

Assessment of Solvation Effects on Calculated Binding Affinity Differences: Trypsin Inhibition by Flavonoids as a Model System for Congeneric Series.

Ana Checa, Angel R. Ortiz,^{*,†} Beatriz de Pascual-Teresa,[‡] and Federico Gago^{*}

Departamento de Farmacología, Universidad de Alcalá, E-28871 Alcalá de Henares, Madrid, Spain

Received April 14, 1997

On the basis of molecular models of the interaction between trypsin and a series of seven structurally congeneric bioflavonoid inhibitors, the influence of solvation effects in the calculation of binding free energy differences in congeneric series has been assessed. The models were derived by making use of the X-ray crystal structure of bovine trypsin and the DOCK program, and the complementarity of the interactions between the functional groups of the docked molecules and the binding site region was corroborated independently with the GRID program. Interaction energies calculated for the complexes using molecular mechanics were found to correlate with the experimental inhibitory activities, although the quality of the correlation was dependent on the molecular modeling protocol. To understand why such correlations could be obtained in the absence of an explicit description of solvent effects, the *in vitro* activities were transformed into binding free energies, and continuum electrostatic theory was used to incorporate solvent effects by approximating them to the electrostatic contribution to the binding free energies. The results of our calculations show that, within this congeneric series, the cost in electrostatic free energy of desolvating both the enzyme binding site and the buried part of the inhibitors (ΔG_{desolv}) is roughly constant within the series. On the other hand, the electrostatic interaction energy ($E_{\text{elc}}^{\text{LR}}$) varies only slightly along the series in comparison with the van der Waals interaction ($E_{\text{VDW}}^{\text{LR}}$), and this variation is mostly solvent-independent, i.e., the reaction field energy of the solvent in the bound state ($E_{\text{srf}}^{\text{LR}}$) makes almost a negligible contribution to the binding free energy differences. As a result, differences in binding free energy are dominated by the van der Waals term, while the electrostatic contribution is, to a good approximation, solvent-independent. A similar scenario may account for the good correlations frequently found between ligand activities and ligand-receptor interaction energies derived using plain molecular mechanics, although generality remains to be determined.

Introduction

One of the major bottlenecks in current structure-based drug design efforts is our inability to evaluate binding free energies correctly in order to rank different ligand candidates.¹⁻⁴ When a series of ligand-receptor complexes is available, as is generally the case in the process of structure-based lead optimization, it becomes particularly important to derive valid correlations between calculated binding energies and experimental binding or activity data.³ If binding affinities of new compounds can be reliably predicted in advance of their synthesis, the drug design process can be made more effective.

For the calculation of ligand-receptor interaction energies, most approaches rely on molecular mechanics force fields that represent van der Waals and Coulombic interactions on the basis of empirical energy functions. Using these potentials, inclusion of explicit solvent molecules and use of free energy perturbation methods can yield good agreement between experimental and calculated free energy differences when the initial and final states are rather similar.⁵ Since this approach is impractical for the comparative study of large series of

receptor-ligand complexes, severe simplifications are often made, such as modeling the screening effect of the solvent throughout the use of a so-called "distance dependent dielectric constant" or ignoring almost systematically the desolvation and entropic contributions. Intermolecular interaction energies calculated this way have been used either *per se*⁶⁻⁹ or through adequate weighting of selected pairwise interactions (COMBINE approach)¹⁰ to yield receptor-based quantitative structure-activity relationships (QSARs).¹¹

Many of these studies have been rather successful despite the fact that they usually consider only differences in binding enthalpies due to ligand-receptor interactions but ignore solvent effects. This is a surprising result, and the reason for the acceptable performance in so many cases of molecular mechanics based QSAR models that neglect solvation effects is unclear since, to the extent that binding represents a desolvation process,¹² ignoring the loss of solute-solvent interactions and the gain in solvent-solvent interactions can be an important source of error. In fact, it has been pointed out that differential solvation/desolvation energies within a closely related series of ligand molecules may vary more significantly than intrinsic ligand-receptor interaction energies.¹³ Nevertheless, attempts to improve the quality of some QSAR models by incorporating solvation effects have been unsuccessful⁸ or only moderately successful.⁹

If we agree that success in correlating affinity or activity data with intermolecular interaction energies

* Address correspondence to either author. F.G.: Telephone: +34-1-885 45 14. Fax: +34-1-885 45 91. E-mail: fgago@fisfar.alcala.es.

† Present address: Department of Molecular Biology, TPC-5, The Scripps Research Institute, La Jolla, CA 92037 (e-mail: ortiz@scripps.edu).

‡ Present address: Departamento de Química, Universidad San Pablo CEU, E-28668 Madrid.

© Abstract published in *Advance ACS Abstracts*, November 15, 1997.

may depend on a large amount of cancellation among the contributions that are neglected and assume that these contributions are not implicitly included in the variables that are considered in the analysis, two different scenarios can be envisaged: either these cancellations are fortuitous and mostly related to enthalpy–entropy compensations in aqueous solutions,³ or some of the neglected terms do not contribute to the binding free energy differences because they are very similar for all the members of the series, especially when they are congeneric.^{6,10} Distinguishing between these two extreme cases is important if we are to improve our methodologies of ligand ranking and design. Thus, if a “random compensation model” holds, there is little hope that the activity of new structures can be reliably predicted. On the contrary, if a “constant term model” is operative, it might be possible to set limits about the reliability of the application of a regression equation to a new compound on the basis of its structural similarity with the training set. Of course, universality is not warranted, and different models can apply to different systems. Moreover, as suggested by one of the reviewers, there could be other cases characterized by mixtures of fortuitous partial cancellation of errors and roughly similar values of neglected terms.

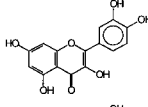
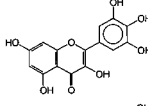
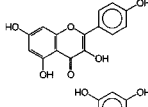
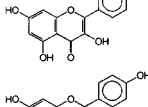
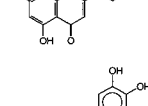
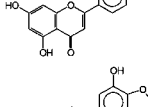
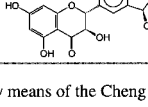
In the following, we have attempted to shed some light on these issues by using trypsin as a model system. Trypsin is a member of the family of serine proteinases, whose catalytic behavior and basis for specificity are well understood in structural terms.¹⁴ The catalytic triad is made up of a serine side chain in close vicinity to a histidine imidazole hydrogen bonded to the carboxylate side chain of a buried aspartic acid. The specificity of trypsin for peptide bonds following an arginine or a lysine residue is a consequence of the accommodation of the side chain of these basic residues into a crevice on the enzyme surface whose bottom is occupied by an aspartic acid side chain (Asp189 in the case of bovin trypsin) which stabilizes the interaction. Both natural and synthetic amine or amidine-based inhibitors are known which make use of this *specificity pocket* to block the access of the polypeptide substrate.^{15,16}

A screening program aimed at discovering new proteinase inhibitors from natural sources recently unveiled that quercetin and other flavonoid compounds devoid of a net positive charge and with no amino or amidino substituents (Table 1) are also potent competitive inhibitors of trypsin.^{17,18} We thought it might be interesting to apply current theoretical methods to both the prediction of the binding mode of these inhibitors and the estimation of their relative binding energies.¹⁹ From a modeling perspective, the trypsin–flavonoid system offers two important advantages: (i) these inhibitors have the torsion about the C2–C1' bond as the only major degree of freedom and (ii) the enzyme conformation is known to be very weakly affected by inhibitor binding, which reflects the rather rigid properties of the active site.^{15,16} From a drug design perspective, flavonoids are attractive as they can be regarded as potential nonpeptidic synthons²⁰ with a wide range of pharmacological actions.²¹

Methodology

(a) Model Building and Parameterization of the Inhibitors. The crystal structures of quercetin,²² nar-

Table 1. Set of Flavonoids Described as Inhibitors of Bovin Trypsin Enzyme Activity^{17,18}

Flavonoid	Chemical structure	IC ₅₀ (μM)	K _i (μM) ^a	ΔG (kcal mol ⁻¹) ^b
quercetin		7.1±2	1.98	-7.88
myricetin		10.2±2	2.84	-7.66
kaempferol		105.9±22	29.53	-6.26
morin		110.8±33	30.9	-6.23
apigenin		141.5±17	39.46	-6.08
luteolin		35.3±6	9.84	-6.92
silychristin		144.4±18	40.27	-6.07

^a Obtained by means of the Cheng and Prusoff equation⁴⁷ from the K_i of quercetin (see text).

