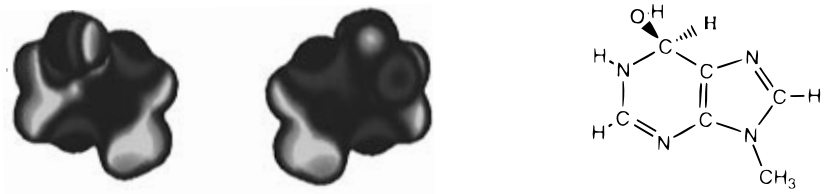
Electrostatic Potential Surfaces for Substrates, Transition States, and Transition State Inhibitors

This section defines the features to which a quantitative scheme for transition state comparisons must be sensitive, using two distinct enzyme chemistries. For both AMP deaminase and AMP nucleosidase, experimental models for the transition state structures were derived from multiple kinetic isotope effects.[32−34] The transition state for adenosine deaminase is modeled upon that for AMP deaminase, based on the near-identity of the amino acids which constitute their catalytic sites.[32]

AMP Deaminase. Stick figures for AMP, for its transition state with AMP deaminase, for a protonated late transition state, and for (*R*)-coformycin 5′-phosphate are shown in Figure 1. The early and late transition states are separated along the reaction coordinate by an unstable intermediate. Enzymatic hydroxyl-

(32) Merkler, D. J.; Kline, P. C.; Weiss, P.; Schramm, V. L. *Biochemistry* **1993**, *32*, 12993−13001.
(33) Mentch, F.; Parkin, D. W.; Schramm, V. L. *Biochemistry* **1987**, *26*, 921−930.
(34) Parkin, D. W.; Mentch, F.; Banks, G. A.; Horenstein, B. A.; Schramm, V. L. *Biochemistry* **1991**, *30*, 4586−4594.

a) 9−methyl derivative of hydrated purine ribonucleoside



b) 1,6−dihydropurine ribonucleoside

Figure 2. Molecular electrostatic potential surfaces for hydrated purine ribonucleoside (a) and 1,6-dihydropurine ribonucleoside (b). These two inhibitors, along with the methyl derivatives of adenosine, (*R*)-coformycin, and the transition state as shown in Figure 1, complete the adenosine deaminase series.

ation at C-6 and protonation at N-1 of AMP represent the highest energetic barrier to catalysis and yield the early transition state structure. It is the early transition state which serves as the target species for inhibitor design and for further comparison with the substrate and the hypothetical late transition state. In this series,[28] the structures as shown have methyl substituents replacing the ribose 5′-phosphate groups at the N-9 positions of AMP, the transition state, and the late transition state and at the N-3 position of (*R*)-coformycin 5′-phosphates. The ribose 5′-phosphate residues are constant for all four entities and are assumed to affect similarly the electrostatic potentials and geometries of the four molecules. The late transition state is derived from the early transition state by increasing the bond order to the attacking hydroxyl nucleophile to unity and by protonation of the leaving group to NH_3^+.[32] While the methyl derivatives of AMP, the transition state, and coformycin are neutral species, the late transition state is more positively charged. This species is thereby electrostatically dissimilar to the early transition state, a feature to be investigated by the similarity measures. The methyl derivative of (*S*)-coformycin is simply the mirror image of the *R*-species and is not shown. The molecular structures used in the calculations are energy minimized—except for bonds influenced by the transition state structure—with the 3-21G, or higher, basis set at the Hartree−Fock level, as discussed in the original work (though, there, STO-3G was used). Because (*R*)-coformycin is a tightly bound transition state mimic, it is assumed to bind with a molecular conformation similar to the transition state's: operationally the inhibitor's hydroxyl group is constrained to reflect this similarity. The simple schematics are given here for convenience.

The transition state structure upon which this work is based stems from catalytic activity of the enzyme on AMP. An inhibition constant for (*S*)-coformycin 5′-phosphate is unavailable to complete the AMP and (*R*)-coformycin 5′-phosphate series. The 2′-hydroxyl on the ribose ring is a significant feature of the recognition process by AMP deaminase,[35] but the similarity measures are calculated without the sugars. Therefore, the augmented set (*R*)-2′-deoxycoformycin 5′-phosphate, (*S*)-2′-deoxycoformycin 5′-phosphate, and 2′-deoxyAMP with K_i (or K_m for dAMP) values of 0.041 μM,[35] 40 μM,[35] and 1200 μM,[36] respectively, will be used in this work. The *R*-inhibitor is bound 1000 times as tightly as is the *S*-inhibitor, while the latter is held only 30 times more tightly than is the substrate dAMP. The dissociation constant for the transition state complex has been estimated from the K_m for dAMP and from the enzymatic rate enhancement over that of the solution phase reaction, based on values for AMP.[1,37,38] K_{TX} is 1.7×10^{-16} M and is a measure of how tightly the transition state is bound. It will be seen that the logarithms of these dissociation constants—that is, the (dimensionless) binding free energies—are correlated to the similarity scores of substrate or inhibitors to the transition state for AMP deaminase. In addition, AMP and (*R*)-coformycin 5′-phosphate, with $K_m = 300$ μM[36] and $K_i =$ 10 pM,[35] respectively, are compared to the transition state for AMP with $K_{TX} = 0.43 \times 10^{-16}$ M.[1,37,38]

Adenosine Deaminase. This enzyme facilitates a chemical hydrolysis (adenosine to inosine) analogous to the one catalyzed by AMP deaminase (AMP to inosine monophosphate). Because both enzymes have similar catalytic sites[35,36] and both are strongly inhibited by (*R*)-coformycin analogs,[38] it is assumed that the transition state for adenosine deaminase with adenosine as substrate is similar to the transition state for AMP deaminase in the region of reaction at C-6 and N-1. Therefore, for adenosine deaminse the methyl derivatives, as shown in Figure 1, of AMP (or, equally, of adenosine) and (*R*)-coformycin and of the transition state for AMP deaminase are used for similarity scoring. Again, the ribose residues of the transition state, of adenosine, and of the inhibitors are replaced by methyl groups. Two of the inhibitors for adenosine deaminase are (*R*)-coformycin (Figure 1c) and hydrated purine ribonucleoside (Figure 2a). There is evidence that purine ribonucleoside binds to adenosine deaminase as the 1,6-hydrate,[38−41] the structure shown in the figure, and this hydrate binds with most of the energy proposed for an ideal transition state analog. The orientation used here for its hydroxyl group is that which matches best that of the transition state's. The 1,6-dihydropurine ribonucleoside, also shown (Figure 2b), binds weakly with a dissociation constant of 5.4×10^{-6} M because it lacks the critical hydroxyl on C-6.

(35) Merkler, D. J.; Brenowitz, M.; Schramm, V. L. *Biochemistry* **1990**, *29*, 8358−8364.
(36) Merkler, D. J.; Schramm, V. L. *Biochemistry* **1993**, *32*, 5792−5799.
(37) Merkler, D. J.; Wali, A. S.; Taylor, J.; Schramm, V. L. *J. Biol. Chem.* **1989**, *264*, 21422−21430.
(38) Radzicka, A.; Wolfenden, R. *Methods Enzymol.* **1995**, *249*, 284−312.
(39) Shih, P.; Wolfenden, R. *Biochemistry* **1996**, *35*, 4697−4703.
(40) Wilson, D. K.; Rudolph, F. B.; Quiocho, F. A. *Science* **1991**, *252*, 1278−1284.
(41) Wilson, D. K.; Quiocho, F. A. *Biochemistry* **1993**, *32*, 1689−1694.

