Prediction of Inhibitor Binding Free Energies by Quantum Neural Networks. Nucleoside Analogues Binding to Trypanosomal Nucleoside Hydrolase†

Benjamin B. Braunheim,‡ Robert W. Miles,§ Vern L. Schramm,*,§ and Steven D. Schwartz*,‡,§

The Departments of Physiology and Biophysics and of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Received April 9, 1999; Revised Manuscript Received August 2, 1999

ABSTRACT: A computational method has been developed to predict inhibitor binding energy for untested inhibitor molecules. A neural network is trained from the electrostatic potential surfaces of known inhibitors and their binding energies. The algorithm is then able to predict, with high accuracy, the binding energy of unknown inhibitors. IU-nucleoside hydrolase from *Crithidia fasciculata* and the inhibitor molecules described previously [Miles, R. W. Tyler, P. C. Evans, G. Furneaux R. H., Parkin, D. W., and Schramm, V. L. (1999) *Biochemistry 38, xxxx-xxxx*] are used as the test system. Discrete points on the molecular electrostatic potential surface of inhibitor molecules are input to neural networks to identify the quantum mechanical features that contribute to binding. Feed-forward neural networks with back-propagation of error are trained to recognize the quantum mechanical electrostatic potential and geometry at the entire van der Waals surface of a group of training molecules and to predict the strength of interactions between the enzyme and novel inhibitors. The binding energies of unknown inhibitors were predicted, followed by experimental determination of K_i values. Predictions of K_i values using this theory are compared to other methods and are more robust in estimating inhibitory strength. The average deviation in estimating K_i values for 18 unknown inhibitor molecules, with 21 training molecules, is a factor of $5 \times K_i$ over a range of 660 000 in K_i values for all molecules. The *a posteriori* accuracy of the predictions suggests the method will be effective as a guide for experimental inhibitor design.

The development of computational and theoretical methods for prediction of the binding constants of proposed inhibitors or analogues for biological receptors would facilitate the discovery of novel bioactive molecules. A general method is tested here that relies on neural network recognition to predict binding affinity. There are other methods to predict binding affinity, notably docking calculations (*1*) and QSAR calculations (*2, 3*). The calculations reported here will show a higher accuracy than recent applications of these methods, with accuracy tested over a much wider range of binding affinities. The method is simple to implement, with no needed assumptions for numerical models or structure activity relationships.

A computational neural network is a computer algorithm that is trained to learn features of input patterns and associate these with an output. After a learning phase, a well-conditioned network is able to predict an output for patterns not in the training set. In the present work, the neural net is trained with substrate and inhibitors for the well-characterized IU-nucleoside hydrolase from *Crithidia fasciculata* (*4*−*6*). The network is trained to associate the three-dimensional molecular electrostatic potential surface properties of the molecules in the training set (the input) with a free energy of binding (the output). The network is then used to predict the free energy of binding for unknown molecules.

Neural networks have been considered previously for the task of simulating biological molecular recognition. Gasteiger et al. (*7*) have used Kohonen self-organizing networks, Wagener et al. (*8*) have used autocorrelation vectors to describe different molecules, and Weinstein et al. (*9*) used neural networks to predict the mode of action of different chemotherapeutic agents. However, none of the previous neural networks have been trained with full three-dimensional information or have been required to predict the properties of unknown molecules in a completely blind test. The present work is a departure from all previous approaches in that it incorporates the quantum mechanical electrostatic potential at the entire van der Waals surface of a molecule as the physicochemical descriptor. This description contains the complete information that defines enzyme-ligand interactions, because the relative electrostatic field at the van der Waals surface describes exactly the quantum mechanical properties of ionic, hydrogen bonding, and hydrophobic sites on the molecular geometry.

IU-NUCLEOSIDE HYDROLASE

Cleavage of the N-ribosidic bonds of nucleosides and nucleotides is an essential function in all organisms. Protozoan parasites lack *de novo* purine biosynthetic pathways and rely on the ability to salvage purines from the nucleosides found in the blood of their host for RNA and DNA synthesis (*10*). Nucleoside hydrolases are common in protozoan parasites and are not found in mammals. The inosine-uridine

† This work was supported by NIH training grant GM08572 and research grants GM41916 and AI34342.

* Address correspondence to these authors. Van L. Schramm: Telephone, 718-430-2813; Fax, 718-430-8565; E-mail, Vern@aecom.yu.edu. Steven D. Schwartz: Telephone, 718-430-2139; E-mail, sschwartz@aecom.yu.edu.

§ Department of Biochemistry.

‡ Department of Physiology and Biophysics.

[1] Abbreviations: IU, inosine/uridine; NH, nucleoside hydrolase; QSAR, quantitative structure activity relationship; pAPIR, para-aminophenyliminoritbitol; Å, angstrom.