^b Obtained by means of the relationship $\Delta G^\circ = RT \ln K_i$.

ingenin,²³ and silybin,²⁴ retrieved from the Cambridge Structural Database,²⁵ were used as templates for constructing the inhibitors.²⁶ The geometry of the molecules was optimized by using the semiempirical molecular orbital AM1 method,²⁷ as implemented in the Spartan program,²⁸ until the energy difference between successive cycles was below 0.0005 kcal mol⁻¹. Atom-centered charges (Supporting Information) were then derived by fitting the calculated molecular electrostatic potential to a monopole–monopole expression.²⁹ The same sets of charges and radii were used in programs DOCK, AMBER, and DelPhi, described below, and all calculations were run on a Control Data Cyber-480 workstation. Covalent and nonbonded parameters for the bioflavonoids were derived, by analogy or through interpolation,³⁰ from those already present in the AMBER database (Supporting Information). One new atom type was introduced (CX), defined as a nonaromatic sp² carbon singly bonded to two other carbon atoms (flavones and flavonols), for which the van der Waals parameters of the AMBER C atom type were taken. The torsional parameters for the C2–C1' bond ($V_n/2 = 2.05$, $\gamma = 180$, $n = 3$) were estimated by a combination of molecular mechanics and quantum mechanics calculations³¹ on 3,4-dihydroxyflavone.

(b) Docking Studies. Programs GRID³² and DOCK³³ were used independently for docking purposes. The structure of bovin trypsin crystallographically determined at 1.55 Å resolution,¹⁵ retrieved from the Brookhaven Data Bank³⁴ (PDB entry 2ptn), was used as the target receptor. Water molecules were not included in the calculation, but the calcium ion was preserved and assigned a charge of +2. For the GRID studies, a lattice of points spaced at 0.5 Å was established throughout and around the active site Ser195 in

order to search for binding sites complementary to the functional groups of the flavonoids. Then, the interactions between the active site pocket and aromatic carbon (C1=), ether oxygen (OC2), carbonyl oxygen (O), and phenolic oxygen (OH) probes were calculated. An aliphatic oxygen probe (O1) was also used for the entire protein in order to assess the goodness of the predicted binding sites for water. The dielectric constants chosen were 4.0 for the macromolecule and 80.0 for the bulk water. The resulting grids were contoured at appropriate energy levels (in kcal mol⁻¹: C1=, -2.5; OC2, -3.5; O, -2.5; OH, -5.5) and graphically displayed.²⁶ The geometries of the inhibitors benzamidine and *p*-amidinophenyl pyruvate, as found in their crystal complexes with trypsin (PDB entries 3ptb and 1tpp),¹⁵ were used as positive controls.

For the DOCK studies, a molecular surface representation³⁵ of all residues within 15 Å of the active site Ser195 of trypsin was obtained by using the MS program³⁶ and a probe of 1.4 Å radius. The surface points and their associated normals were used by program SPHGEN^{33a} to fill the active site with spheres, whose volume provided a geometric description of the volume available to the ligands. No sphere radius was larger than 4 Å or smaller than 1.4 Å. This cluster containing 69 spheres was enclosed in a 23 × 20 × 25 Å³ box which was gridded so that each point in the cubic lattice could be evaluated as making favorable or unfavorable contacts with receptor atoms (*contact scoring*) or used to store steric and electrostatic information of the receptor atoms within a 10.0 Å cutoff of the point (*force field scoring*). Polar and nonpolar contact limits for filtering were 2.3 and 2.8 Å. The ligand molecules were then input to DOCK in the AMBER-optimized conformation in order to examine their preferred binding orientations.

(c) Molecular Mechanics Calculations. All-atom AMBER force field parameters^{30,37} were used for the inhibitors and those trypsin residues within 15 Å of the active site Ser195. Nonpolar hydrogens of the rest of the protein were treated by way of the united-atom approach.³⁸ The complexes were refined by means of a steepest descent energy minimizer, using a cutoff of 10.0 Å and a distance-dependent dielectric constant ($\epsilon = r_{ij}$ or $\epsilon = 4r_{ij}$) in two steps: first, only the flavonoid molecule and the hydrogen atoms of the protein were allowed to move, and then the flavonoid and any protein residues within a 10 Å radius sphere were allowed complete relaxation. Those amino acids comprised in a 10–12 Å buffer zone were restrained to their crystallographic positions by means of a harmonic potential with a force constant of 5 kcal mol⁻¹ Å⁻² and all protein residues beyond 12 Å of Ser195 were restrained to their starting locations although they were included in the determination of the forces. The energy minimization proceeded until the root-mean-square value of the potential energy gradient was below 0.1 kcal mol⁻¹ Å⁻¹. Additional models were built which incorporated either 4 (416, 702, 710, 808) or 10 (402, 408, 414, 415, 416, 562, 702, 703, 704, 705) crystallographic water molecules in the vicinity of the inhibitors. A structural role for some of these buried water molecules has been proposed.³⁹ Since a part of each inhibitor molecule is directed toward the solvent, a spherical "cap" of about 530 TIP3P water molecules, generated from a Monte

Carlo simulation,⁴⁰ was also added to the complex centered on the O γ oxygen atom of Ser195. The solvated complexes were minimized in two steps: first, only the water molecules were allowed to reorient, and then the water molecules, the flavonoid, and those protein residues containing any atoms within this water sphere were allowed to relax whereas the residues outside were fixed at their starting locations.

(d) Continuum Electrostatics Calculations. A macroscopic solvent model based on the Poisson equation can be efficiently used to describe the dynamically averaged dielectric behavior of the solvent environment.^{41,42} Here, finite difference solutions to the linearized Poisson equation,⁴³ as implemented in the DelPhi module of Insight-II,²⁶ were used to calculate electrostatic potentials and energies. Since we were interested in dissecting out the different contributions to the electrostatic free energy differences (see below), and some the components, such as the reaction field energy, can be correctly computed in DelPhi only at zero ionic strength, the Poisson equation rather than the Poisson–Boltzmann equation was used. Cubic grids with a resolution of 0.5 Å were centered on the molecular systems considered, and the charges were distributed onto the grid points.⁴⁴ Solvent-accessible surfaces,³⁵ calculated with a spherical probe of 1.4 Å radius, defined the solute boundaries, and a minimum separation of 11 Å was left between any solute atom and the borders of the box. The potentials at the grid points delimiting the box were calculated analytically by treating each charge atom as a Debye–Hückel sphere.⁴⁵ Since the Poisson equation was solved, this amounts to calculating the electrostatic potential generated by the rest of the charges at these points. The accuracy of the calculated electrostatic potentials was subsequently improved by defining a smaller grid with a lower resolution (0.25 Å spacing) and using boundary potentials linearly interpolated from those calculated in the previous run.⁴⁴ By using this finer grid, the numerical results are much less sensitive to the orientation of the molecules on the lattice of points.⁴⁶

The atomic coordinates employed were those of the free inhibitors and the AMBER-optimized complexes that yielded the best correlation with inhibitory activity. The interior of the protein, the ligands and the complexes was considered a low dielectric medium ($\epsilon = 4$) whereas the surrounding solvent was treated as a high dielectric medium ($\epsilon = 80$).