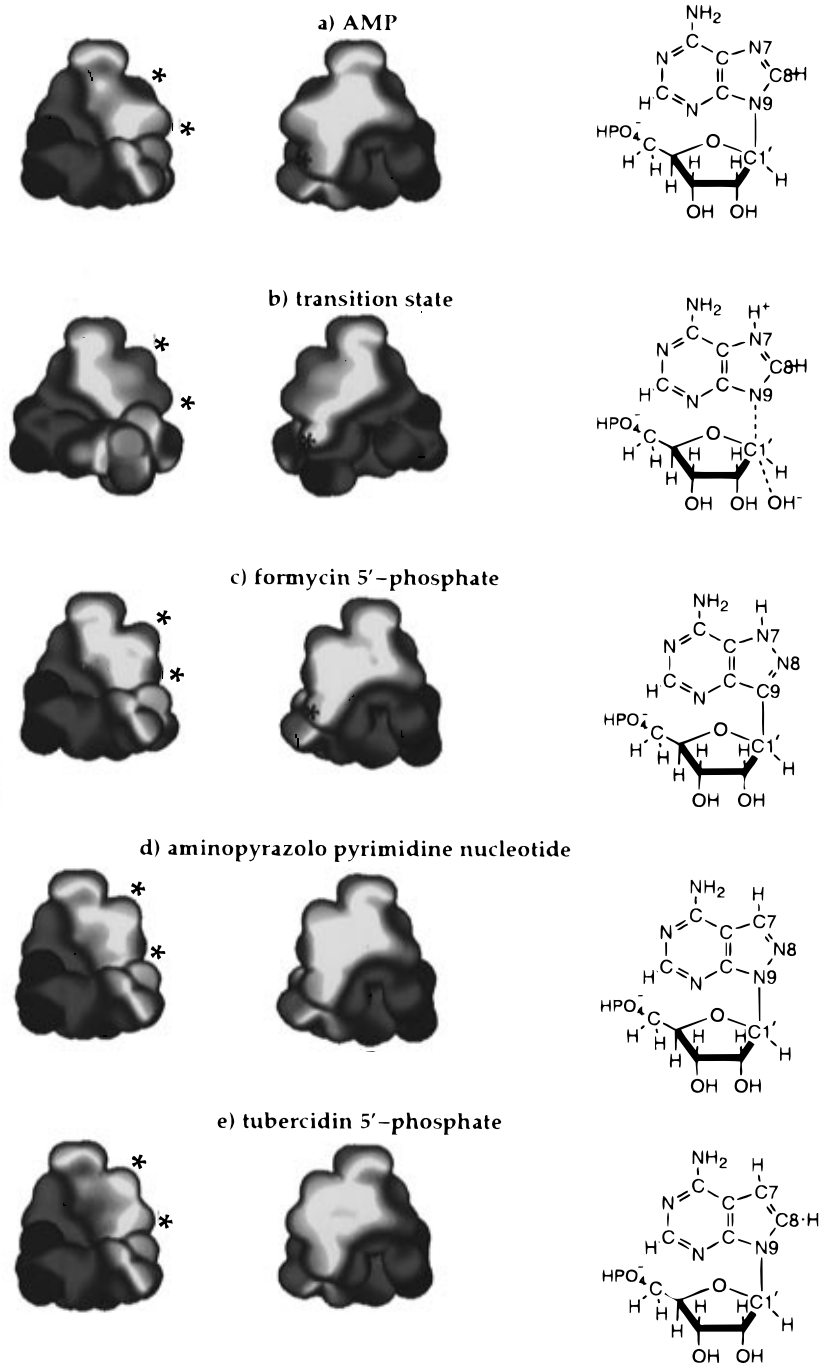


Figure 3. Molecular electrostatic potential surfaces for substrate (a), transition state (b), and transition state inhibitors formycin 5′-phosphate (c), 4-aminopyrazolo[3,4-*d*]pyrimidine-1-ribonucleotide (d), and tubercidin 5′-phosphate (e) of the AMP nucleosidase reaction. The left-hand colored images show the front and rear views of the molecular electrostatic potential surfaces. The black asterisks locate the N-7 locus on AMP, on the transition state, and on the inhibitors. The blue asterisks point to the C-8 locus on AMP, on the transition state, and to the corresponding atoms on the inhibitors. The asterisks on the rear views indicate a region of similarity between the transition state and the inhibitor. The color spectrum from red to blue is in the direction of decreasing relative positive charge. The column of structures on the right have the same orientations as the colored figures on the left.

The equilibrium binding constants for the adenosine deaminase series[38] are $K_{TX} = 1.5 \times 10^{-17}$ M for the transition state and $K_i = 3.0 \times 10^{-13}$ M for the hydrate of purine ribonucleoside, (*R*)-coformycin is less tightly bound with $K_i = 1 \times 10^{-11}$ M, 1,6-dihydropurine ribonucleoside is a weak inhibitor with $K_i = 5.4 \times 10^{-6}$ M, and $K_M = 3 \times 10^{-5}$ M for the substrate adenosine. Note that analogous molecules in the adenosine deaminase and AMP deaminase series have different binding affinities to their respective enzymes.

AMP Nucleosidase. Figure 3 shows AMP, the transition state for AMP nucleoside, formycin 5′-phosphate (the series studied by Ehrlich and Schramm[29]), and two additional inhibitors, 4-aminopyrazolo[3,4-*d*]pyrimidine-1-ribonucleotide and tubercidin 5′-phosphate. Note that the transition state is derived from AMP by protonation at the N-7 position and the inclusion of an attacking, though weakly bonded, hydroxyl nucleophile. The C-1′ to N-9 bond is partially broken to give a relatively low bond order of 0.2 at the transition state. In the AMP nucleosidase series, all five molecules are negatively charged through the phosphate residue.

The equilibrium constants are[42] $K_{TX} = 2 \times 10^{-17}$ M for the transition state, $K_i = 43 \times 10^{-9}$ M for formycin 5′-phosphate, $K_i = 10 \times 10^{-6}$ M for aminopyrazolopyrimidine ribonucleotide,

(42) DeWolf, W. E., Jr.; Fullin, F. A.; Schramm, V. L. *J. Biol. Chem.* **1979**, *254*, 10868−10875.
(43) Schramm, V. L.; Baker, D. L. *Biochemistry* **1985**, *24*, 641−646.

$K_i = 51 \times 10^{-6}$ M for tubercidin 5′-phosphate, and $K_M = 120 \times 10^{-6}$ M for AMP.

Molecular Electrostatic Potential Surfaces. After calculation of the electron densities and electrostatic potentials of these molecules using the CUBE function of the Gaussian 94 package,[44] the electrostatic potential surfaces were visualized with the AVS Chemistry Viewer (Advanced Visual Systems Inc. and Molecular Simulations Inc.). The restriction of the visualization procedure to the van der Waals surface is achieved by choosing on AVS the 0.002 electrons/bohr^3 isosurface. This contour contains at least 95% of the total electron density and corresponds well with the van der Waals surface.[11] The STO-3G basis set was judged to be adequate for calculation of the electrostatic potentials (though the geometry minimizations are done at a higher basis set). Some of the surfaces have been previously published and the reader is referred to the original references,[28,29] but similar surfaces are reproduced here for connection to the present work. Figure 1 is for the AMP deaminase series; Figure 2 is for adenosine deaminase to which must be added the methyl-substituted transition state, substrate, and (*R*)-coformycin of Figure 1 for completeness; and Figure 3 shows the AMP nucleosidase series. Each left−right pair of electrostatic pictures represents the front and back views of a molecule with respect to the plane of the purine ring. The most positive regions of electrostatic potential are represented as red and the most negative as blue, without regard to absolute magnitudes. This method of scaling the potentials permits molecules with a net charge, like the late transition state of AMP deaminase or the substrate, transition state, and inhibitors of the AMP nucleosidase series, to be observed with the appropriate color patterns on the surface. In these representations, molecular similarities are exposed via a comparison of the patterns of relative electropositive, neutral, and negative regions. This important point is expanded below as the algorithm for the quantitative similarity measure is defined.