10.1021/bi990830t CCC: $18.00
Published on Web 11/12/1999

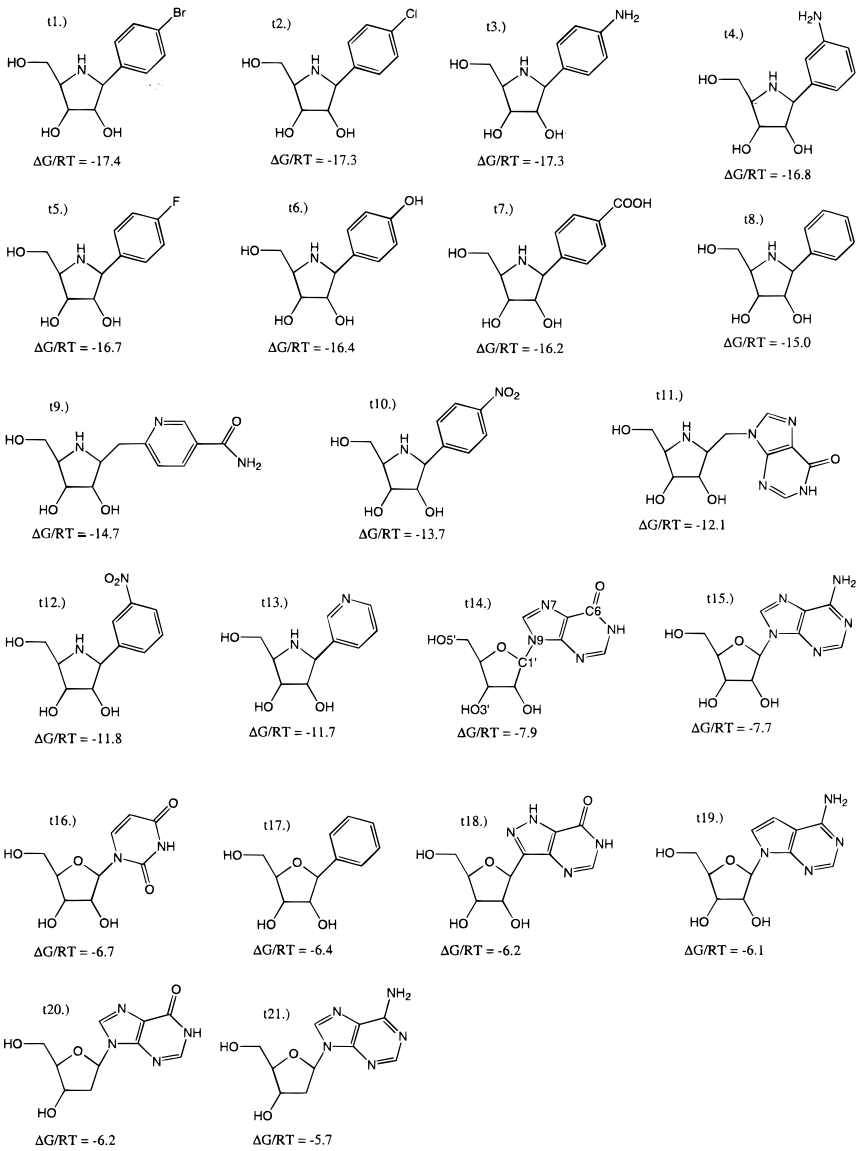

FIGURE 1: Molecules in the training set, binding energies in dimensionless units of $\Delta G/RT$. Most of these values are taken from reference *5*, *11*, and *15*. For all studies, $T = 303°$ K. The molecular electrostatic potential was calculated for the neutral structures shown. The three reference atoms chosen for inosine (t14) are: $\alpha =$ N9, $\beta = 3'$O $\gamma = 5'$O.

nucleoside hydrolase (IU-NH) found in *Crithidia fasciculata* catalyzes the N-ribosyl hydrolysis of all commonly occurring purines and pyrimidines (*6*, *11*). The active site of the enzyme has two binding regions, one region binds ribose and the other binds the leaving group. The IU-NH lowers the energy of activation required for hydrolysis of inosine by $\Delta\Delta G =$ 17.7 kcal/mol. Most of this energy (13.1 kcals/mol) is used in activation of the ribosyl group, and only 4.6 kcals/mol is used for activation of the hypoxanthine leaving group (*5*). The transition state structure is ribocarbenium-ion in character, and analogues that resemble the inosine transition-state both geometrically and electronically have proven to be powerful competitive inhibitors of this enzyme (*5*, *12*−*15*).

*Ribose Binding Region.*The transition state for IU-NH features an oxocarbenium-ion achieved by the polarization of the C4′ oxygen and C1′ carbon bond of ribose. The C4′ oxygen is proposed to be sandwiched between the lone pair from the C5′ oxygen and carboxamide of Asn168 during transition-state stabilization (*4*). This results in a partial double bond between the C4′ oxygen and the C1′ carbon causing the oxygen and the carbon to share a partial positive charge. Nucleoside analogues with iminoribitol groups have a secondary amine in place of the C4′ oxygen of ribose and have proven to be effective inhibitors of IU-NH.

Leaving-Group Binding Region. IU-NH acts on all naturally occurring ribonucleosides. The lack of specificity for the leaving groups results from the small number of amino acids in this region that might form specific interactions, e.g., Tyr229, His82, and His241 (*4*). The only crystal structure data available concerning the configuration of bound inhibitors was generated from a study of the enzyme bound to *p*-aminophenyliminoribitol (pAPIR) Figure 1, t3. Tyr229 relocates during binding and moves above the phenyl ring of pAPIR. The side chain hydroxyl group of Tyr229 is directed toward the cavity that would contain the six member ring of a purine, were it bound. His82 is 3.6 Å from the phenyl ring of pAPIR, and in the proper position for positive charge-π interactions to occur. None of these interactions are expected to provide strong leaving-group assistance. Mutagenesis of His241 has demonstrated a role in leaving-group activation in the hydrolysis of inosine, presumably as

the proton donor for transition-state formation, but it is not in contact with pAPIR (*4, 16*).

ELECTROSTATIC POTENTIAL SURFACES

The molecular electrostatic potentials calculated at the van der Waals surface of existing or proposed inhibitors are used as the descriptors of chemical structure and properties. Electrostatic properties define the electronic potential for interactions between molecules and the active site (*17*). Molecular regions with electrostatic potentials close to zero are capable of van der Waals interactions. Atoms with a partial positive or negative charge can serve as hydrogen bond donor or acceptor sites. Regions with greater positive or negative potentials may be involved in H-bonding and Coulombic interactions. The electrostatic potential also conveys information concerning the likelihood that a particular region can undergo electrophilic or nucleophilic attack (*18*). The quantum mechanical electrostatic potential at the van der Waals surface is calculated using 3-21G basis set (*Gaussian 96* was used to generate all quantum chemical data (*19*)). Since molecules described by quantum mechanics have finite electron density in all space, molecular geometry is defined at the van der Waals surface by selecting the space around the molecule where the electron density is close to $0.002 \pm \delta$ *electrons/bohr*^{3} where δ is the acceptance tolerance. These points describe a surface under which approximately 95% of the electron density resides. δ is adjusted so that 15 points per atom are accepted, creating a well-defined molecular surface (*20*). The information for each molecular surface is described by a matrix with dimensions of $4 \times n$ where n is the number of points for the molecule and the row vector of length 4 contains the *x,y,z*-coordinates of a given point and the electrostatic potential at that point.