(e) Calculation of Free Energy Differences. Given that the mechanism of inhibition appears to be the same for all the inhibitors studied and the assays were performed under the same conditions,^{17,19} direct comparison of IC₅₀ values was considered sufficient to determine the relative efficacies. Moreover, since the dissociation constant (K_I) for the quercetin–trypsin complex (Q) was known,¹⁶ use of the relationship⁴⁷

$$\frac{(\text{IC}_{50})_F}{(\text{IC}_{50})_Q} = \frac{K_{I_F}}{K_{I_Q}}$$

allowed us to calculate the inhibition constants (K_I) for the rest of the flavonoids (F), which were then transformed into binding free energies by means of the equation

$$\Delta G^\circ = RT \ln K_I$$

The differences in free energies of binding to a common receptor were calculated for each flavonoid (F) relative to quercetin (Q) as

$$\Delta\Delta G_F = \Delta\Delta G_{\text{vdw}} + \Delta\Delta G_{\text{ele}}$$

where $\Delta\Delta G_{\text{vdw}} = \Delta G_{\text{vdw}}(\text{F}) - \Delta G_{\text{vdw}}(\text{Q})$ and $\Delta\Delta G_{\text{ele}} = \Delta G_{\text{ele}}(\text{F}) - \Delta G_{\text{ele}}(\text{Q})$, assuming similar conformational and vibrational entropy differences for all the complexes considered. This is reasonable given the congeneric nature of the ligands and the rigidity of the binding site.^{15,16} The van der Waals interaction energies (ΔG_{vdw}) were taken from the molecular mechanics calculations whereas the changes in electrostatic free energy on molecular association (ΔG_{ele}) were calculated using the continuum method, which takes into account both the desolvation of ligands and receptor and the effects of the surrounding solvent. ΔG_{ele} can be separated into three components:⁴⁸ (i) the change in solvation energy of the receptor upon binding ($\Delta G_{\text{desolv}}^{\text{R}}$), (ii) the change in solvation energy of the ligand upon binding ($\Delta G_{\text{desolv}}^{\text{L}}$), and (iii) the ligand–receptor interaction energy in the presence of the surrounding solvent ($E_{\text{ele}}^{\text{LR}}$). The first two terms, corresponding to differences in electrostatic free energy of solvation, can be readily calculated by considering the effects on the respective electrostatic energies of replacing the high dielectric medium of the solvent with the low dielectric medium of the other molecule in those regions that are occupied by the binding partner in the complex (Figure 1). The last term can be obtained from:

$$E_{\text{ele}}^{\text{LR}} = \sum_i^L \phi_i \cdot q_i$$

where ϕ_i is the potential generated by the charges on the receptor atoms at the location of the charge of each ligand atom (q_i) and the summation is over all of the atoms of the ligand (L).

To estimate the precision of the calculation, the overall electrostatic free energy change upon binding (ΔG_{ele}) was also calculated in an alternate way from the total electrostatic energy of the system by running three consecutive calculations on the same grid: one for all the atoms in the complex ($G_{\text{ele}}^{\text{LR}}$), one for the ligand atoms alone ($G_{\text{ele}}^{\text{L}}$), and a third one for the receptor atoms alone ($G_{\text{ele}}^{\text{R}}$). Since the grid definition is the same in the three calculations, the artifactual grid energy cancels out and the electrostatic contribution to the binding free energy can be expressed as the difference in energy between the product and the reactants: $\Delta G_{\text{ele}} = G_{\text{ele}}^{\text{LR}} - (G_{\text{ele}}^{\text{L}} + G_{\text{ele}}^{\text{R}})$.

In turn, two different contributions can be considered for the electrostatic energy of the ligand–receptor complex ($G_{\text{ele}}^{\text{LR}}$), as well as for the electrostatic component of the ligand–receptor interaction energy ($E_{\text{LR}}^{\text{ele}}$): the electrostatic energy in a dielectric medium equal to that of the complex, and that which arises from the reaction field caused by the surrounding solvent. The contribution of the reaction field of the solvent to $E_{\text{LR}}^{\text{ele}}$ ($E_{\text{sr}}^{\text{LR}}$) was estimated as the difference between the energy calculated with the exterior dielectric set to 80

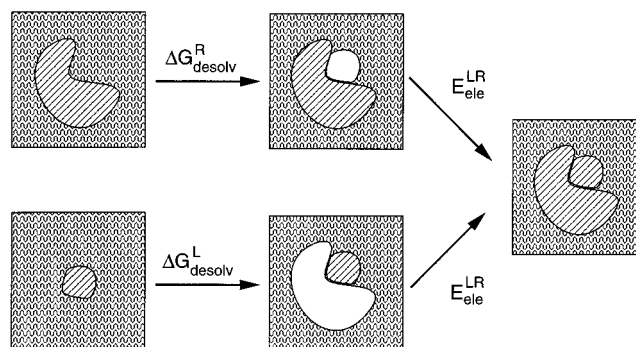


Figure 1. Schematic representation for calculating the electrostatic contribution to the free energy of binding (ΔG_{ele}). The wavy pattern represents the high dielectric solvent ($\epsilon_s = 80$) that surrounds the low dielectric solutes ($\epsilon_m = 4$). Hatched or blank areas in the solutes denote the presence or absence, respectively, of atomic charges in the calculation. ΔG_{desolv} stands for the loss of electrostatic interaction between the solvent and either the receptor (R) or the ligand (L) whereas $E_{\text{ele}}^{\text{LR}}$ represents the electrostatic interaction between ligand and receptor in the presence of surrounding solvent.

and that calculated using a value of 4 for both the complex environment and the dielectric interior.

The electrostatic contribution to the solvation free energy of each inhibitor ($\Delta G_{\text{ele}}^{\text{solv}}$) was also calculated. It was obtained by subtracting the total electrostatic free energy in vacuo ($G_{\text{ele}}^{\text{vac}}$) from the total electrostatic free energy in water ($G_{\text{ele}}^{\text{wat}}$).^{46b,48} To this end, two calculations with identical grid mappings and the same interior dielectric ($\epsilon = 4$) were run for each inhibitor. The exterior dielectric was set to either 80, to simulate the water environment, or 1, to reproduce vacuum conditions. This method was useful to estimate the differences in electrostatic solvation free energies among different conformers (e.g. free and bound conformations).