To summarize the features of the transition states, the hydroxyl group at C-6 on the transition state of AMP in the deaminase reaction (Figure 1) is mimicked electrostatically and geometrically by the hydroxyl group at C-8 of the methyl derivative of (*R*)-coformycin, the inhibitor—note the similarity of the color patterns in Figure 1, parts b and c. Furthermore, the enzymatic protonation at N-1 in the transition state introduces positive charge which is reproduced by the hydrogen at N-6 of (*R*)-coformycin. The substrate AMP lacks these geometrical and electrostatic features and so is less related to the transition state in both shape and charge distribution than is the transition state inhibitor. AMP is therefore less tightly bound than is (*R*)-coformycin. (*S*)-Coformycin, which is not pictured, is not as good an inhibitor because its hydroxyl group at C-8 is in the *S*-configuration while the transition state's attacking nucleophile has *R*-stereochemistry. Indeed, the enzyme adenosine deaminase, which shares catalytic-site similarity to AMP deaminase,[35,36] binds (*R*)-deoxycoformycin 1.3×10^7 times tighter than it does (*S*)-deoxycoformycin.[43] The backside view of the late transition state shown in Figure 1d is different electrostatically than the backside of the proposed transition state (Figure 1b) because of the presence of the positively charged NH_3^+ group on the former. In addition, the late transition state is positively charged and thus bears less resemblance to the earlier species. In developing quantitative molecular comparisons, it is important to consider that the relative color schemes mask to a certain extent the differences between differently charged molecules. Note, however, that geometrically these two species are very similar.

For (*R*)-coformycin, adenosine, and the transition state in the adenosine deaminase series, the identical discussion given above holds. It was seen that the 1,6-hydrate of purine ribonucleoside is more tightly bound than is (*R*)-coformycin by a factor of 100. This inhibitor, with its hydroxyl on C-6 in the proper orientation and with N-1 protonated to mimic the transition state, is more similar to the transition state than is coformycin because the latter features a nonplanar seven-membered ring while the hydrate has a six-membered ring, as does the transition state. Indeed, the only difference between the hydrated purine ribonucleoside and the transition state is the inhibitor's lack of the NH_2 group on C-6. The electrostatic potential surface for this inhibitor shown in Figure 2 is most similar to the transition state's shown in Figure 1. If the hydroxyl group of the hydrate is replaced with a hydrogen, 1,6-dihydropurine ribonucleoside results (Figure 2b), which because of the missing OH, cannot be as good an inhibitor as coformycin or the hydrate. Crystallographic work[40,41] has shown that the hydroxyl group in the *R*-stereochemistry is necessary for coordination with an enzymatic zinc moiety.

In the AMP nucleosidase series (Figure 3), AMP, the transition state, formycin 5′-phosphate, the aminopyrazolopyrimidine ribonucleotide, and tubercidin 5′-phosphate are geometrically similar. One major difference appears in the ribose residue. AMP and the three inhibitors are shown with their sugar moieties in the same conformation, but this conformation is different from that of the ribose in the transition state. By comparing the inhibitors and AMP to the same transition state, the geometrical and electrostatic differences between the ribose residues of the inhibitors and the transition state will be the same as those between the sugars of AMP and the transition state. The presence of a proton on N-7 of AMP at the transition state introduces a partial positive charge which is absent in the substrate AMP. Formycin has a geometrically similar charge on its analogous N-7. The backside views at the positions of the asterisks demonstrate that formycin 5′-phosphate and the transition state share greater electrostatic similarities. However, the proton on C-8 of both AMP and the transition state causes a local electron deficiency. This position is occupied by N-8 of formycin 5′-phosphate, and there is a lack of positive charge as shown in the relative color scheme. Tubercidin and aminopyrazolopyrimidine ribonucleotide are both less tightly bound than is formycin, but are more so than is the substrate. Both of these inhibitors feature a proton on C-7, in a spatially analogous position to the proton on N-7 of the transition state. However, because the carbon-bound protons of these inhibitors are not as electropositive as the nitrogen-bound protons of the transition state and formycin, these inhibitors are less similar at this important locus to the transition state than is formycin. When C-8 of tubercidin is replaced with a nitrogen to give the aminopyrazolo compound, a slightly better inhibitor results since the hydrogen on C-7 is now more electropositive ($K_i = 51$ μM for tubercidin and 10 μM for aminopyrazolo). A goal of this work is to compare the features of the three inhibitor−transition state pairs to each other and to those of the AMP−transition state pair to see whether binding free energies can be correctly ranked on the basis of similarity comparisons with the transition state. This is accomplished by the application of an algorithm which incorporates the actual values of the electrostatic potentials.

(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; DeFrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision C.2*; Gaussian, Inc.: Pittsburgh, PA, 1995.

This work is a departure from all previous studies since the geometry of the transition state is experimentally derived from kinetic isotope effects. Thus, the transition state models reflect the action of the enzymatic groups in stabilizing a defined structure. However, because transition state inhibitors and the target transition states are proposed to have similar interactions with the enzyme, these explicit interactions between active-site moieties and the transition states or the inhibitors are omitted in the calculations of molecular electrostatic potential surfaces. Since enzymatic active sites are largely hydrophobic, bulk solvent water is excluded from the molecular surfaces involved. Molecules bearing strong resemblance to the transition state will be strong inhibitors and will be found with the present method. In conclusion, these gas phase calculations are expected to have sufficient predictive power for inhibitor design.

Calculation of Molecular Similarity

The comparison of substrates, transitions states, and inhibitors can be made over entire molecular surfaces or can emphasize only those atoms which are known to be in contact with the enzyme catalytic site. In most cases little is known about the details of enzymatic−transition state contacts; therefore, the global characteristics of a test molecule and the transition state are compared. (For systems with established transition-state interactions, see the work on adenosine deaminase[40,41] or cytidine deaminase[45].)

Gaussian 94 electron density and electrostatic potential calculations yield three-dimensional matrices of numbers for these quantities. In this work, a 64 × 56 × 56 collection of entries for each physical quantity is established to define the space around and within the molecule in the *x*-, *y*-, and *z*-coordinates. At each spatial point *i*, there is an associated electron density, ρ_i, and potential, ϵ_i. These entries cover points both within the molecular core and beyond the van der Waals molecular surface. To find a numerical description of the *surface* of a molecule, an algorithm was written to scan the electron density matrix and to calculate and record the spatial coordinates of all points having density $0.002 \pm \delta$ electrons/bohr^3. δ is the acceptance tolerance since no Gaussian output will have $\rho_i = 0.002$ exactly. By this procedure, the surface is defined by a collection of points, each point lying on the van der Waals surface and the *i*th one having coordinates (x_i, y_i, z_i). From the electrostatic potential matrix, ϵ_i is assigned to each point (x_i, y_i, z_i), with *i* constrained to be on the surface. As δ becomes larger, more points are accepted; however, if δ is too large, the surface becomes deformed via inclusion of points not on the physical van der Waals surface. Typically, δ is adjusted so that each atom of a molecule is represented on average by 17 surface points. This gives good surface representation, and the molecular shapes thus obtained are faithful reproductions, via *selected* surface points, of the surfaces shown in Figures 1−3. Recall that the original surfaces were made by analyzing the Gaussian data directly by AVS and choosing the proper isosurface without any knowledge of the coordinates of the surface being produced. The color scheme, used to identify each surface point, is created by ranking the values of the electrostatic potentials of the points from most positive to most negative. The color spectrum from red to blue is arbitrarily divided into 11 distinct colors, with the most positive surface points assigned red, the next group a reddish-orange, and so on. Thus, the surface color patterns shown in Figures 1−3 (defined by functions in the AVS algorithm) are reproduced by the selected surface points. The visualization is through the Geometry Viewer of AVS (discussed below).