PREPARATION OF THE INPUT PATTERNS

The surface features of molecular structures are oriented for maximum geometric coincidence prior to input into the neural network. This is accomplished by choosing three reference atoms from each molecule that are common to all molecules. The reference atoms are labeled α, β, and γ. The α atom is translated to the origin, followed by translations of β and γ, and the surface points. The basis set is then rotated to place β on the positive z axis. Finally, the basis set is rotated to place γ in the positive xz plane.

Input patterns for a neural network are presented in the form of a vector with entries (l_1, l_2, ..., l_n). Since the molecules are represented by a $4 \times n$ matrix, a method is needed to discard the *x,y,z*-coordinates but to maintain spatial information. This is accomplished by mapping the surface points of every molecule onto the reference surface. Mapping ensures that similar regions on different molecules enter the same part of the neural network. To minimize the loss of geometric information, we augment the network with this geometric information. Geometrical information was generated by a reference surface that was larger, in all regions, than the surfaces used in the study (a sphere of diameter slightly larger than the largest molecule). During the nearest neighbor mapping, the shortest distance between each point on the reference and each point on the inhibitor is calculated. The input to the neural network includes the electrostatic potential at many points, and the distance of that point from the nearest point on the reference surface. Using a reference surface larger than the other surfaces permits a uniform outward mapping. In the limit, with an infinite number of points, all mappings are normal to the inhibitor's surface, and the mapping distances will be as small as possible. To approach this limit a 10-fold excess of points was selected to describe the molecules. After the mappings, the electrostatic potential and geometry vectors are culled by removing nine out of 10 points. For a study where the average molecule has 27 atoms, the molecular surface would be described by 4000 points. After mapping to the sphere, the molecular inputs would be represented by 400 electrostatic potential scalars and 400 geometry scalars.

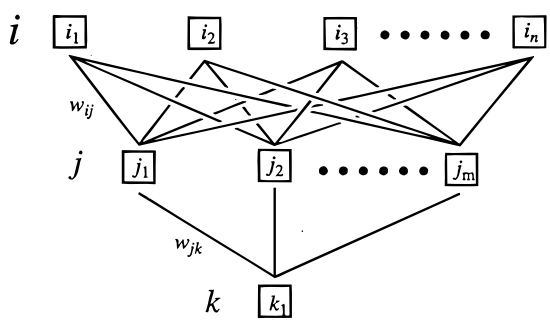

FIGURE 2: Schematic representation of a neural network. Row i represents the input layer. Row j is the hidden layer and row k is the output layer. The layers interact by means of a weight matrix, w_{ij} for the hidden layer, and w_{jk} for the output layer. The input to the network is descriptions of the molecules and the output is the binding free energy of the molecule with the enzyme, $\Delta G/RT$.

NEURAL NETWORK CONSTRUCTION

Computational neural networks are composed of many simple units operating in parallel. These units and their interactions are inspired by biological nervous systems. The network function is largely determined by the interactions between units. Networks learn by adjusting the values of the connections between elements (*21*). The neural network employed in this study is a feed forward with back-propagation of error network, which learns with momentum. This neural network has three layers: an input layer, a hidden layer, and an output layer. The input layer is where the different input vectors are transferred. The link between the layers of the network is one of multiplication by a weight matrix where every entry in the input vector is multiplied by a weight and sent to every hidden layer neuron, so that the hidden layer weight matrix has the dimensions $m \times n$, where n is the length of the input vector and m is the number of hidden layer neurons. A bias is added to scale all the arguments before they are input into the transfer function (*21*). The value of the weights and biases are under the control of the neural network learning algorithm. The actual values of the weights and biases are adjusted during training.

The schematic for the input layer is represented by the squares at the top of Figure 2. The weights are represented by the lines connecting the layers: w_{ij} is the weight between the i^{th} neuron of the input layer and j^{th} neuron of the hidden layer and w_{jk} is the weight between the j^{th} neuron of the hidden layer and the k^{th} neuron of the output layer. In this diagram, the output layer has only one neuron because the target pattern is a single number $-\Delta G/RT$, the dimensionless binding energy for the ligand (also equal to $\ln K_i$ in molar units).

Back-propagation was created by a generalization of the Widrow-Hoff learning rule applied to multiple-layer networks and nonlinear differentiable transfer functions (*22*). Input vectors and the corresponding output vectors are used to train until the network can approximate a function (*22*). The strength of a back-propagation neural network is its ability to form internal representations through the use of a hidden layer of neurons. For example, the "exclusive or" problem demonstrates the ability of neural networks, with hidden layers, to form internal representations and to solve complex problems. An example is provided by four input patterns [(0,1) (0,0) (1,0) (1,1)] with output targets [1, 0, 1, 0], respectively. A perceptron or other single layer system would be unable to simulate the function described by these four input/output pairs. The only way to solve this problem is to learn that the two types of inputs work together to affect the output. In this case, the least similar inputs cause the same output, and the more similar inputs have different outputs (22). The ability required to solve this problem is not unlike that required to find the best inhibitor when it is not closely related to transition state of the targeted enzyme. It is this inherent ability of neural networks to solve complex puzzles that make them well conditioned for the task of simulating biological molecular recognition.

