A Possible Involvement of Solvent-Induced Interactions in Drug Design

Hongwu Wang[†] and Arieh Ben-Naim*

Department of Physical Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Received June 12, 1995[⊗]

We propose to study a new factor in designing new drugs. Most approaches to the drug design problem focus on the *direct* interactions between the drug and the corresponding target. We propose to study specific solvent-induced effects that can contribute to the binding Gibbs energy between the drug and its target. We estimate that these *indirect* effects will contribute significantly to the binding affinity and hopefully improve the clinical efficiency of the drugs.

1. Introduction

Recently, an intense effort has been directed at developing rational methods for designing new drugs.[1−5] The basic idea underlying the modern rational methods is to improve upon the binding affinity between a drug and its target. The improvement is expected to be achieved by first examining the *binding domain* between the drug and the target and then suggesting ways of modifying the strength of the interaction energy—hence the binding affinity—between the two binding partners.

Most of the effort has been focused on improving the *direct interaction* between the drug and the corresponding site on the target. These interactions are sometimes classified into two types: van der Waals or dispersion forces on one hand, and complementary functional groups on the other hand. These two types are shown schematically in Figure 1. The first is sometimes referred to as the lock and key model, since it involves "geometrical fit" between the drug and the target.

In both cases the focus is on the binding domain where all the interacting groups reside. By modifying some of these groups one can achieve an improved binding energy, hence binding affinity and hopefully also clinical effects.

In this article we propose a novel approach to the structure-based drug design which we hope can improve on the presently existing paradigm. The main idea is to search for specific solvent-induced effects that can potentially improve the binding affinity between the drug and the target. It is true that some solvent effects have been considered in the past, in the context of "hydrophobic (HΦO) interactions", but these, as the direct interactions, involve groups in the binding domain. It should be realized that HΦO interactions is only one of the many possible solvent-induced effects.[6−11] We shall describe these HΦO interactions in the context of all possible solvent-induced effects in the next section.

The main novelty of the approach described in this article is to add a new factor that can contribute to the binding affinity of the drug. This factor does not depend on the functional groups (FG) in the binding domain, but depends on the specific interactions between water molecules and hydrophilic (HΦI) groups on both the drug and the target. Figure 2 shows one such solvent-induced effect. Two HΦI groups, one on each of the binding partners, can be bridged by a water molecule forming two simultaneous hydrogen bonds. We stress that these two HΦI groups (a carbonyl and a hydroxyl in Figure 2) do not belong to the binding domain. This means that direct interactions between these two groups are negligible and therefore these FG's do not contribute to the *direct* binding affinity. It is only in an aqueous solvent that the *indirect* interactions, by means of water bridges, are expected to be operative. Recently we have examined various solvent effects on protein−protein binding and protein folding.[6−11] We found that there are several HΦI effects that can contribute significantly to the driving forces for these processes. We believe that these forces can also be used to enhance the binding of drugs to their targets.

In the following sections we shall describe how to use this additional "degree of freedom" to suggest modifications in the structure of the drug. These modifications would not affect the direct interaction, but may well add a significant contribution to enhance the overall binding affinity. We also hope that the suggested modifications will improve the clinical effects of the drugs.

In the next section we describe the theoretical framework within which we search for possible specific solvent effects on the binding Gibbs energy. We shall focus on one of these solvent-induced effects referred to as HΦI interaction, which we believe is the most important. In the following section we present several examples where such modification in the HΦI interaction can be implemented to achieve an enhanced binding affinity between the drug and the target.

2. Identifying Specific Solvent-Induced Effects

Consider the general binding of a ligand L (i.e. the drug) to a specific site on a protein P (the corresponding target). The reaction is

$$\mathrm{P} + \mathrm{L} \rightarrow \mathrm{PL} \tag{2.1}$$

The corresponding free energy change is

$$\Delta G = \mu_{\mathrm{PL}} - \mu_{\mathrm{P}} - \mu_{\mathrm{L}} \tag{2.2}$$

For simplicity we assume the followings:

(1) No conformational changes of either L or P occurs upon binding.