Results

(a) Docking Studies. Trypsin-bound benzamidine^{15a} and *p*-amininophenyl pyruvate^{15b} were used as positive controls to test the programs used for docking. The GRID map obtained for the aromatic carbon probe highlighted the favorable binding energy of the phenyl ring of the inhibitors sandwiched between segments Trp215–Gly216 and Ser190–Gln192, which form the *ceiling* and *floor* of the specificity pocket, respectively,¹⁵ and the O1 probe correctly predicted the position of the buried water molecule (WAT 805) that is found hydrogen bonded to Asp189 and Ser190 in free trypsin. DOCK, on the other hand, produced a variety of binding orientations when supplied with the coordinates of ligand-free trypsin and the bound conformations of the inhibitors if only *contact scoring*^{33a,33b} was used, but was able to reproduce the crystallographic binding orientations of both ligands when this function was supplemented with a measure of electrostatic complementarity (*force-field scoring*).^{33c} For benzamidine, the root-mean-square deviation (rmsd) from the crystallographic position for the best scoring conformation was only 0.52 Å. In the case of *p*-amininophenyl pyruvate, a larger value was found (rmsd = 1.39), which was attributed to the intermolecular covalent bond present in the crystal structure. Thus, the orientation found for this ligand, which is consistent with the observed binding mode, can

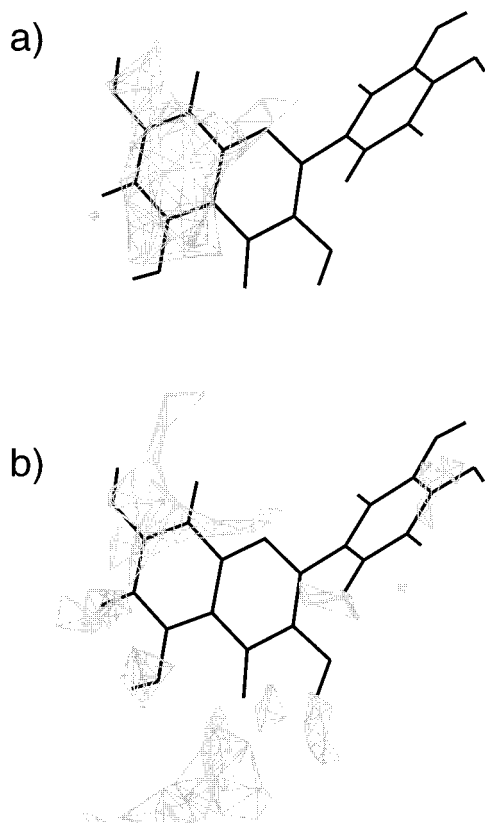


Figure 2. Correspondence between the GRID maps for (a) the sp^2 carbon (contoured at $-2.5 \text{ kcal mol}^{-1}$) and (b) phenolic OH (contoured at $-5.5 \text{ kcal mol}^{-1}$) probes and the orientation of bound quercetin as suggested by program DOCK using force-field scoring.

be reasonably expected to be the conformation of the inhibitor prior to covalent bond formation. When the same procedure was employed for the flavonoids, a distinctly preferred orientation was obtained for binding each inhibitor, and good matches were found between the functional groups of the ligands in the bound conformation and the GRID maps obtained for free trypsin (Figure 2).

The resulting binding orientations (Figure 3) had in common (i) the burial of the chromone ring system into the specificity pocket and the displacement of WAT 805, whose position is occupied by the phenolic OH at C7, (ii) establishment of a good hydrogen bond between the OH of the active site Ser195 and the γ -pyrone carbonyl, and (iii) projection of the substituent on C2 toward the solvent, which allows the best inhibitors to have the 3'-hydroxyl group on the phenyl ring engaged in a hydrogen-bonding interaction with the NH of Gly219. Remarkably, this NH group is involved in a tight hydrogen bond to a water molecule (WAT 804) from a symmetry-related copy in the crystal structure of free trypsin. In the set of complexes studied, its engagement in a hydrogen bond with the 3'-OH group of a given flavonoid will be modulated by both the nature of the C2-C1' bond (sp^2 carbon- sp^2 carbon in flavones and derivatives vs sp^3 carbon- sp^2 carbon in flavanones) and the presence or absence of a hydroxyl at C3.

The extent of conformational freedom in this series of inhibitors is largely limited to rotation around the C2-C1' bond, which determines the twist of the exocyclic phenyl ring and the existence of two possible low-energy conformations for each flavonoid. The ones

selected for binding to trypsin display values for the O1-C2-C1'-C2' torsional angle ranging from around -30° (apigenin and luteolin, flavones with a hydrogen atom on C3) to around -40° (all the rest, except morin), in good agreement with values found in X-ray crystal structures of flavonoids.⁴⁹ Morin is a flavonol with a hydroxyl group at the 2' position (Table 1) which, in our modeled structures, cannot make the same hydrogen-bonding interaction with the NH of Gly219 as quercetin and myricetin can, and for which the best binding orientation is found to be that with the phenyl ring at an angle of $\sim 40^\circ$ (see legend to Figure 3). As will be shown below, this distinction may have important implications as the two virtually isoenergetic conformers show differences in their solvation free energies that may favor one over the other in solution. Saturation of the C2-C3 bond brings about the puckering of the γ -pyrone ring, an increase in the length of the C2-C1' bond, and the equatorial positioning of the exocyclic phenyl ring, which is thus out of plane with the rest of the molecule. As a result, the distance between the 3'-OH of silychristin and the NH of Gly-219 is increased by about 1 Å, and no hydrogen bond can be formed. These changes in geometry are consistent with the lower inhibitory activity of flavanones and flavanols relative to flavones and flavonols.^{17,18}

(b) Correlation of Inhibitory Potencies with Calculated Binding Energies. The complexes resulting from the previous docking study were refined by following different energy minimization strategies (see Methodology). The calculated intermolecular interaction energies were found to correlate well with the inhibitory activity these flavonoids showed in vitro, although the quality of the regression was sensitive to the minimization protocol (Figure 4). Interestingly, and contrary to some suggestions, introduction of only a few crystallographic water molecules around the bound ligand slightly deteriorated the correlation (Figure 4b), presumably as a result of the incomplete treatment of the protein-solvent interactions in the binding site, which gives rise to small distortions that affect the energetics of the ligand-protein interaction. In fact, when a cap of water molecules completely surrounded the active site region, a considerably better behavior was observed (Figure 4c). The worst results, however, were obtained with the strong damping of electrostatic interactions provided by the $4r_{ij}$ dielectric constant, which would indicate that a correct treatment of electrostatics is essential in order to model ligand-receptor complexes correctly. A mild energy refinement using a dielectric value equal to the interatomic distance produced the best correlation coefficient and the correct ranking of all the inhibitors (Figure 4a). Analysis of the interactions with a dielectric constant value of either 2 or 4 slightly worsened the correlation ($r^2 = 0.884$ and 0.863 , respectively). By comparing the qualities of the different correlations, it can be argued that it is probably wise to keep changes to just the essential induced-fit conformational rearrangements and that analysis of the interaction energies using a distance-dependent dielectric constant ($\epsilon = r_{ij}$) is probably the most adequate method. If crystallographic molecules are to be included around the binding site, our results would suggest that it is preferable to include a whole "cap" of water



Figure 3. Stereoview of the quercetin-trypsin complex. For clarity, only the α -carbon backbone atoms (thin lines), the active site Ser195, His57 and Asp102 residues, and Asp189 of trypsin (thick lines) are shown; the cross represents the bound calcium ion. In the morin-trypsin complex, the exocyclic phenyl ring of morin is rotated counterclockwise so that the O1-C2-C1'-C2' dihedral angle is positive instead of negative.