METHODS

Choosing Molecular Configurations. Inhibitors of IU-NH were selected with a variety of base and ribose analogues. The only crystal structure data available to provide the configuration of bound inhibitors is that obtained in a study of the enzyme bound to *p*-aminophenyliminoribitol (*4*). The crystal structure provides reliable information for the relative position of the phenyl group to the iminoribitol, and the structure of bound iminoribitol. The phenyl group has 2-fold rotational symmetry (identical on rotation of 180°), thus the structure offers ambiguous information as to how molecules that are unambiguous on rotation should be oriented. Structure t14 is inosine, one of the substrates. Catalysis is proposed to involve interaction between N7 of the hypoxanthine group and His241, the proposed general acid for proton donation (*4*). However, there is no absolute way to establish which geometry is correct from the existing data. To determine if geometric orientation discrepancies introduce misinformation, we used both orientations in the set of training and test molecules. The neural network was trained with the patterns of each molecule in both orientations.

In addition to using both configurations about the ribosidic bond, there are additional degrees of rotational freedom that were made consistent for all molecules. The assumption is made that the enzyme will bind all molecules in a similar conformation to minimize the energy of the complex. The neural network is relatively insensitive to configuration as long as a consistent pattern is presented to the neural network. The known conformation of *p*-aminophenyliminoribitol in the crystal structure with IU-NH is used as the model conformation for the other molecules.

The transition-state structure for inosine hydrolysis is achieved through the creation of an unstable ribooxocarbenium ion (*6*). The groups surrounding the ribooxocarbenium transition state serve to neutralize the transition-state structure. To simulate this aspect of the active site, a negatively charged fluoride ion was placed at the same relative position of the nearest oxygen of the carboxyl group of the active site to polarize the inhibitors similar to the catalytic site environment. Calculations of the electrostatic potential at the van der Waals surface used these slightly polarized molecules. The fluoride ion polarizes, but does not impart significant electron density to the polarized molecule. The partial charge of the fluoride at this location, is greater than −0.99 as calculated by the Mulliken method, confirming that its electron density remains localized to the fluoride ion.

Simulation Constraints. A rigorous test of the neural network recognition method in drug discovery involved analysis of the structures of 21 inhibitor molecules whose K_i values were known. These 21 molecules provided the training set. The binding energy of another 18 molecules were unknowns to be tested by algorithm. The K_i values had not been determined, or were not disclosed to the investigators until the K_i values had been predicted by the neural net algorithm. The 21 molecules in the training set included eight substrates and substrate analogues of IU-NH (*23*), eleven inhibitors from previous work (*5*), and two inhibitors with previously unpublished binding energies. The neural network was trained with all 21 molecules as inputs. The output (training target) is the binding free energy of each molecule with the enzyme.

Training a neural network requires the adjustment of four parameters; (1) the number of hidden layer neurons, (2) the learning rate, (3) the momentum constant, and (4) the number of training iterations. There are no set values for these parameters. A robust training procedure is developed by minimizing the errors using a known training set. Prediction error is the difference between the target$_i$ value for a molecule (experimentally determined binding energy), and the number the neural network predicted for that molecule$_i$, and summing the absolute value of this number for all the molecules in the training set. As training progresses the four parameters are refined and training set prediction error decreases. Minimizing training set error may have negative consequences. Overtraining occurs when the network trains for too many iterations, has too many hidden layer neurons, and has too large of a learning rate or too small of a momentum constant. Detecting an overtrained neural network is accomplished by making predictions for inhibitors not in the training set. A well-conditioned network can generalize from the information contained in the input/output pairs of the training set and apply that information to a novel molecule. Optimal training is accomplished by using one of the 21 training set molecules as a calibration molecule. This molecule is deleted during training and used to determine if the neural network was overtrained. The neural network is trained until the prediction set error decreases to a constant value. Training is then completed and the neural network is tested with the calibration molecule. If the neural network predicts the binding energy of the calibration molecule to within 5% of the value for $\Delta G/RT$, then the neural network is saved. If the prediction is in error by more than 5% in $\Delta G/RT$, then different values for the four parameters are chosen. This procedure is repeated until values for the four parameters generate a network that can predict binding of the calibration molecule to the selected limits. This procedure is repeated systematically, using each of the 21 molecules in the training set as the calibration molecule and the