(2) One mode of binding, i.e. binding on one specific site.

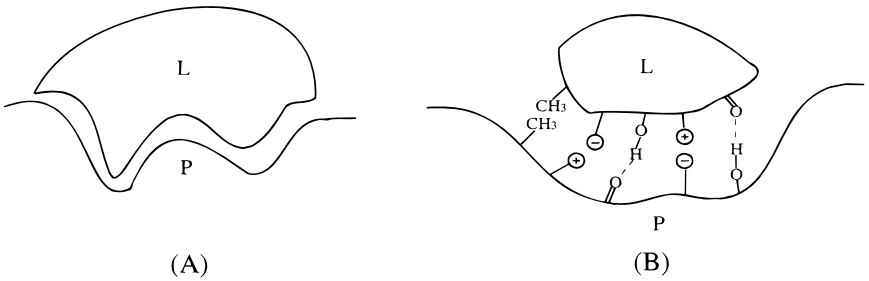
If the binding occurs in vacuum, i.e. in the absence of a solvent, then

* To whom all correspondence should be addressed.

[†] Present address: Research Institute, FF1, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195.

[⊗] Abstract published in *Advance ACS Abstracts,* March 1, 1996.

0022-2623/96/1839-1531$12.00/0



(A) (B)

Figure 1. The "lock and key" model for binding a ligand L to a protein P. In A the geometrical fit is equivalent to a maximum van der Waals interaction. In B the maximum interaction is achieved by a complementary pattern of functional groups on the surfaces of L and P.

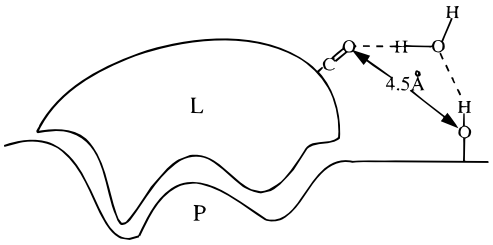


Figure 2. Same as in Figure 1, but now we added a new factor, a water molecule bridging between two functional groups. This originates from the presence of the solvent. Note that the two functional groups do not belong to the binding domain of PL and therefore do not contribute to the binding energy between P and L.

$$\Delta G^{g} = U_{PL} + kT \ln \frac{\rho_{PL} q_{PL}^{-1} \Lambda_{PL}^{3}}{(\rho_P q_P^{-1} \Lambda_P^3)(\rho_L q_L^{-1} \Lambda_L^3)} \quad (2.3)$$

where Λ_α^3 is the momentum partition function (PF), q_α is the internal PF of the species α, ρ_α is the number density of α, k is the Boltzmann constant, and T is the absolute temperature. U_{PL} is the interaction energy between P and L at the specific site of binding.

In a conventional thermodynamic treatment one usually defines some standard state (say 1 M solution for each species) and considers the standard Gibbs energy of the process.[12] Instead we write the chemical potential (CP) of each species as

$$\mu_\alpha = \mu_\alpha^* + kT \ln \rho_\alpha \Lambda_\alpha^3 \quad (2.4)$$

The quantity μ_α^* is referred to as the pseudochemical potential of the species α.[12] Eliminating the term $kT \ln \rho_\alpha \Lambda_\alpha^3$ is equivalent to freezing the translation degrees of freedom of α.