(45) Betts, L.; Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, C. W., Jr. *J. Mol. Biol.* **1994**, *235*, 635−656.

A molecule can be compared to another either geometrically or electrostatically, but ideally, a similarity measure will contain a mixture of both. Consider first the measure

$$S_e = \frac{\sum_{i=1}^{nA}\sum_{i=1}^{nB}\epsilon_i^A\epsilon_j^B \exp(-\alpha r_{ij}^2)}{\sqrt{\sum_{i=1}^{nA}\sum_{j=1}^{nA}\epsilon_i^A\epsilon_j^A \exp(-\alpha r_{ij}^2)}\sqrt{\sum_{i=1}^{nB}\sum_{j=1}^{nB}\epsilon_i^B\epsilon_j^B \exp(-\alpha r_{ij}^2)}} \quad (1)$$

where ϵ_i^A is the electrostatic potential at surface point *i* of molecule A, ϵ_j^B defines point *j* of molecule B, and in the numerator r_{ij}^2 is the spatial distance squared between point *i* on A and *j* on B. *n*A and *n*B refer to the number of surface points on each molecule. The double summation is therefore over all possible interactions between points on the two molecules, and α is the length scale for the interaction between *i* and *j*. The numerator compares A to B for a particular orientation of molecule B relative to molecule A. The denominator serves as a normalization factor for the comparison of A to itself and for B to itself. Here, r_{ij}^2 refers to the distance between *i* and *j* on the same molecule. The distance between points is squared to decrease computation time. Consider also a second, purely geometrical measure:

$$S_g = \frac{\sum_{i=1}^{nA}\sum_{j=1}^{nB} \exp(-\alpha r_{ij}^2)}{\sqrt{\sum_{i=1}^{nA}\sum_{j=1}^{nA} \exp(-\alpha r_{ij}^2)}\sqrt{\sum_{i=1}^{nB}\sum_{j=1}^{nB} \exp(-\alpha r_{ij}^2)}} \quad (2)$$

The mathematical structures of S_e and S_g are similar to the alignment function used by Kearsley and Smith[46] and by Klebe et al.,[47] who build upon the former work and use the normalizing denominator. There is, however, a crucial difference: these workers sought to align molecules through matchings of atomic partial charges and van der Waals radii. Here, the electrostatic potential on and the geometry of the molecular van der Waals surfaces are compared. The philosophies are therefore quite different but are implemented within similar mathematical frameworks.

The measures S_e and S_g work in the following way. Assume that molecules A and B are superimposed in space with their "centers of mass" coincident and at the origin of the coordinate system. The molecular center of mass is calculated by assigning unit mass to each surface point and proceeding in the usual way to obtain the molecule's spatial center. Molecule A is fixed while B rotates and translates about its center of mass. For either measure, S_e or S_g, we locate a point *i* on A and sum on interactions with all points *j* on B—but with interaction strengths modulated by the exponential $\exp(-\alpha r_{ij}^2)$. For S_g this geometrical factor is the only interaction while S_e includes as well the product of the electrostatic potentials. The closer the point *j* to *i*, the more important its relative contribution. As α becomes smaller, a point will interact with a greater number of points on the second molecule. It is easy to choose the range of the parameter α to make physical sense, as will be discussed below.

Molecule B is allowed to reorient with respect to A, and it is clear that for each new orientation S_e and S_g will have different values. At some orientation S_e will be maximized, and at

(46) Kearsley, S. K.; Smith, G. M. *Tetrahedron Comput. Methodol.* **1990**, *3*, 615−633.

(47) Klebe, G.; Mietzner, T.; Weber, F. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 751−778.

another S_g will have reached its maximum value. These respective orientations are not in general the same since S_g lacks the electrostatic component of S_e. For the choice of either S_e or S_g, depending on the physical information to be explored, the maximum value of the measure will correspond to the best orientation of B with respect to A. For S_e this is easily seen as follows: if A and B each have regions of negative and positive electrostatic potentials, the numerator, and therefore S_e, is maximized when like-signed regions of A and B are placed close to each other in orientation space. Point *i* on A then interacts strongly with points on B with the same sign. If, on the other hand, both A and B have all negative points, then S_e will be maximized by aligning B relative to A such that A's more negative points are closer to B's, while the less negative points become preferentially superimposed. In terms of the color schemes used to represent the surfaces of the molecules, this means that B will be reoriented with respect to A until global maximum overlap of the same-colored regions is achieved: red on A will be placed near red on B. For S_g the argument is even easier: the best orientation of one molecule with respect to the other will align the two surfaces as coincidently as possible. In similar molecules, any given point on one molecule is close to surface points on the other, leading to a maximum in the measure through summation of relatively large exponential terms. It is clear that when the two molecules under consideration are identical and perfectly oriented, the value for either measure is unity. It is found that for the molecules studied here the most useful application of the above equations arises from an initial geometrical alignment via S_g and a subsequent similarity scoring with S_e *under the constraint of the initial optimal geometrical matching*.

A more quantitative way to see the utility of eqs 1 and 2, revealing also the importance of the denominators, is as follows. Fix the spatial position of molecule A. Imagine that the points on B are free to move, independently of each other and without conservation of B's geometry, and free to take on any electrostatic values. Of course, in actual application of eqs 1 and 2, the molecular shape and electrostatic nature of any molecule are fixed. If the points of B were free to rearrange in any fashion, how must they individually align relative to those of A for S_e to be maximum? Extremize through the usual construction:

$$\frac{\partial S_e}{\partial \vec{r}_k^{\,B}} = 0 \quad \text{and} \quad \frac{\partial S_e}{\partial \epsilon_k^B} = 0 \quad \text{for all } k$$

where $\vec{r}_k^{\,B}$ is the coordinate of the *k*th point on B, with $r_{ij}^2 = (\vec{r}_i^{\,A} - \vec{r}_j^{\,B})^2$. It becomes clear that there are at least two extrema: the case where B becomes identical with A and oriented with respect to it for perfect matching so that $S_e = 1$ and the case where B is geometrically identical to A but with $\epsilon_i^B = -\epsilon_i^A$ for all *i* and $S_e = -1$. Can there exist a molecule that when compared with A has a similarity measure greater than unity? Begin by aligning two copies of A perfectly such that $S_e = 1$. Now, multiply all electrostatic potentials on one of the replicas by a positive integer so that ϵ_i^A becomes $n\epsilon_i^A$ on that molecule. Clearly, the numerator of S_e increases but $S_e = 1$ because of the normalizing denominator. A second approach is also illuminating. Orient perfectly two copies of a molecule, but let $\epsilon_k^A \rightarrow \epsilon_k^A + \gamma$, without disruption of the geometry, for only the *k*th point on *one* them. In the limit of a small perturbation γ, we find

$$S_e = 1 - \frac{\gamma^2}{2Q^2}(Q - P^2)$$

to leading order in γ with

$$Q = \sum_i \epsilon_i^A \epsilon_i^A + \sum_{i \neq j} \sum_{j \neq i} \epsilon_i^A \epsilon_j^A \exp(-\alpha r_{ij}^{\,2})$$

and

$$P^2 = \sum_i \epsilon_i^A \epsilon_i^A \exp(-2\alpha r_{ik}^{\,2}) + \sum_{i \neq j} \sum_{j \neq i} \epsilon_i^A \epsilon_j^A \exp[-\alpha(r_{ik}^{\,2} + r_{jk}^{\,2})]$$