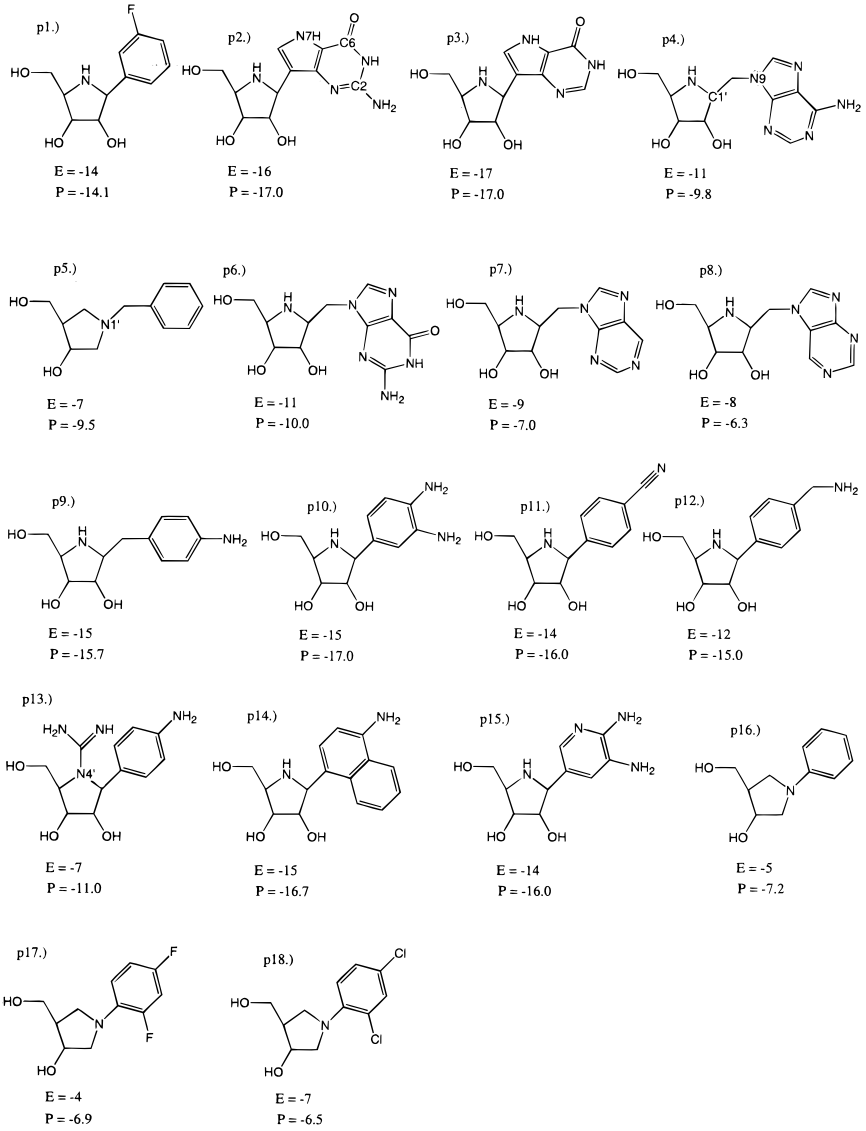

FIGURE 3: Molecules of unknown $\Delta G/RT$ used to determine the predictive power of the trained neural network. After the predictions (P) were made, the experimental $\Delta G/RT$ (E) were revealed. The inhibitor constants for p2-p4, p6-p8, p14, and p15 are reported in ref *15*. Inhibition constants for the remaining molecules were determined as described in ref *15* and are reported here for the first time.

remaining 20 as the training molecules. The 21 best neural network constructions are used to predict binding energies for each of the 18 molecules that are unknown to the neural network. The final reported result is the average of the predictions for each of the network constructions.

RESULTS

The structures of the 21 molecules used as the training set are shown in Figure 1 together with their experimental binding energies for IU-NH. Figure 3 shows the 18 molecules in the unknown molecule set, with their experimentally determined binding energies (*E*) and their predicted binding energies (*P*). Figure 4 shows a graphical representation of the data which relates the neural network prediction for $\Delta G/RT$ as a function of the experimentally determined binding energy. The neural network prediction method is effective over the entire range of binding energies from -4 to -17 $\Delta G/RT$ (18 mM$-$4 nM). The average deviation is 1.7 $\Delta G/RT$, a factor of 5, over an inhibitory range of 442 000 for the unknowns, and the largest deviation is 4 $\Delta G/RT$, a factor of 55 over the range of 442 000 in K_i values.

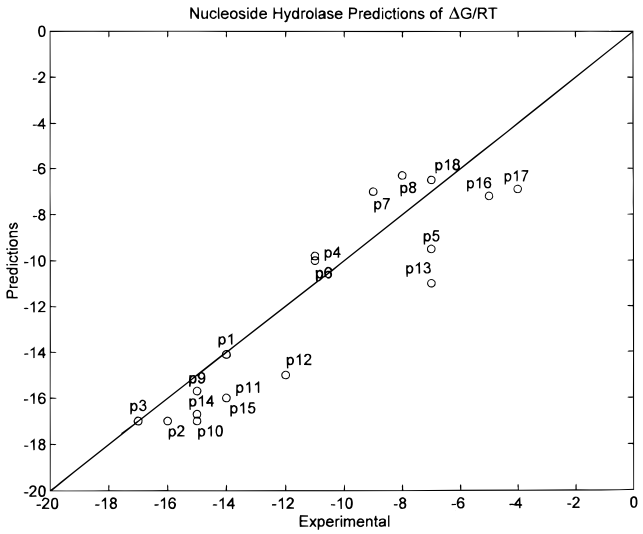

FIGURE 4: Correlation of $\Delta G/RT$ values predicted by the neural network with those determined from experimental data. The numbers correspond to the compounds of Figure 3.

The accuracy of the results demonstrates that molecular electrostatic potential surfaces can be used to train a mathematical algorithm that describes the relationship between structure and inhibition. There are specific successes that are the product of synthetic reasoning on the part of the neural network, which warrant discussion. The prediction of $\Delta G/RT$ by the neural network for unknown set molecule number one (p1 of Figure 3) is an example of synthetic reasoning by the neural network. Training set molecules numbers 10 and twelve (t10 and t12 of Figure 1) have nitro groups at the *para* and *meta* positions, respectively. Nitro groups are strongly electron withdrawing. The neural network recognized the negative effects of electron withdrawing substitutions to the phenyl substituent of pyrimidine analogues from the nitro substituted inhibitors in the training set. The algorithm also recognized the negative effects of *meta* verses *para* substitutions from studying the nitro inhibitors. Since *p*-fluorophenyliminoribitol (t5) was in the training set, the network was able to quantitate this information and subtract about 2 $\Delta G/RT$ from the binding energy of t5 when generating a prediction of m-flourophenyliminoribitol (p1). An interesting point about the prediction of p2 and p3 is that these molecules are tight binding 9-deaza purine analogues and are identified as such by the neural network. There are, however, no tight binding 9-deaza purine analogues in the training set. The neural network was able to combine rules from the binding of the weaker binding purines with the rules of the tight binding pyrimidines to generate predictions for tight binding purine analogous. The prediction of p3 was also generated by analogous reasoning. Inhibitor p3 has three features that are conducive to tight binding; an electron withdrawing group at C6, a partially positive charged hydrogen at N7, and an iminoribitol group. The neural network combined these pieces of information and predicted a large binding energy. The prediction of binding energy for p4 also required a combination of information from several training molecules. Compound p4 has two qualities that make it a poor inhibitor of IU-NH: (1) p4 has a methylene-bridged C1′-N9 glycosidic bond, the neural network could have learned about the negative effects of this feature from t9 and t11, and (2) p4 has an electron donating group at C6, these groups decrease the binding of purine analogues. Although p4 does have an iminoribitol group that predicts tight binding, the molecule is a poor inhibitor because of the electrostatic potential at N7 and the geometry that extends the purine ring to an unfavorable position relative to tight-binding inhibitors. The compound is correctly identified as such by the neural network.