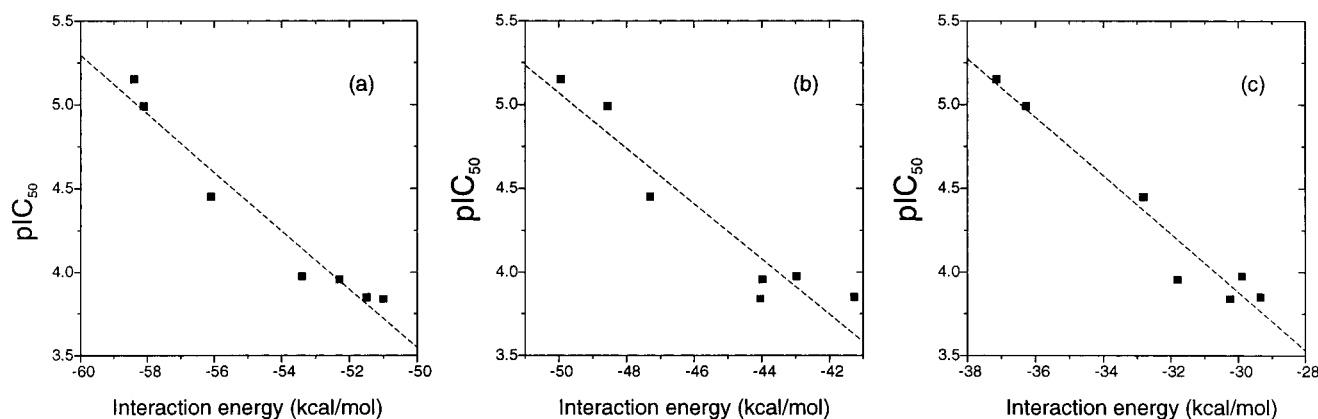


Figure 4. Correlations between in vitro activities and calculated binding energies using different modeling protocols and a distance-dependent dielectric model ($\epsilon = r_{ij}$): (a) no explicit waters ($r^2 = 0.953$), (b) four crystallographic water molecules ($r^2 = 0.897$), (c) water "cap" ($r^2 = 0.951$). The model with 10 crystallographic water molecules yielded a $r^2 = 0.768$. See Methodology for details.

Table 2. Electrostatic Energy Contributions to the Binding Free Energy

flavonoid	$\Delta G_{\text{desolv}}^{\text{R}}$	$\Delta G_{\text{desolv}}^{\text{L}}$	ΔG_{desolv}	$E_{\text{ele}}^{\text{LR}}(4, 80)^a$	ΔG_{ele}^b	ΔG_{ele}^c	$E_{\text{ele}}^{\text{LR}}(4, 4)^d$	$E_{\text{ele}}^{\text{LR}} \text{ AMBER}^e$	$E_{\text{srf}}^{\text{LR}}{}^f$
quercetin	5.24	3.03	8.27	-6.35	1.92	1.99	-5.95	-5.88	-0.40
myricetin	5.00	3.10	8.10	-6.29	1.81	1.85	-5.79	-5.77	-0.50
kaempferol	5.05	2.85	7.90	-5.37	2.53	2.61	-4.76	-4.66	-0.61
morin	5.18	2.88	8.06	-5.38	2.68	2.72	-4.86	-4.79	-0.52
apigenin	5.28	2.85	8.13	-5.23	2.90	2.74	-4.90	-4.70	-0.33
luteolin	5.47	2.99	8.46	-6.02	2.43	2.49	-5.72	-5.54	-0.30
silychristin	4.63	3.30	7.93	-5.51	2.42	2.42	-5.82	-5.64	0.31

^a Delphi calculation; interior dielectric = 4, exterior dielectric = 80. ^b $\Delta G_{\text{ele}} = \sum(q_i \phi) + \Delta G_{\text{desolv}}$. ^c $\Delta G_{\text{ele}} = G_{\text{ele}}^{\text{LR}} - (G_{\text{ele}}^{\text{L}} + G_{\text{ele}}^{\text{R}})$. ^d Delphi calculation; interior dielectric = exterior dielectric = 4. ^e AMBER calculation; dielectric constant = 4; no cutoff. ^f $E_{\text{srf}}^{\text{LR}} = E_{\text{ele}}^{\text{LR}}(4, 80) - E_{\text{ele}}^{\text{LR}}(4, 4)$.

molecules in order to properly balance the solute-solvent interactions.

(c) Precision of the Calculated Electrostatic Contributions to the Binding Free Energies. In this work, the contributions of the solvent to the binding process have been accounted for by representing the solvent as a homogeneous dielectric continuum. This has allowed us to compute both the influence of the surrounding solvent on the calculated electrostatic interaction energies and the changes in solute-solvent interaction energies that take place upon binding (Table 2). A measure of the possible uncertainty introduced in the continuum calculations when mapping the charges and the dielectric constants onto a grid of finite size was provided by comparing the AMBER electrostatic inter-

action energies ($\epsilon = 4$) with those calculated by DelPhi when a uniform dielectric of 4 was used for both the complex interior and the surrounding medium. The average difference amounted to about 0.1 kcal mol⁻¹, and this is also the degree of discrepancy in electrostatic interaction energies calculated by either $\Delta G_{\text{ele}} = G_{\text{ele}}^{\text{LR}} - (G_{\text{ele}}^{\text{L}} + G_{\text{ele}}^{\text{R}})$ or $\Delta G_{\text{ele}} = E_{\text{ele}}^{\text{LR}} + (\Delta G_{\text{desolv}}^{\text{L}} + \Delta G_{\text{desolv}}^{\text{R}})$ (Table 2).

(d) Calculation of Binding Free Energy Differences. The largest experimental free energy difference among the inhibitors studied amounts to about 2 kcal mol⁻¹ (Table 1) whereas the corresponding difference in calculated binding energies is of about 3 kcal mol⁻¹ (Table 3). The agreement between observed and pre-

Table 3. Differences in Free Energy Components Relative to Quercetin^a

flavonoid	$\Delta\Delta G_{\text{vdw}}$	$\Delta E_{\text{ele}}^{\text{LR}}(4, 4)$	$\Delta E_{\text{srf}}^{\text{LR}}$	$\Delta\Delta G_{\text{desolv}}$	$\Delta\Delta G$ (calcd)	$\Delta\Delta G$ (exptl)
myricetin	0.01	0.16	-0.10	-0.17	-0.12	0.22
kaempferol	1.93	1.19	-0.21	-0.37	2.56	1.62
morin	0.52	1.09	-0.12	-0.21	1.26 ^b	1.65
apigenin	2.11	1.05	0.07	-0.14	2.87	1.80
luteolin	1.44	0.23	0.10	0.19	1.94	0.96
silychristin	2.49	0.13	0.71	-0.34	2.92	1.81

^a See Table 2 and Methods. ^b Prior to the solvation energy correction for the bound state.

Table 4. Electrostatic Contribution to the Solvation Energy^a

flavonoid	conformer 1 ^b	conformer 2 ^c
quercetin	-9.76	-9.83
myricetin	-10.70	-10.12
kaempferol	-8.75	-8.65
morin	-9.11	-10.45
apigenin	-8.55	-8.60
luteolin	-9.67	-9.57
silychristin	-10.11	-10.47

^a Calculated on the AMBER-optimized geometries in the unbound state. ^b Negative O1-C2-C1'-C2' dihedral angle. ^c Positive O1-C2-C1'-C2' dihedral angle.