If $Q - P^2$ is positive, increasing or decreasing γ away from zero—making the two molecules dissimilar-decreases the similarity continuously from $S_e = 1$. Because the first term on the right hand side of Q is the only one not subject to exponential decay, it is dominant with $Q - P^2$ positive. Going to the large α limit provides the final demonstration that both S_g and S_e are maximized in comparisons between identical molecules. In this limit a point on molecule A will interact with a point on B only if the latter is superimposed spatially on the first. Assume that B has the same number of surface points as does the target A, but that they are not coincident with any of the target's. $S_e \cong 0$ and $S_g \cong 0$ as the numerators of eqs 1 and 2 are vanishingly small. Deforming one point of B to the position and corresponding electrostatic potential of a chosen target point will clearly increase the values of the two measures (the only measurable contribution being between the transposed B-point and the target A-point since α is large). Subsequent point-by-point distortions will increase both S_g and S_e incrementally to their maximum values of unity. Therefore, it is impossible for $S_e > 1$, and if B is different from A, $S_e < 1$. Because $S_e = 1$ is the maximum value achieved and decreasing values reflect comparison of less related molecules, we are confident that S_e (and S_g since the arguments for it are similar) can be used both to orient the molecules and to give a numerical value to the similarity.

As discussed in the Introduction, an often-used family of measures for molecular comparison[4−10] has been based on scoring the electron density or electrostatic potential at every point in space, without specialization to the molecular surfaces. For example, with $\epsilon(\vec{r})$ describing the potential at position $\vec{r}$, the numerator of eq 1 would take the form $\int \epsilon^A(\vec{r})\epsilon^B(\vec{r})\, d\vec{r}$ with the integration spanning all space. Note that these point-by-point integrations do not allow for a given spatial region of one molecule to explore the corresponding spatial *neighborhood* of the other. Adopting such a scheme to our surfaces would have limited the interaction of point *i* on one molecule to its *nearest* neighbor *j* on the other (for any chosen orientation). This procedure would not have been as sensitive to *regions* of electrostatic and geometric similarity. The exponential term in our measure is responsible for this regional sensitivity and also serves another important purpose. The points extracted from the Gaussian output do not create a molecular surface with uniform coverage at all regions: such imperfections are smoothed by this regional comparison.

It was shown that the electrostatic similarity measure is insensitive to the transformation $\epsilon_j^B \rightarrow n\epsilon_j^B$ of all the potentials on the B-points (but with conservation of the coordinates). This feature means that S_e looks for patterns in electrostatic potential distributions and not for magnitudes. (Reference 8 addresses this issue for the other measures described above.)

Two molecules can be oriented relative to each other and the similarity scored via (1) the purely geometric measure S_g or (2) the electrostatic S_e which contains a geometric component through the exponential, or (3) a weighted average of the two. For the molecules tested here, the best results were achieved by first orienting the two molecules with respect to each other using S_g alone. This achieves a good geometrical fit. Then,

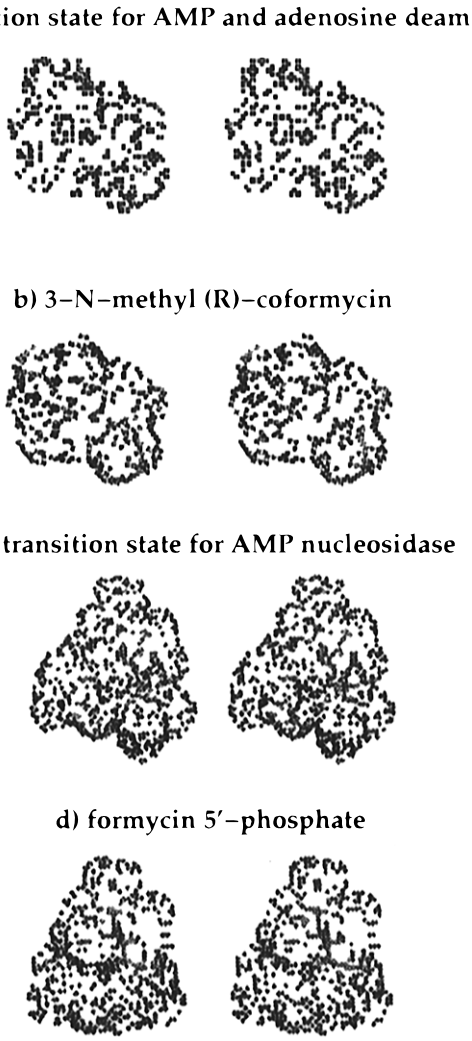


Figure 4. Surface-point stereoviews of the molecular electrostatic potential surfaces for the transition state for AMP or adenosine deaminase (a), (*R*)-coformycin (b), transition state for AMP nucleosidase (c), and for formycin 5′-phosphate (d). The color spectrum from red to blue is in the direction of decreasing relative positive charge. Surfaces for all other molecules are of similar quality. The transition state for AMP nucleosidase and its inhibitor are protonated on their phosphates.

for this geometrically optimized alignment, use S_e to measure the similarity. The ability of S_e to differentiate between both electrostatic and geometric differences is important in this analysis. For example, the geometric S_g would orient two identical copies of a molecule perfectly, and S_e would score them with a value of unity. If one of the molecules were simply expanded to become slightly larger without changing the potentials of the points, S_e would quantitate this decrease in similarity. This scheme is adequate for the molecules within the AMP deaminase, adenosine deaminase, or the AMP nucleosidase series, which are geometrically rather similar. If the molecules were sterically more dissimilar, option 3 from above could also be applied.

Results

Surface Constructions and Rotations of Test Molecules. Figure 4 depicts the electrostatic potential surfaces of the transition states and representative inhibitors for the AMP deaminase, adenosine deaminase, and AMP nucleosidase reactions. Omitted surfaces are of similar quality to those shown. The *S*-stereoisomer is created directly from the surface points of (*R*)-coformycin via the transformation $x_i \rightarrow -x_i$, $y_i \rightarrow y_i$, $z_i \rightarrow z_i$, and $\epsilon_i \rightarrow \epsilon_i$ for all points *i*. The method successfully reproduces, with discrete points, the surfaces in Figures 1−3.[48,49] The transition state structure for AMP deaminase, adenosine deaminase, or AMP nucleosidase is positioned and frozen with its center of mass at the origin of the coordinate system. The N-9-methyl transition state for AMP deaminase serves as the target structure to which the methyl derivatives of AMP, (*R*)- and (*S*)-coformycin 5′-phosphate, and the late transition state are compared. Adenosine deaminase features the same N-9-methyl transition state as does AMP deaminase, and to it are compared the methyl derivatives of adenosine, (*R*)-coformycin, 1,6-dihydropurine ribonucleoside, and hydrated purine ribonucleoside. Formycin 5′-phosphate, aminopyrazolopyrimidine ribonucleotide, tubercidin 5′-phosphate, and AMP are similarly scored against the transition state for AMP nucleosidase. As described above, within an enzyme series all molecules are geometrically similar: for the AMP deaminase structures, the substrate, the transition state, and the late transition state for the deaminase reaction are structurally related to each other through addition or deletion of a small number of atoms. Therefore, relative geometrical orientation between the target and test species results in a close matching for the bodies of the molecules when their respective centers of mass are coincident. This close matching of molecular volumes is an advantageous feature of these structurally related entities and, to emphasize, it is this initial purely geometric orientation using S_g upon which similarity ranking with S_e is built. The algorithm employed rotates the test molecule around its center of mass with calculation of S_g at each new orientation. The goal is to find the maximum value for the measure and thereby the best relative alignment. A single rotation requires three euler angles: each is generated randomly, with 100 000 random reorientations providing adequate sampling of rotation space. Because the similarity measure is assessed at each new orientation and so involves multiple evaluations of the exponential in the point-to-point distance, computation time can become substantial. To facilitate computation, the quantity $\exp(-\alpha r_{ij}^2)$ is calculated only if $\alpha r_{ij}^2 \leq 4.5$ for the points *i* on A (the transition state) and *j* on B (the test molecule). This means that all point−point interactions for which $\exp(-\alpha r_{ij}^2) \leq 0.011$ are ignored, an approximation that introduces negligible error in the final relative orientation of the two molecules. After this alignment S_g or S_e can be calculated without the truncation in the double sum, but no significant change is found over the truncated numbers.