The pseudochemical potential may be interpreted as the Gibbs energy change for placing α at a fixed position in the liquid.[12]

$$\mu_\alpha^* = \mu_\alpha - kT \ln \rho_\alpha \Lambda_\alpha^3 \quad (2.5)$$

In a liquid, when internal degrees of freedom are separable, we have $\mu_\alpha^* = W(\alpha|l) - kT \ln q_\alpha$, where $W(\alpha|l)$ is the coupling work of α to the entire liquid l. When such separate cannot be assumed, one needs to define μ_α^* for each conformer, and then take an average over all possible conformations.[10]

We now focus on the process (2.1), but instead of freely moving particles we perform the process of binding P and L from fixed positions at infinite separation to the final configuration of PL. Thus we have, instead of (2.3)

$$\Delta G^{g} = U_{PL} + kT \ln\left(\frac{q_L q_P}{q_{PL}}\right) \quad (2.6)$$

We next perform the same process in a solvent and further assume that

(3) The internal PF of the species are not affected by the presence of the solvent, i.e. all q_α are invariant under the addition of the solvent to the system.

(4) That apart from the direct interaction energy, the internal PF's of P and L are preserved in the binding process,[13] i.e.

$$q_{PL} = q_P q_L \quad (2.7)$$

The Gibbs energy change for the same process is thus

$$\Delta G^{l} = U_{PL} + \delta G \quad (2.8)$$

where U_{PL} is referred to as the direct interaction energy between P and L at PL, and δG is the solvent effect on the Gibbs energy. (Here we exclude any solvent effects on the internal degrees of freedom of P and L.) The solvent contribution to the Gibbs energy change is defined by[10]

$$\delta G = \Delta G^{l} - \Delta G^{g} = \Delta\mu_{PL}^* - \Delta\mu_P^* - \Delta\mu_L^* \quad (2.9)$$

where $\Delta\mu_\alpha^*$ is the solvation Gibbs energy of the species α.

In the drug design problem most workers have focused on the direct interaction energy U_{PL}. This quantity is built up of the contribution of all the groups on L and P that are within the range of intermolecular forces. In some cases U_{PL} can be written as a sum of all pair interactions

$$U_{PL} = \sum U_{ij} \quad (2.10)$$

where i and j run over all the FG's on the surfaces of P and L, respectively. Modifications in the drug has been proposed in such a way as to change one of the quantities U_{ij}. This, in turn, would change the binding constant.

The binding constant, as measured experimentally through

$$K_{PL}^{l} = \left[\frac{\rho_{PL}}{\rho_P \rho_L}\right]_{eq} = e^{-\beta \Delta G_{PL}^{0l}} \quad (2.11)$$

where $\beta = (kT)^{-1}$ and ΔG_{PL}^{0l} is the standard Gibbs energy of the binding reaction, given by

$$\Delta G^{0l}_{PL} = \mu^{0l}_{PL} - \mu^{0l}_{P} - \mu^{0l}_{L} = \mu^{*l}_{PL} - \mu^{*l}_{P} - \mu^{*l}_{L} + kT \ln \frac{\Lambda^3_{PL}}{\Lambda^3_P \Lambda^3_L} \quad (2.12)$$

similarly, in the gaseous phase

$$K^{g}_{PL} = e^{-\beta \Delta G^{0g}_{PL}} \quad (2.13)$$

where

$$\Delta G^{0g}_{PL} = \mu^{0g}_{PL} - \mu^{0g}_{P} - \mu^{0g}_{L} = \mu^{*g}_{PL} - \mu^{*g}_{P} - \mu^{*g}_{L} + kT \ln \frac{\Lambda^3_{PL}}{\Lambda^3_P \Lambda^3_L} \quad (2.14)$$

(Classically, Λ^3_α is unchanged in the presence of the solvent.)

The ratio of the two binding constants is thus

$$\eta = \frac{K^{l}_{PL}}{K^{g}_{PL}} = e^{-\beta(\Delta G^{0l}_{PL} - \Delta G^{0g}_{PL})} = e^{-\beta \delta G} \quad (2.15)$$

where

$$\delta G = (\mu^{*l}_{PL} - \mu^{*g}_{PL}) - (\mu^{*l}_{P} - \mu^{*g}_{P}) - (\mu^{*l}_{L} - \mu^{*g}_{L}) = \Delta\mu^{*}_{PL} - \Delta\mu^{*}_{P} - \Delta\mu^{*}_{L} \quad (2.16)$$

Thus the ratio of the two experimental binding constants η is related to δG or to the difference in the solvation Gibbs energies between the final and the initial states.