DISCUSSION

The neural network forms its ability to predict $\Delta G/RT$ for unknown inhibitor molecules through the generation of rules. There are eight main rules that the neural network applies to be able to generate accurate prediction. The neural network must recognize the effects: (1) that the different purine analogue moieties have on enzyme binding [p2-p4, p6-p8, p14] and not apply these rules to pyrimidine analogues; (2) of substitutions to pyrimidine analogues, electron donating or withdrawing groups or both [p1, p5, p9-p13, p15-p18] and that these substitutions could occur at the *ortho*, *meta*, or *para* positions and not apply these rules to purines; (3) of having a secondary amine to replace the C4′ oxygen of ribose [p1-p4, p6-p12, p14, p15]; 4) of substitutions that lengthen the C1′-N9 glycosidic bond [p4-p9]; 5) of C2′ deoxyribosyl groups [p5, p16-p18]; (6) of a planar configuration at the C4′ nitrogen of the iminoribitol group [p13]; (7) of bulky substitutions to the C4′ nitrogen of the iminoribitol group [p13]; (8) of a methyl group in place of the C4′ oxygen and nitrogen in place of the C1′ carbon of ribose [p5, p9, p16-p18].

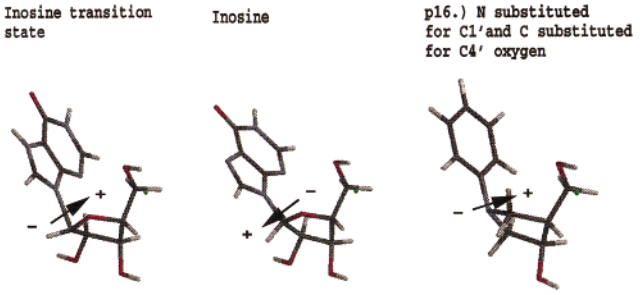

FIGURE 5: Even though the dipole of inhibitor p16 is in the same direction as that of transition state for inosine hydrolysis and opposite that of inosine substrate, p16 is a poor inhibitor. This feature is readily recognized by the neural network, even though no molecules in the training set contained nitrogen as replacement for C1′.

There are six main rules that the neural network could establish from the input/output pairs in the training set. The neural network could incorporate the effects: (1) that different purine and purine analogue moieties have on enzyme binding [t11, t14-t15, t18-t21]; (2) of substitutions to pyrimidine analogues, electron donating or withdrawing groups or both, and that these substitutions can occur at the *ortho*, *meta*, or *para* positions [t1-t10, t12, t13, t16, t17]; (3) of having variations at the C4′ oxygen moiety of ribose, such as a secondary amine, or the usual ether oxygen [t1-t21]; (4) of substitutions that lengthen the C1′-N9 glycosidic bond [t9, t11]; (5) of C2′ deoxyribosyl groups [t20, t21]; (6) of a planar configuration at the C1′ [t9, t11].

There are two examples where the neural network was able to establish rules that were not apparent from the training set: (1) The inhibitor p13 has a bulky guanidino group that causes the C4′ nitrogen to become planar, but sterically inaccessible. Planar character in this region suggests tight binding because the transition-state geometry is similar. However, the bulky group will certainly cause steric hindrance, since close interactions around O4′ are required for transition-state stabilization. The binding energy is, therefore, lower than otherwise expected. The most similar molecule in the training set is t3, which has a binding energy of -17.3, compared to the p13 binding energy of -7. The neural network prediction of -11.0 is a deviation of 4 ($\Delta G/RT$). This prediction has the greatest error but is an excellent guess considering the powerful binding of t3 and the weak binding of p13; (2) Inhibitors p5, p16 through p18 have nitrogen in place of the C1′ carbon of ribose, and carbon in place of the C4′ oxygen. This atomic arrangement is not like any in the training set. The weakest binding species in the training set are the substrates, with ground state carbon oxygen bonds at those positions that have partial charges in the transition state. On the basis of this information, one could argue that inhibitors p5 and p16 through p18 would bind tightly because the C4′ carbon has a small partial positive charge and the N1′ nitrogen has a partial negative charge (Figure 5). Both