Choices of α for similarity calculations are adjusted so that a point on the target will see a neighborhood of points on the test molecule about a bond length (or less) away. The Gaussian outputs are in units of the Bohr radius, equal to 0.529 Å, so that α has units of bohr^{-2}. Values of $\alpha = 0.1$, 0.3, and 0.5 bohr^{-2} provide exponential decays to e^{-1} in a point-to-point distance of 1.67, 0.97, and 0.75 Å, respectively. All of these α values give similar ranking of molecular similarity, demonstrating the robust nature of the algorithm.

AMP Deaminase Series. It is found that the orientation of a given test molecule, e.g., the methyl derivative of AMP, (*R*)- or (*S*)-coformycin, or the late transition state, relative to the transition state is essentially the same for any α (0.1, 0.3, or 0.5). As a representative example, alignment of (*R*)-coformycin with the transition state is shown in Figure 5. All other test molecules show comparable alignments relative to the transition

(48) Indeed, the surface topologies are more easily appreciated when physically rotating the molecules with the visualization package.

(49) In the case of formycin 5′-phosphate, the extraction procedure retrieved from the *Gaussian* files a single point that was not on the surface but which had, nonetheless, an electron density ≅ *0.002 electrons/bohr*^{3}. Visualization uncovered it, and it was deleted. This point corresponded to a region of low electron density in the molecular core close to a nucleus.

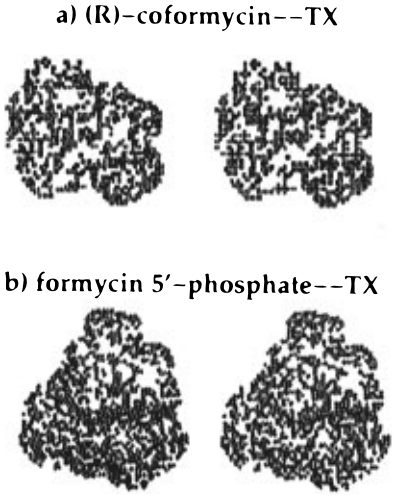


Figure 5. Stereoviews of the orientation of 9-methyl-(*R*)-coformycin with respect to the transition state for AMP (or adenosine) deaminase (a) and of formycin 5′-phosphate with respect to the AMP nucleosidase transition state (b). In both cases the transition states (called TX) are in red. The orientation is achieved through the purely geometric construction S_g. The color scheme for the electrostatic coding is suppressed. Orientations for all other molecules are comparable.

Table 1. S_e and S_g Values for the Similarity to the Relevant Transition State[a]

	α = 0.1		α = 0.3		α = 0.5	
enzyme/molecule	S_e	S_g	S_e	S_g	S_e	S_g
AMP Deaminase						
transition state	1.000	1.000	1.000	1.000	1.000	1.000
(*R*)-coformycin 5′-phosphate	0.692	0.960	0.598	0.850	0.523	0.759
(*S*)-coformycin 5′-phosphate	0.655	0.963	0.535	0.854	0.457	0.759
AMP	0.522	0.945	0.410	0.821	0.352	0.725
late transition state	0.019	0.970	0.080	0.888	0.091	0.811
Adenosine Deaminase						
transition state	1.000	1.000	1.000	1.000	1.000	1.000
hydrated purine ribonucleoside	0.809	0.978	0.765	0.923	0.721	0.865
(*R*)-coformycin	0.692	0.960	0.598	0.850	0.523	0.759
1,6-dihydropurine ribonucleoside	0.754	0.970	0.675	0.889	0.603	0.807
adenosine	0.522	0.945	0.410	0.821	0.352	0.725
AMP Nucleosidase						
transition state	1.000	1.000	1.000	1.000	1.000	1.000
formycin 5′-phosphate	0.527	0.930	0.418	0.772	0.357	0.652
aminopyrazolopyrimidine ribonucleotide	0.354	0.932	0.305	0.778	0.270	0.663
tubercidin 5′-phosphate	0.327	0.935	0.299	0.790	0.268	0.681
AMP	0.178	0.939	0.178	0.792	0.163	0.677

[a] Each label indicates the molecular species that is scored against the transition state, with the transition state−transition state comparison serving as the control. The test molecules are aligned relative to the transition state with S_g. Results are shown for three values of α.

state. In this example each molecule has a single color to permit easy comparison of alignment. When the points are also color-coded for their electrostatic potentials, the preferential overlap of electrostatic similarities for the two superimposed molecules is apparent.

The values of S_e and S_g for these sterically matched orientations are reported in Table 1. As a control, the transition state is randomly reoriented against itself to find the perfect match with both S_g and S_e having the maximum value of unity. For the three values of α studied, (the methyl derivative of) (*R*)-coformycin is more similar to the transition state—as expressed through S_e—than is either AMP or (*S*)-coformycin. This correlates qualitatively with the great affinity of the enzyme for the transition state inhibitor. Indeed, the average value of S_e (taken over the three α's) for (*R*)-coformycin is 0.604, while that for AMP is 0.428 with 0.549 for (*S*)-coformycin. The geometries of AMP and (*R*)- and (*S*)-coformycin as compared with the transition state are quite similar by the measure S_g, with an average of 0.856 for (*R*)-coformycin, 0.830 for AMP, and 0.859 for the *S*-inhibitor. Therefore, *the measure S_e is more closely related to binding affinities*, given that the molecules are subject to an initial geometrical alignment. Although the late transition state is sterically most similar to the transition state (see the visualizations or values for S_g), S_e affirms that it is a poor match to the transition state. The late transition state is part of the reaction pathway: a point in the pathway has been reached where the reaction-coordinate species is losing electrostatic similarity to the transition state, with, presumably, the relaxation of binding energy. As the value of α increases, a point on one molecule sees less of a neighborhood on the other, and in this sense the similarity measures explore more local features. Therefore, S_e for the late transition state increases with increasing α because information about the global electrostatic dissimilarity is lost. S_e for AMP and the two isomers of coformycin decreases with increasing α, reflecting that all share similar electrostatic neighborhoods with the transition state.