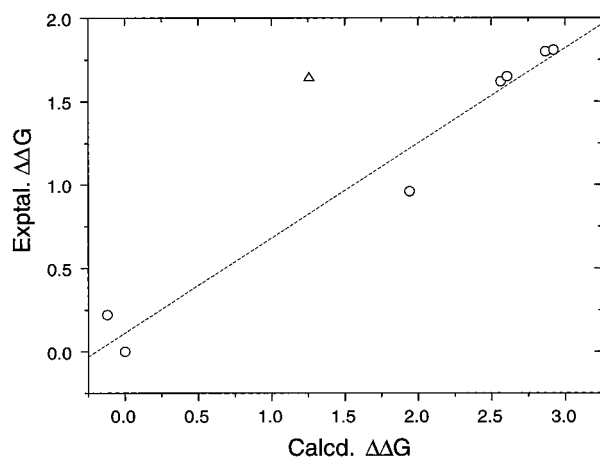


Figure 5. Correlation between experimental and calculated binding free energy differences. Values are relative to those of quercetin. Two values are given for morin, before (Δ) and after (\circ) correcting for the difference in the electrostatic contribution to the solvation free energy of the molecule in the bound state (Table 4).

dicted values is reasonably good but, interestingly, for morin it is necessary to correct for the difference in the electrostatic contribution to the solvation free energy between the two possible conformers in the unbound state. This term favors the "alternative" conformation, which can be presumed to be the major form in solution, over the selected bound conformation by about 1 kcal mol⁻¹ (Table 4). Introduction of this correction decreases the net binding energy for morin and improves the agreement between experimental and calculated free energy differences (Figure 5).

The level of agreement achieved in the computation of the different electrostatic contributions to the binding free energies (Table 2) allows us to feel confident about the significance of the partitioning scheme outlined in the Methodology section. The effect of the solvent on the computed binding free energy differences along the series is reflected in two sets of values (Table 3): the differences in the reaction field contributions to the ligand-receptor interaction energies in the bound state ($\Delta E_{\text{srf}}^{\text{LR}}$), and the differences in the overall electrostatic contributions to the desolvation energies upon binding ($\Delta\Delta G_{\text{desolv}}$). It is apparent that the values within each

set are rather small and roughly comparable for all the inhibitors even though it is interesting to note that both effects seem to be similarly important for this particular series. The screening of the solvent, which opposes binding, is particularly noticeable for silychristin, the largest flavonoid in the series, and that which retains a significant portion of the molecule in the bulk solvent upon complex formation. On the other hand, the solvent-independent van der Waals component shows a larger variation along the series and appears to dominate the binding free energy differences. The differences in the contribution of the solvent-independent electrostatic interaction energy, which are mainly related to the presence and orientation of the 3'-hydroxyl group, are comparatively much smaller, but they also provide a good ranking of the inhibitors' potencies.

Discussion

To rank the estimated binding affinities of a series of inhibitors, use is often made of differences in binding enthalpies that are usually approximated as differences in molecular mechanics-based interaction energies.⁶⁻⁹ In these calculations, the electrostatic energy term is crucially dependent on the dielectric constant chosen to represent the environment of the ligand in the bound state and treatment of the ligand in the unbound state is often neglected. A distance-dependent dielectric model, while useful for dampening electrostatic interactions during the refinement stage, does not take into account the effect of the discontinuity between the low dielectric solute and the high dielectric solvent. Moreover, the electrostatic interaction energy term is likely to be overestimated in force field calculations as it is not balanced by the desolvation penalty associated with the removal of the interacting polar groups from water. These binding free energy components, as well as more elusive entropic effects, are seldom included on the assumption that they are of similar magnitude within the series^{6,10} and/or they randomly cancel each other due to enthalpy-entropy compensations.³ If our goal is to compute free energy differences correctly, it is of interest to distinguish between these two alternatives, and the trypsin-flavonoid complexes analyzed here are particularly well suited for sorting this out, as they fulfill a number of criteria that are usually deemed to be advantageous:⁵⁰ the inhibitors belong to a family of closely related compounds, no major conformational changes are expected to take place in the protein upon complex formation, and no side chain motions are frozen in the protein binding site relative to the unliganded state. The results of our calculations suggest that, within this congeneric series, the cost in electrostatic energy of desolvating both the enzyme binding site and the buried part of the inhibitors (ΔG_{desolv}) is roughly constant within the series (Table 2). On the other hand,

the third component of the electrostatic binding free energy (Figure 1), the electrostatic interaction energy ($E_{\text{ele}}^{\text{LR}}$), varies only slightly along the series in comparison with the van der Waals interaction, and this variation is mostly solvent-independent (Table 3). As a result, the differences in free energy of binding are dominated by the van der Waals term and the electrostatic contribution to the binding free energy differences is, to a good approximation, solvent-independent. This may account for the good correlation between calculated interaction energies and inhibitory activities initially found when only the simple molecular mechanics approach was considered (Figure 4).

The establishment of a good hydrogen bond between the NH of Gly219 and the 3'-hydroxyl group on the phenyl ring of quercetin, myricetin, and luteolin (Table 1) very likely involves the restriction of one internal rotation of the ligand, which is entropically unfavorable. Neglect of this entropic contribution in the present study does not appear to lead to large inaccuracies, as the agreement between the calculated and experimental free energy differences is within the error inherent to this type of experiments³ and comparable to that obtained using the more computationally demanding free energy perturbation calculations.⁵ At all events, we note that the average difference between calculated and experimental $\Delta\Delta G$ values comes to about 0.8 kcal mol⁻¹, which is roughly the energetic cost of fixing a torsional degree of freedom at 300 K.³ This means that introduction of this entropic penalty into the calculated ΔG values for quercetin, myricetin, and luteolin would provide even better agreement with the experimental values.

Dissection of the electrostatic binding free energy into components allowed us to examine the variation in the contribution of the different terms along the series (Table 2). All in all, the results presented clarify the energetics of ligand binding for these flavonoid inhibitors, and may help explain the success of other molecular mechanics-based QSAR studies on different ligand-receptor complexes. That is, our results would suggest that the good correlations obtained by different authors for congeneric series using only force field calculations⁶⁻⁹ could be accounted for, at least in part, by the fact that solvation effects tend to be of similar magnitude within the series. In the particular case studied in this paper, the interaction of a series of flavonoids with trypsin, van der Waals interactions appear to be the dominant energy contribution for the correct ranking of the inhibitors, even though the electrostatic term (found to be largely solvent-independent) follows a similar trend and may dominate the driving of the ligands and their binding orientations.

Finally, it is interesting to note the lack of correspondence between the electrostatic free energies of solvation (Table 4) and the electrostatic free energies of desolvation upon binding (Table 2). This arises from the fact that only a portion of each inhibitor is actually buried in the receptor upon binding. This observation merits some comment as 3D-QSAR-based methods (e.g. comparative molecular field analysis) that try to incorporate solvent effects by introducing the total solvation free energy of the ligands as an additional descriptor are likely, at least in cases with binding modes similar to the one described here, to introduce noise into the

energy matrix. In one example,⁵¹ it was actually found that whereas this contribution loaded heavily into the CoMFA model, it did not result in increased predictive ability.

Conclusions

Molecular models of the enzyme bovine trypsin complexed to a series of bioflavonoids that act as competitive inhibitors have been built. The deep pocket with Asp189 at its end accommodates the chromone ring, which interacts with the active site Ser195 through its γ -pyrone carbonyl, and the NH of Gly219 provides a hydrogen bond to the 3'-hydroxy on the phenyl ring of the best inhibitors. To our knowledge, this would be the first example where no charged ammonium or guanidinium groups are found interacting with Asp189 in the S1 pocket of trypsin, unlike the case of thrombin for which this is not a prerequisite for potent inhibition.⁵² Instead, a phenolic OH occupies the position of the water molecule that is found hydrogen bonded to Asp189 and Ser190 in free trypsin.