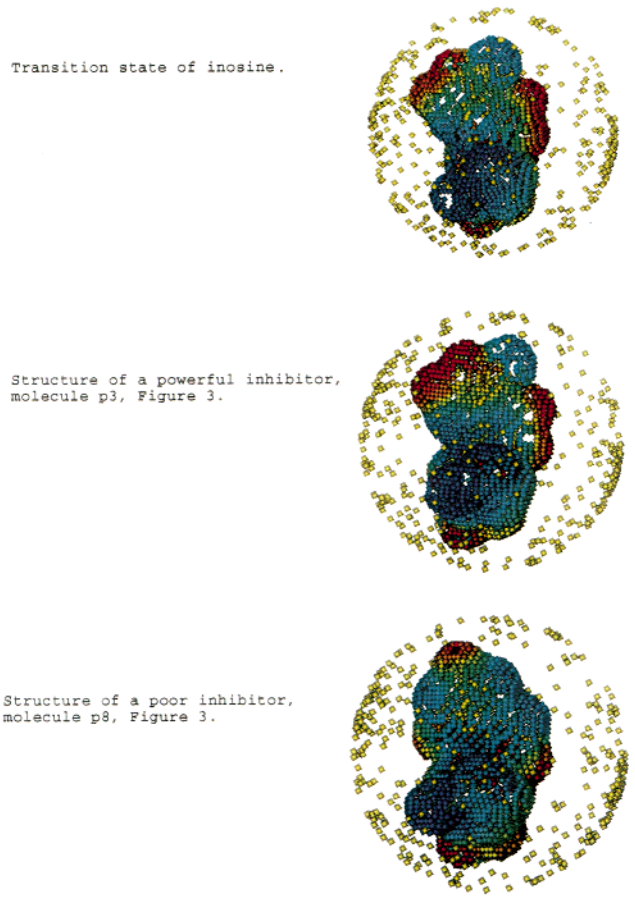

FIGURE 6: A representation of the molecular electrostatic potential and geometric input for inhibitor recognition by the neural network. The yellow points surrounding the molecules represent the surface to which the molecular electrostatic potentials are imprinted. The molecular electrostatic potential is shown at the van der Waals surface of the inhibitor molecules. Regions on the molecular surface that are colored red have a partial positive charge and regions colored blue have a partial negative charge. The structure on the top is the transition state for inosine hydrolysis, it has a high affinity for the enzyme. The structure in the middle is compound p3, a powerful inhibitor. The structure on the bottom is compound p8, a poor inhibitor.

features are suggestive of the electrostatic potential surface of the transition state. Inhibitor p5 and p16 through p18 have an average binding energy of −5.8 and the neural network predictions have an average of −7.5. The neural network (Figure 6) correctly predicts that the magnitude of the partial charge is more important than dipole direction. These compounds also lack a 2-hydroxyl group. Inhibition studies have shown that this group contributes at least 1000-fold to binding (*15*).

CONCLUSIONS

The quantum neural network technology is capable of accurate predictions of enzyme/inhibitor binding constants in a complex biochemical system. The predictive strength of the network is facilitated by including the hypothetic binding energy of the substrate bound in the Michaelis complex. The transition-state structure for inosine is created by polarization of the ribosyl group across the C4′ oxygen C1′ bond, and protonation of N7 of the purine group. Protonation of a stable pyrimidine ring (e.g., uridine base in t16) does not occur since the electrons of the ring nitrogen are involved in ring conjugation. Therefore, it is clear that pyrimidine analogues can develop only limited features of the inosine transition-state structure. Substituted phenyl groups at the 1′-position of iminoribitol decrease in binding energy with increasingly electron withdrawing substitutions. The opposite trend is seen with purine analogues. The neural network is able to predict binding energy for both purine and pyrimidine analogues and for substituted phenyl groups. Pattern classification and creation of priority trees is a known strength of neural networks and these features function well for the inhibitor substituents. The neural network identifies two or more groups that have one mutually exclusive characteristic that have an affect on the output. This characteristic relates specifically to geometric or electrostatic features and their affect on $\Delta G/RT$. Once two distinct inhibitor categories have been identified, rules for evaluating the two can be learned, the rules of one group can be the same or different from the rules of the other group. Examples of this are; methyl substitutions that lengthen the glycosidic bond decrease binding energy for purines and pyrimidines, but extremely electronegative groups (like protons and double bonded oxygens) increase binding energy for purines ($\Delta G/RT$ for t14 > t15, t18 > t19, t20 > t21) but electron withdrawing groups decrease the binding energy of pyrimidines ($\Delta G/RT$ for t1 > t5 > t10). As shown earlier with the "exclusive or" problem, the ability of a neural network to solve a problem when it appears to contain contradictory information is not a new finding. The novelty of this quantum-chemistry neural net application is that this ability is being employed to solve the real world problem of inhibitor design. Intuitive prediction of K_i values by experienced enzymologists is considerably less accurate than the correlation demonstrated in Figure 4.

The strength of neural network approach lies in the independence from specific identification of structure activity relationships. The algorithm independently chooses the important features of the quantum mechanical properties of the putative inhibitor molecules and generates the mathematical relationship between the feature and the inhibitory activity. It is possible in individual cases that simple polynomial fitting to a discrete set of properties would match the accuracy of prediction generated by the neural network, but the approach described in this paper removes error and uncertainty resulting from this choice.

To access the accuracy of the method, it is instructive to review several recent publications that also attempt to predict binding affinity of small organics to biomolecules. Two methodologies are currently popular: QSAR activity fitting and docking. In recent studies based on catalytic site docking, ligand binding affinity was estimated for inhibitors of cAMP-dependent protein kinase (*1*). These studies used first principles docking energetic and thermodynamic methodologies. A series of 19 ligands were analyzed to predict calculated binding energies. The accuracy of the predicted energies varied widely, with some predictions in error $\Delta G/RT$ of >28. The average error in these calculations was a $\Delta G/RT$ of >21. From the perspective of an investigator wishing to narrow a focus on the best inhibitors, a major weakness in this and related computational approaches is that the errors are not uniform. For example, in the study of protein kinase, one of the strongest binders with an experimental binding energy of −10.3 kcal/mol was theoretically

found to be among the weaker inhibitors. At the same time, some of the weaker binders are predicted to have the highest binding energy. Thus, there is both quantitative error and qualitative error. In contrast, the average error in our predictions has a $\Delta G/RT < 2$ and the inhibitors are appropriately ranked in order of inhibitory strength.

A second recent study (*2*) utilized Comparative Molecular Field Analysis with QSAR to predict known binding energies of cyclic urea derivatives when used as HIV-1 protease inhibitors. As described above, QSAR analysis requires that the investigator postulate a QSAR model, and so QSAR predictions are not unique. Because QSAR is model-based, three separate predictions were made for groups of inhibitors with known binding energies corresponding to three QSAR models. In the tests of simple cyclo urea derivatives, for example, when the three models were trained with 40 molecules in the first case and 43 in each of the second and third, the three QSAR predictions yielded average errors of 2.07, 1.73, and 2.40 dimensionless $\Delta G/RT$ units, respectively. The range of experimental reported K_i values for the cyclo urea inhibitors varied by a factor of 64 000. Our inhibitors analyzed by the neural network varied in K_i by a factor of 660 000. Thus, the 3 QSAR models after training with more than twice as many known closely related inhibitors, produced predictions comparable to the neural net in one case, and significantly less accurate in the two other cases. The range of inhibitor variability was smaller than that considered in the research reported in this paper. When large families of unknown inhibitors are to be screened, the selection of a QSAR model becomes problematic, but is a strength for the model-independent approach of the neural network.