Table 2. Dimensionless Free Energy of Binding, $\Delta G/RT$, and the Average Similarity Measure, $\langle S_e \rangle$[a]

enzyme/molecule	$\Delta G/RT$	$\langle S_e \rangle$
AMP Deaminase		
transition state (for dAMP)	−36	1.000
(*R*)-deoxycoformycin 5′-phosphate	−17	0.604
(*S*)-deoxycoformycin 5′-phosphate	−10	0.549
dAMP	−6.7	0.428
transition state (for AMP)	−38	1.000
(*R*)-coformycin 5′-phosphate	−25	0.604
AMP	−8.1	0.428
Adenosine Deaminase		
transition state	−39	1.000
hydrated purine ribonucleoside	−29	0.765
(*R*)-coformycin	−25	0.604
1,6-dihydropurine ribonucleoside	−12	0.677
adenosine	−10	0.428
AMP Nucleosidase		
transition state	−39	1.000
formycin 5′-phosphate	−17	0.434
aminopyrazolo pyrimidine ribonucleotide	−12	0.310
tubercidin 5′-phosphate	−9.9	0.298
AMP	−9.0	0.173

[a] References for the free energy values are given in the text.

Discussion of AMP Deaminase Results. A quantitative correlation can be made between the S_e values for the methyl derivatives of (*R*)-coformycin, (*S*)-coformycin, and AMP and their respective binding free energies to AMP deaminase. The binding free energies to AMP deaminase and the average values for the similarity scorings of these molecules with the transition state are summarized in Table 2. The free energies listed are for both the 2′-deoxy molecules (dAMP, (*R*)-deoxycoformycin 5′-phosphate, (*S*)-deoxycoformycin 5′-phosphate, and the transition state for dAMP) and for the series AMP, (*R*)-coformycin 5′-phosphate, and the transition state for AMP. Figure 6 shows a plot of the dimensionless free energy of binding, $\Delta G/RT$, as a function of the average similarity measure, $\langle S_e \rangle$. Each transition state is scored against itself, giving a similarity of 1.00. Note the near-linear correlation for the 2′-deoxy series (black circles). The difference between the similarity measures for (*R*)-deoxycoformycin scored against the transition state and for the transition state scored against itself is 0.396, and the difference in the corresponding binding free energies is 19, giving a ratio of 48. For dAMP and the transition state the ratio is 51. (*S*)-Deoxycoformycin gives a ratio of 58. The *S*-inhibitor is held only 30 times more tightly by the enzyme than is the substrate dAMP. This means that the *S*-isomer and the substrate should exhibit approximately the same degree of

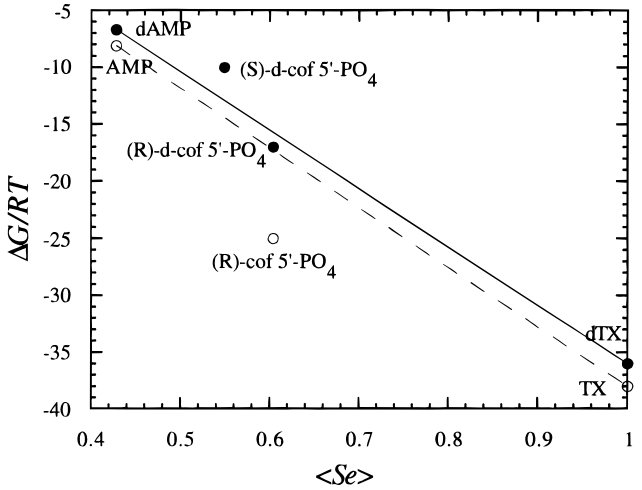


Figure 6. Binding free energy *vs* the average value of the similarity measure. The AMP deaminase series: black circles label the transition state of dAMP (called dTX), (*R*)- and (*S*)-deoxycoformycin 5′-phosphate, and dAMP. The solid line connects the substrate to dTX. Open circles are for the transition state of AMP (called TX), (*R*)-coformycin 5′-phosphate, and AMP. A broken line connects TX to AMP. The deoxy series shows better correlation between binding free energy and the similarity measure.

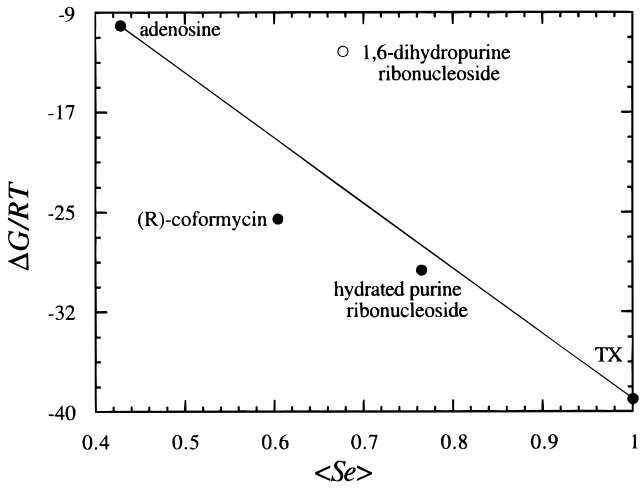


Figure 7. Binding free energy *vs* the average value of the similarity measure. The adenosine deaminase series: transition state (called TX), 1,6-hydrated purine ribonucleoside, (*R*)-coformycin, and adenosine. A line connects TX to substrate. 1,6-Dihydropurine ribonucleoside, an outlier, is shown with an open circle (see the text).

global similarity to the transition state. In the absence of any knowledge concerning the possible conformation of its hydroxyl group in the enzymatic cleft, this *S*-enantiomer is constructed as the mirror image of the *R*-form. It is therefore expected that the S_e values for the *S*-species and the substrate—which reflect the similarity of these test molecules to the transition state—would be more nearly the same if such information were available. The strong correlation between $\Delta G/RT$ and $\langle S_e \rangle$ is found as well for adenosine deaminase and AMP nucleosidase. In the series AMP, its transition state, and (*R*)-coformycin 5′-phosphate (open circles), the correlation is not as strong since the transition state inhibitor falls significantly below the line connecting substrate to transition state. Indeed, for this case the inhibitor is bound 3×10^7 times as tightly as is the substrate, while in the deoxy series, which shows stronger correlation, (*R*)-deoxycoformycin 5′-phosphate is bound only 3×10^4 times tighter than dAMP. This underscores the enzyme's remarkable ability to differentially stabilize almost identical transition states, while binding with nearly the same affinity the related substrates AMP and dAMP.

It is also possible to orient a test molecule relative to the transition state using the electrostatic measure S_e. The orientation thus achieved by maximizing S_e is the same as that from application of S_g only for test molecules that are similar both electrostatically and geometrically to the target. Thus, (*R*)-coformycin can be aligned to the transition state with S_e or S_g: the same relative orientation is found with either and for all values of α. For the electrostatically dissimilar late transition state, however, orientation through S_e does not lead to near atomic overlap with the target for any value of α. The substrate AMP serves as an intermediate case for which good atomic superpositioning with the transition state is possible for $\alpha = 0.5$ only. Because AMP, the transition state, the late transition state, and the inhibitor for the AMP deaminase reaction share global structural similarities, it is most comprehensive to demand above all good steric overlaps prior to similarity scoring with S_e.