Using these complexes, a good correlation was found between the calculated molecular mechanics interaction energies and the experimental inhibitory potencies. To understand the origin of the observed correlation, the inhibitory activities were transformed into experimental free energies, and calculated free energy differences were approximated by considering the sum of the van der Waals term from the molecular mechanics force field and the electrostatic contribution computed using the continuum method. Partitioning of the electrostatic free energy contributions indicated similar desolvation penalties in the formation of all the complexes studied. The calculated and experimental free energy differences were in good agreement, although for one of the inhibitors it was necessary to take into account the difference in electrostatic free energy of solvation between bound and free states. A similar scenario in other systems may account for the good correlations frequently found between ligand activities and ligand-receptor interaction energies derived using plain molecular mechanics. Research is already in progress to assess the effect of introducing solvation/desolvation effects into more complex 3D-QSAR models derived using comparative binding energy (COMBINE) analysis.¹⁰

Acknowledgment. We thank Prof. Irwin Kurtz (Molecular Design Institute) and Dr. Peter J. Goodford (Molecular Discovery Ltd.) for provision of the DOCK and GRID programs, respectively. Biosym/Molecular Simulations, Inc. contributed a software license. We are also grateful to our colleagues Josefina Parellada and María C. Guinea for bringing these flavonoids to our attention. Financial support from the Spanish CICYT (Projects SAF94-630 and SAF96-231) and Comunidad Autónoma de Madrid (fellowship to A.R.O.) is gratefully acknowledged.

Supporting Information Available: AMBER parameters and partial atomic charges for the bioflavonoids (6 pages). Ordering information is given on any current masthead page.

References

- (1) Straatsma, T. P.; McCammon, J. A. Theoretical Calculations of Relative Affinities of Binding. *Methods Enzymol.* **1991**, *202*, 497-511.
- (2) Janin, J. Elusive Affinities. *Proteins* **1995**, *21*, 30-39.

- (3) Ajay; Murcko, M. A. Computational Methods to Predict Binding Free Energy in Ligand-Receptor Complexes. *J. Med. Chem.* **1995**, *38*, 4953-4967.
- (4) Böhm, H.-J.; Klebe, G. What Can We Learn from Molecular Recognition in Protein-Ligand Complexes for the Design of New Drugs? *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2588-2614.
- (5) (a) Straatsma, T. P.; McCammon, J. A. Computational Alchemy. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407-435. (b) Kollman, P. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev. (Washington, D.C.)* **1993**, *93*, 2395-2417.
- (6) Menziani, M. C.; De Benedetti, P. G.; Gago, F.; Richards, W. G. The Binding of Benzenesulfonamides to Carbonic Anhydrase Enzyme. A Molecular Mechanics Study and Quantitative Structure-Activity Relationships *J. Med. Chem.* **1989**, *32*, 951-956.
- (7) Grootenhuis, P. D. J.; van Galen, P. J. M. Correlation of Binding Affinities with Non-Bonded Interaction Energies of Thrombin-Inhibitor Complexes. *Acta Cryst.* **1995**, *D51*, 560-566.
- (8) Holloway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M. D.; Vacca, J. P.; Dorsey, B. D.; Levin, R. B.; Thompson, W. J.; Chen, L. J.; deSolms, S. J.; Gaffin, N.; Ghosh, A. K.; Giuliani, E. A.; Graham, S. L.; Guare, J. P.; Hungate, R. W.; Lyle, T. A.; Sanders, W. M.; Tucker, T. J.; Wiggins, M.; Wiscount, C. M.; Woltersdorf, O. W.; Young, S. D.; Darke, P. L.; Zugay, J. A. A Priori Prediction of Activity for HIV-1 Protease Inhibitors Employing Energy Minimization in the Active Site. *J. Med. Chem.* **1995**, *38*, 305-317.
- (9) Carson, M.; Yang, Z.; Babu, Y. S.; Montgomery, J. A. Calculation of Relative Binding Affinities of Purine Nucleoside Phosphorylase Inhibitors. *Acta Crystallogr.* **1995**, *D51*, 536-540.
- (10) Ortiz, A. R.; Pisabarro, M. T.; Gago, F.; Wade, R. C. Prediction of Drug Binding Affinities by Comparative Binding Energy Analysis. *J. Med. Chem.* **1995**, *38*, 2681-2691.
- (11) For a review, see: *3D-QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM Science Publishers B.V.: Leiden, 1993.
- (12) Chervenak, M. C.; Toone, E. J. A Direct Measure of the Contribution of Solvent Reorganization to the Enthalpy of Ligand Binding. *J. Am. Chem. Soc.* **1994**, *116*, 10533-10539.
- (13) Lybrand, T. P. Ligand-Protein Docking and Rational Drug Design. *Curr. Opin. Struct. Biol.* **1995**, *5*, 224-228.
- (14) Kraut, J. Serine Proteases: Structure and Mechanism of Catalysis. *Annu. Rev. Biochem.* **1977**, *46*, 331-358.
- (15) (a) Bode, W.; Schwager, P. Refined Crystal Structure of Bovine Beta-Trypsin at 1.8 Å Resolution. II. Crystallographic Refinement, Calcium Binding Site, Benzamidine Binding Site and Active Site at pH 7.0. *J. Mol. Biol.* **1975**, *98*, 693-717. (b) Walter, J.; Bode, W. The X-ray Crystal Structure Analysis of the Refined Complex formed by Bovine Trypsin and *p*-Amidinophenylpyruvate at 1.4 Å Resolution. *Hoppe-Seyler's Physiol. Chem.* **1983**, *364*, 949-959. (c) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. The Geometry of the Reactive Site and the Peptide Groups in Trypsin, Trypsinogen and Its Complexes with Inhibitors. *Acta Crystallogr.* **1983**, *B39*, 480-490.
- (16) Kurinov, I. V.; Harrison, R. W. Prediction of New Serine Proteinase Inhibitors. *Nature Struct. Biol.* **1994**, *1*, 735-743.
- (17) Parellada, J.; Guinea, M. Search for New Non-Peptidic Serine Proteinase Inhibitors of Plant Origin. *Pharm. Pharmacol. Lett.* **1995**, *2*, 66-69.
- (18) Parellada, J.; Guinea, M. Flavonoid Inhibitors of Trypsin and Leucine Aminopeptidase: A Proposed Mathematical Model for IC₅₀ Estimation. *J. Nat. Prod.* **1995**, *58*, 823-829.
- (19) A preliminary account of this work was presented at the *York Meeting on Molecular Interactions*, Molecular Graphics Society: York University, 1996.
- (20) Thaisrivongs, S.; Romero, D. L.; Tommasi, R. A.; Janakiraman, M. N.; Strohbach, J. W.; Turner, S. R.; Biles, C.; Morge, R. R.; Johnson, P. D.; Aristoff, P. A.; Tomich, P. K.; Lynn, J. C.; Horng, M.-M.; Chong, K.-T.; Hinshaw, R. R.; Howe, W. J.; Finzel, B. C.; Watenpugh, K. D. Structure-Based Design of HIV Protease Inhibitors: 5,6-Dihydro-4-hydroxy-2-pyrones as Effective, Non-peptidic Inhibitors. *J. Med. Chem.* **1996**, *39*, 4630-4642.
- (21) (a) Ballesteros, J. F.; Sanz, M. J.; Ubeda, A.; Miranda, M. A.; Iborra, S.; Payá, M.; Alcaraz, M. J. Synthesis and Pharmacological Evaluation of 2'-Hydroxychalcones and Flavones as Inhibitors of Inflammatory Mediators Generation. *J. Med. Chem.* **1995**, *38*, 2794-2797. (b) Ares, J. J.; Outt, P. E.; Randall, J. L.; Murray, P. D.; Weisshaar, P. S.; O'Brien, L. M.; Ems, B. L.; Kakodkar, S. V.; Kelm, G. R.; Kershaw, W. C.; Werchowski, K. M.; Parkinson, A. Synthesis and Biological Evaluation of Substituted Flavones as Gastroprotective Agents. *J. Med. Chem.* **1995**, *38*, 4937-4943. (c) Brinkworth, R. I.; Stoermer, M. J.; Fairlie, D. P. Flavones are Inhibitors of HIV-1 Proteinase. *Biochem. Biophys. Res. Commun.* **1992**, *188*, 631-637. (d) Wei, Y.; Zhao, X.; Kariya, Y.; Fukata, H.; Teshigawara, K.; Uchida, A. Induction of Apoptosis by Quercetin: Involvement of Heat Shock Protein. *Cancer Res.* **1994**, *54*, 4952-4957.
- (22) Rossi, M.; Rickles, L. F.; Halpin, W. A. The Crystal and Molecular Structure of Quercetin: A Biologically Active and Naturally Occurring Flavonoid. *Bioorg. Chem.* **1986**, *14*, 55-69.
- (23) Shin, W.; Lah, M. S. Structure of (R, S)-Naringenin. *Acta Crystallogr., Sect. C (Cryst. Struct. Commun.)* **1986**, *C42*, 626-628.
- (24) Lotter, H.; Wagner, H. Stereochemistry of Silybin. *Z. Naturforsch., C: Biosci.* **1983**, *38*, 339-341.
- (25) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. The Cambridge Crystallographic Data Centre: Computer-Based Search, Retrieval, Analysis and Display of Information. *Acta Crystallogr.* **1979**, *B35*, 2331-2339.
- (26) Insight II, release 95.0 (1995), Molecular Simulations Inc., 9685 Scranton Road, San Diego, CA 92121-3752.
- (27) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902-3909.
- (28) SPARTAN, version 3.0.1 (1993), Wavefunction, Inc., 18401 von Karman Avenue, Irvine, CA 92715.
- (29) Chirlian, L. E.; Francl, M. M. Atomic Charges Derived from Electrostatic Potentials: A Detailed Study. *J. Comput. Chem.* **1987**, *8*, 894-905.
- (30) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230-252.
- (31) Hopfinger, A. J.; Pearlstein, R. A. Molecular Mechanics Force-Field Parametrization Procedures. *J. Comput. Chem.* **1984**, *5*, 486-499.
- (32) (a) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849-857. (b) GRIN, GRID, and GRAB, version 12.0, Molecular Discovery Ltd., 1994.
- (33) (a) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A Geometric Approach to Macromolecule-Ligand Interactions. *J. Mol. Biol.* **1982**, *161*, 269-288. (b) Shoichet, B. K.; Bodian, B.; Kuntz, I. D. Molecular Docking Using Shape Descriptors. *J. Comput. Chem.* **1992**, *13*, 380-397. (c) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated Docking with Grid-Force Evaluation. *J. Comput. Chem.* **1992**, *13*, 505-524. (d) DOCK, version 3.0, Molecular Design Institute, University of California, San Francisco, 1992.
- (34) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F. Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A Computer-Based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112*, 535-542.
- (35) Richards, F. M. Areas, Volumes, Packing and Protein Structure. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151-176.
- (36) Connolly, M. L. Analytical Molecular Surface Calculation. *J. Appl. Crystallogr.* **1983**, *16*, 548-558.
- (37) AMBER (UCSF): Assisted Model Building with Energy Refinement, version 4.0, 1991. Department of Pharmaceutical Chemistry, University of California, San Francisco.
- (38) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, C.; Profeta, S.; Weiner, P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.
- (39) (a) Sreenivasan, U.; Axelsen, P. H. Buried Waters in Homologous Serine Proteases. *Biochemistry* **1992**, *31*, 12785-12791. (b) Finer-Moore, J. S.; Kossiakoff, A. A.; Hurlley, J. H.; Earnest, T.; Stroud, R. M. Solvent Structure in Crystals of Trypsin Determined by X-Ray and Neutron Diffraction. *Proteins* **1992**, *12*, 203-222.
- (40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.
- (41) Honig, B.; Nicholls, A. Classical Electrostatics in Biology and Chemistry. *Science* **1995**, *268*, 1144-1149.
- (42) Resat, H.; Marrone, T. J.; McCammon, J. A. Enzyme-Inhibitor Association Thermodynamics: Explicit and Continuum Solvent Studies. *Biophys. J.* **1997**, *72*, 522-532.
- (43) Nicholls, A.; Honig, B. A Rapid Finite Difference Algorithm, Utilizing Successive Over-Relaxation to Solve the Poisson-Boltzmann Equation. *J. Comput. Chem.* **1991**, *12*, 435-445.
- (44) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the Electrostatic Potential of Molecules in Solution: Method and Error Assessment. *J. Comput. Chem.* **1987**, *9*, 327-335.
- (45) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. Focusing of Electric Fields in the Active Site of Cu-Zn Superoxide Dismutase: Effects of Ionic Strength and Amino acid Modification. *Proteins* **1986**, *1*, 47-59.
- (46) (a) Mohan, V.; Davis, M. E.; McCammon, J. A.; Pettitt, B. M. Continuum Model Calculations of Solvation Free Energies: Accurate Evaluation of Electrostatic Contributions. *J. Phys.*