A third example describes application of a state-of-the-art 3D-QSAR algorithm to predict inhibitory potential (*3*). The approach was applied to a series of 21 steroids in the training set to predict binding affinity 10 realted steroids to the corticosteroid binding globulin. With similar size training and predicting sets, the errors here were on average 1.7 $\Delta G/RT$ the same as our value of 1.7 $\Delta G/RT$ reported in Figure 4. The similarity of the steroid ring systems simplifies binding prediction. The range of binding energy for the test steroids (a limited structural range) is only 5.75 $\Delta G/RT$ units as compared to the inhibitors tested here, with a range of 13.4 $\Delta G/RT$. This QSAR method is clearly accurate and does not mis-predict strong and weak binders over a small range of K_i values and with limited structural variation. This 3D-QSAR approach is the most sophisticated QSAR currently available, and it is apparent that our method is of comparable accuracy with greater variability in the structure of the unknown compounds, and accurate over a much larger range of binding affinities.

Reliable prediction of inhibitor strength must recognize the atomic forces at enzymatic catalytic sites, and therefore, depends heavily on quantum mechanically derived features. Inspection of the weaker inhibitors in the training set emphasizes this point. The weakest-binding iminoribitol in the training set is 3-pyridineiminoribitol (t13). It binds with less energy than iminoribitol alone (iminoribitol $\Delta G/RT = -12.3$, 3-pyridine-iminoribitol $\Delta G/RT = -11.7$). The pyridine group interaction with the leaving group site is unfavorable, since the phenyl substituent (t8) binds more favorably. The pyridine nitrogen alters the favorable electrostatic potential of the inhibitor. The neural network is superior at recognizing the important electrostatic features molecules need to bind IU-NH. The information is presented to the neural network during training and is used by the network to make accurate predictions. The network functions as a sophisticated structure−activity relationship machine. It determines the quantum structural features important for binding and forms a reliable association of these with binding energy.

ACKNOWLEDGMENT

The authors acknowledge the generosity of Drs. P.C. Tyler, G.B. Evans, and R.H. Furneaux in providing the iminoribitol analogues, and Dr. C. -H. Wong for p5, p16, p17, and p18.

REFERENCES

1. Hunenberger, P. H., Helms, V., Narayana, N., Taylor, S. S., and McCammon, J. A. (1999) *Biochemistry 28*, 2358−2366.
2. Debnath, A. K. (1999) *J. Med. Chem. 42*, 249−259.
3. Kubinyi, H., Hamprecht, F. A., and Mietzner, T. (1998) *J. Med. Chem 41*, 2553−2564.
4. Degano, M., Almo, S. C., Sacchettini, J. C., and Schramm V. L. (1998) *Biochemistry 37*, 6277−6285.
5. Parkin, D. W., Limberg, G., Tyler, P. C., Furneaux, R, H., Chen, X. Y., and Schramm, V. L. (1997) *Biochemistry 36*, 3528−3534.
6. Horenstein, B. A., Parkin, D. W., Estupinan, B., and Schramm, V. L. (1991) *Biochemistry 30*, 10788−10795.
7. Gasteiger, J., Li, X., Rudolph, C., Sadowski, J., and Zupan, J. (1994) *J. Am. Chem. Soc. 116*, 4608−4620.
8. Wagener, M., Sadowski, J., and Gasteiger, J. (1995) *J. Am. Chem. Soc. 117*, 7769−7775.
9. Weinstein, J. N., Kohn, K. W., Grever, M. R., Viswanadhan, V. N., Rubinstein, L. V., Monks, A. P., Scudiero, D. A., Welch, L., Koutsoukos, A. D., Chiausa, A. J., and Paull, K. D. (1992) *Science 258*, 447−451.
10. Hammond, D. J., and Gutteridge, W. E. (1984) *Mol. Biochem. Parasitol. 13*, 243−261.
11. Parkin, D. W., and Schramm, V. L. (1995**)** *Biochemistry 34*, 13961−13966.
12. Horenstein, B. A., and Schramm, V. L. (1993) *Biochemistry 32*, 7089−7079.
13. Horenstein, B. A., and Schramm, V. L. (1993) *Biochemistry 32*, 9917−9925.
14. Boutellier, M., Horenstein, B. A., Semeyaka, A., Schramm, V. L., and Ganem, B. (1994) *Biochemistry 33*, 3994−4000.
15. Miles, R. W., Tyler, P. C., Evans, G., Furneaux, R. H., Parkin, D. W., and Schramm, V. L. (1999) *Biochemistry, 38*, xxxx-xxxx.
16. Gopaul, D. N., Meyer, S. L., Degano, M., Sacchettini, J. C., and Schramm, V. L. (1996) *Biochemistry 35*, 5963−5970.
17. Politzer, P., and Truhlar, D. G., Eds. (1981) Chemical Applications of Atomic and Molecular Electrostatic Potentials, Plenum Press, New York.
18. Sjoberg, P., and Politzer, P. (1990) *J. Phys. Chem. 94*, 3959−3961.
19. *Gaussian 96, Revision C2;* Gaussian, Inc.: Pittsburgh, PA, 1995.
20. Bagdassarian, C. K., Schramm, V. L., and Schwartz, S. D. (1996) *J. Am. Chem. Soc. 118*, 8825−8836.
21. Fausett, L. *Fundamentals of Neural Networks;* Prentice Hall: New Jersey, 1994.
22. Rumelhart, D. E., Hinton, G. E., and Williams, R. J *Parallel Distributed Processing*, Vol. 1; MIT Press: Massachusetts, 1986.
23. Parkin, D. W., Horenstein, B. A., Abdulah, D. R., Estupiñán, B., and Schramm, V. L. (1991) *J. Biol. Chem. 266*, 20658−20665.

BI990830T