Adenosine Deaminase Series. This enzyme shares with AMP deaminase, as input to the similarity calculations, the methyl derivatives of (*R*)-coformycin, AMP (or adenosine), and a common transition state. Table 1 lists S_e for these molecules and for 1,6-dihydropurine ribonucleoside and hydrated purine ribonucleoside as compared to the transition state. The relative orientations of the inhibitors and the substrate with respect to the transition state are found through S_g, and a representative example is shown in Figure 5. Table 2 summarizes the average values of the S_e measure (taken over the parameter α) and the free energies of binding of the five molecules. Focusing on the transition state, the hydrated purine ribonucleoside, (*R*)-coformycin, and AMP, it is found that the similarity measures for these molecules decrease in that order (1.000, 0.765, 0.604, 0.428) with the binding free energies $\Delta G/RT$ to adenosine deaminase becoming less favorable (−39, −29, −25, −10). Figure 7 shows the near-linear correlation between these quantities. Note that though (*R*)-coformycin, the transition state, and the substrate appear in the AMP deaminase series as well, there they have different binding affinities, and adenosine deaminase therefore constitutes a separate test case. 1,6-Dihydropurine ribonucleoside is an outlier: its similarity measure is larger than that of coformycin even though the dihydro inhibitor is considerably less tightly bound. The explanation for this comes from the crystal structure of cytidine deaminase complexed with 3,4-dihydrozebularine.[50] This inhibitor is the pyrimidine ribonucleoside analog of 1,6-dihydropurine ribonucleoside and has an energetically costly interaction with the zinc-bound water of cytidine deaminase. Adenosine deaminase has a zinc-bound water at the analogous position, which would be expected to interact sterically with one of the hydrogens on C-6 of the inhibitor. Therefore, though 1,6-dihydropurine ribonucleoside is fairly similar to the transition state as measured by S_e, it is less tightly bound to adenosine deaminase since it cannot be hydrated by the enzyme to form transition state contacts.

AMP Nucleosidase Series. Another test of the algorithm for scoring similarities is to differentiate between the binding affinities of AMP nucleosidase for formycin 5′-phosphate, aminopyrazolopyrimidine nucleotide, tubercidin 5′-phosphate, and for AMP. All molecules, as well as the transition state target, are shown in Figure 3 with a negative charge on their phosphate groups. This corresponds to the solution phase ionization of these test molecules. On average the similarity of formycin 5′-phosphate to the transition state as determined via S_e is 0.688; for AMP the similarity scores to 0.683. It is clear why S_e cannot convincingly differentiate between these

(50) Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, C. W., Jr. *Biochemistry* **1995**, *34*, 4516−4523.

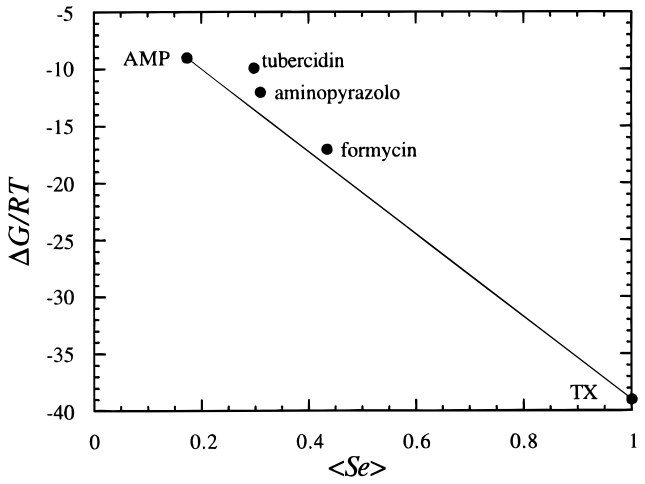


Figure 8. Binding free energy *vs* the average value of the similarity measure. The AMP nucleosidase series: transition state (called TX), formycin 5′-phosphate, 4-aminopyrazolo[3,4-*d*]pyrimidine-1-ribonucleotide, tubercidin 5′-phosphate, and AMP. Similarity measures are for molecules with protonated phosphates. A line links TX to substrate.

negatively charged molecules. For any given molecular surface, the charge on the phosphate serves as a source of electronegativity and gives a negative electrostatic potential to virtually all surface points. This effect overwhelms the localized and subtle differences in the molecules. Experimental evidence suggests that, in the catalytic cleft, the phosphate group of a bound species interacts with a positive residue from the enzyme.[42] Protonation of the phosphate groups of the three inhibitors, AMP, and the transition state and recalculation of the similarity measures leads to the results shown in Table 1. Figure 5 gives a representative example of the orientation of a test molecule relative to the transition state, as determined through the maximum in S_g. The relative orientations of the other molecules to the transition state are similar. Table 2 lists the average S_e for each molecule along with its binding constant to AMP nucleosidase, and Figure 8 shows the correlation between these quantities. Through the entire series, a larger value for the similarity between a test molecule and the transition state corresponds to a greater affinity for AMP nucleosidase. For the transition state, formycin, aminopyrazolopyrimidine ribonucleotide, tubercidin, and AMP, $\langle S_e \rangle$ is 1.000, 0.434, 0.310, 0.298, and 0.173, respectively, while $\Delta G/RT$ for these molecules is −39, −17, −12, −9.9, and −9.0. S_e distinguishes between aminopyrazolopyrimidine nucleotide ($K_i = 10 \times 10^{-6}$ M), tubercidin 5′-phosphate ($K_i = 51 \times 10^{-6}$ M), and AMP ($K_M = 120 \times 10^{-6}$ M), even though tubercidin is bound only twice as tightly as is the substrate and only 5 times less tightly than is the aminopyrazolo compound. To increase further the quantitative correlation, the value of S_e for tubercidin (0.298) needs to be closer to that for AMP (0.173) and more distant from the measure for the aminopyrazolo compound (0.310). A similarity measure focused locally on the areas of the inhibitors and substrate corresponding to the proton on N-7 of the transition state can be expected to provide this. Though both tubercidin and AMP have a proton on C-8, as does the transition state, aminopyrazolopyrimidine nucleotide does not. If the similarity to the transition state at the C-8 or N-8 regions were ignored (assuming that this loci is unimportant for recognition by the enzyme), this last compound would be more like the transition state, effectively achieving the goal.

Given a putative transition state inhibitor for any of these enzymes and calculation of its value for S_e, it is possible to estimate its free energy of binding—therefore, its effectiveness as an enzymatic inhibitor—from the constructions shown in Figures 6−8.

Concluding Remarks

The wealth of physical and chemical information contained within molecular electrostatic potential surfaces can be extracted and quantified for similarity comparisons between enzymatic substrates, inhibitors, and transition states. For the three reactions studied, those of AMP deaminase, adenosine deaminase, and AMP nucleosidase, the transition state inhibitors are scored more similar to the transition states than are the substrates. Furthermore, there is general correlation between the similarity of an inhibitor to the transition state and the binding free energy to the enzyme. The similarity measures introduced are sensitive to geometric rearrangement of an inhibitor into its stereoisomer; such a procedure diminishes the potency of the transition state inhibitor (*R*)-coformycin 5′-phosphate.

Although refinements of the scheme introduced are possible, the method is intuitive and robust. First of all, the point density on any molecular surface is greatest on those regions having largest curvatures. One can imagine the creation of a mesh on the van der Waals surface with an extrapolation procedure to find the electrostatic potential and coordinates at uniformly distributed mesh points. As discussed above, however, the exponential term in the similarity measures serves to compensate for nonuniform coverage. A simulated annealing scheme to find the maximum value of the measures might be employed if a greater surface coverage is used. This approach has been found to be successful in speeding approach to an extremum in many-dimensional optimizations.[51] Finally, if experimental data can establish the specific interaction sites within the transition state−enzyme complex, then those points on a putative inhibitor thought to be unessential in the recognition and binding process can be systematically deleted.

Acknowledgment. C.K.B. thanks Joel Ehrlich and Paul Kline for help with the Gaussian 92 and AVS packages and for kindly sharing the original Gaussian outputs for the molecules they studied. John Mazzella, Paul Weakliem, and Paul Berti have been remarkably patient in explaining various graphics programs. Thanks also to Paul Berti for sharing his enzymological savvy. This work was supported by NIH grant GM 41916, by funds from the Mather Foundation, and by the U.S. Army Medical Research and Development Command under Contract No. DAMD 17-93-C-3051. The views, opinions and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other documentation.

JA952781N

(51) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. *Science* **1983**, *220*, 671−683.