Referring again to eq 2.8, where U_{PL} is the total binding *energy* of P and L at the specific binding site. We note that U_{PL} depends on all the FG's that belong to the surfaces of P and L. On the other hand δG depends on the properties of the solvent and is not an additive sum of the "interactions" between groups on P and L as in (2.10). Note also that δG also depends on groups on P and L that are far apart so that no direct interaction between them exists, but they are close enough so that they can affect the quantity δG; one such an example is described below.

We shall now focus on δG only. The statistical mechanical expression for δG is[8−10]

$$\delta G = -kT \ln \frac{\langle e^{-\beta B_{PL}} \rangle_0}{\langle e^{-\beta B_P} \rangle_0 \langle e^{-\beta B_L} \rangle_0} \quad (2.17)$$

the quantities $\langle\ \rangle_0$ are averages over all possible configurations of the solvent molecules in the absence of the solute. As these quantities stand there exist no obvious way of rewriting these quantities as sums of free energy terms, each pertaining to a pair of groups on P and L, i.e. we cannot claim a pairwise additive scheme for δG as for U_{PL} in (2.12).

In the drug design problem one starts with a given P (the target) and make changes in L (the drug) so as to modify the binding energy of L to P. For instance group 1 in Figure 3 is in the binding domain; hence a change in group 1 will affect the binding energy U_{PL}. In contrast, change in groups such as 2 or 3 in Figure 3 will not affect U_{PL}.

We now pose the following question. Suppose we replace or modify one FG on L say from R to R′; how this will affect δG? Clearly $\langle e^{-\beta B_P} \rangle_0$ remains unchanged (since no change has occurred to P). Hence for such a modification we write

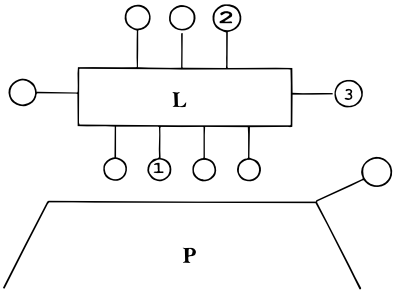


Figure 3. A schematic complex of a ligand L bound to protein P. Three representatives of functional groups are indicated. Group 1 in the binding domain (I), group 2 in the exterior (E) and group 3 in the joint (J) domain.

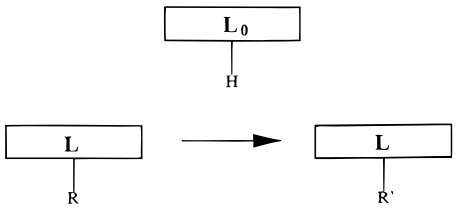


Figure 4. Ligand L with a functional group R being replaced by a group R′. The bare ligand, where the functional group is replaced by a hydrogen atom is indicated as L_0.

$$\delta\delta G = \delta G(R') - \delta G(R) = -kT \ln \frac{\langle e^{-\beta B'_{PL}} \rangle_0 \langle e^{-\beta B_L} \rangle_0}{\langle e^{-\beta B_{PL}} \rangle_0 \langle e^{-\beta B'_L} \rangle_0} \quad (2.18)$$

We now examine three possible changes in L. We make the distribution between three classes of FG's on L that are exposed to the solvent. These were referred to as classes E, I, and J (in refs 7−10), and the distinction between the FG's belonging to each class is according to the extent of the change in the solvation of the FG upon binding. In class E (E for external, e.g. group 2 of Figure 3), the solvation of a FG does not change upon binding. A FG in class I (I for internal, i.e. in binding domain, e.g. group 1 of Figure 3) loses completely its solvation, while in class J (J for joint, e.g. group 3 of Figure 3) the change in the solvation is only partial.