- Chem.* **1992**, *96*, 6428–6431. (b) Ewing, T. J. A.; Lybrand, T. P. A Comparison of Perturbation Methods and Poisson–Boltzmann Electrostatics Calculations for Estimation of Relative Solvation Free Energies. *J. Phys. Chem.* **1994**, *98*, 1748–1752.
- (47) Cheng, Y.-C.; Prusoff, W. H. Relationship Between the Inhibition Constant (K_i) and the Concentration of Inhibitor which Causes 50 Percent Inhibition (I_{50}) of an Enzymatic Reaction. *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
- (48) (a) Gilson, M. K.; Honig, B. Calculation of the Total Electrostatic Energy of a Macromolecular System: Solvation Energies, Binding Energies, and Conformational Analysis. *Proteins* **1988**, *4*, 7–18. (b) Jackson, R. M.; Sternberg, M. J. E. A Continuum Model for Protein-Protein Interactions: Application to the Docking Problem. *J. Mol. Biol.* **1995**, *250*, 258–275.
- (49) Glusker, J. P.; Rossi, M. Molecular Aspects of Chemical Carcinogens and Bioflavonoids. In *Plant Flavonoids in Biology and Medicine: Biochemical, Pharmacological, and Structure–Activity Relationships*; A. R. Liss, Inc.: New York, 1986; pp. 395–410.
- (50) (a) Kuntz, I. D.; Meng, E. C.; Shoichet, B. K. Structure-Based Molecular Design. *Acc. Chem. Res.* **1994**, *27*, 117–123. (b) Verlinde, C. L. M. J.; Hol, W. G. J. Structure-Based Drug Design: Progress, Results and Challenges. *Structure* **1994**, *2*, 577–587.
- (51) Waller, C. L.; Marshall, G. R. Three-Dimensional Quantitative Structure–Activity Relationship of Angiotensin-Converting Enzyme and Thermolysin Inhibitors. II. A Comparison of CoMFA Models Incorporating Molecular Orbital Fields and Desolvation Free Energies Based on Active-Analogue and Complementary-Receptor-Field Alignment Rules. *J. Med. Chem.* **1993**, *36*, 2390–2403.
- (52) Tapparelli, C.; Metternich, R.; Ehrhardt, C.; Cook, N. S. Synthetic Low-Molecular Weight Thrombin Inhibitors: Molecular Design and Pharmacological Profile. *Trends Pharmacol. Sci.* **1993**, *14*, 366–376.

JM970245V