Clearly any change in a FG in E, say group 2 in Figure 3, will cause the same change in the two ratios in eq 2.18. Therefore modification in FG's belonging to E will have no effect on δG.[10]

Changes in FG's belong to I, e.g. group 1 in Figure 3, will affect U_{PL} as noted above. In addition it will also affect the ratio $\langle e^{-\beta B'_L} \rangle_0 / \langle e^{-\beta B_L} \rangle_0$ in (2.18). Note that since all the groups in I are not exposed to the solvent, the first ratio in (2.18) will not be affected by such a change. Solvent effects involving FG's in region I were discussed in literature in the context of the HΦO interactions. However as we shall soon demonstrate, there are other solvent-induced effects, involving HΦI groups that can contribute significantly to δG.

Suppose now we change a FG in the I region from R to R′. Let L_0 by L without the FG (Figure 4), then

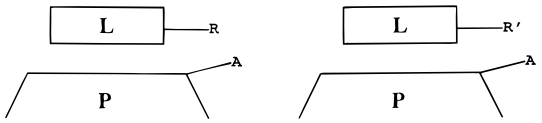


Figure 5. A replacement of a functional group R → R′ in the joint (J) domain.

$$\frac{\langle e^{-\beta B_L}\rangle_0}{\langle e^{-\beta B'_L}\rangle_0} = \frac{\langle e^{-\beta(B_{L0+R})}\rangle_0}{\langle e^{-\beta(B_{L0+R'})}\rangle_0} = \frac{\langle e^{-\beta B_{L0}}\rangle_0 \langle e^{-\beta B_R}\rangle_{BB}}{\langle e^{-\beta B_{L0}}\rangle_0 \langle e^{-\beta B_{R'}}\rangle_{BB}} = e^{-\beta(\Delta G^*_{R/BB} - \Delta G^*_{R'/BB})} \quad (2.19)$$

$\langle\ \rangle_{BB}$ is a conditional average, it is similar to $\langle\ \rangle_0$ in the sense that it is an average over all configurations of the solvent molecules. The only difference between $\langle\ \rangle_{BB}$ and $\langle\ \rangle_0$ is the use of conditional distribution of solvent configurations given the backbone.[10] $\Delta G^*_{\alpha/BB}$ means the conditional solvation Gibbs energy of α given the backbone BB.

If the change R → R′ has been the only change in L, then from (2.18) and (2.19) we have

$$\delta\delta G = -kT \ln \frac{\langle e^{-\beta B_R}\rangle}{\langle e^{-\beta B_{R'}}\rangle} = \Delta G^*_{R/BB} - \Delta G^*_{R'/BB} \quad (2.20)$$

Thus replacing R → R′ will cause a change in δG equivalent to the difference in the *conditional* Gibbs energies of solvation of R and R′. For example if R = CH_3 and R′ = OH, then we have the following approximate values[10,14]

$$\Delta G^*_{CH_3/BB} \approx +0.5$$

$$\Delta G^*_{OH/BB} \approx -6.0$$

$$\delta\delta G \approx +0.5 - (-6.0) = +6.5 \quad (2.21)$$

which means an unfavorable contribution to ΔG^l. Replacing R = OH by R′ = CH_3 will be favorable. This is a well-known effect. The solvent effect will prefer the HΦI groups to remain exposed to the solvent and the HΦO groups to be unexposed to the solvent. Any replacement that makes a HΦI group exposed or a HΦO group unexposed will therefore be favorable. Note that although the HΦO effect in this region contribute favorably, its absolute magnitude is small. On the other hand, the HΦI effect will in general contribute unfavorably to δG, but can contribute to the specificity of the binding, i.e. in the selection of the binding site on P.[10] Note that HΦO effect has been considered, for instance between group 1 in Figure 3 and a corresponding HΦO group on P. In our classification of solvent effects the HΦO group 1 will lose its solvation and hence provide negative contribution to ΔG^l. On the other hand if at 1 we have a HΦI group then we expect a relatively large loss of the solvation Gibbs energy, which will contribute a positive quantity to ΔG^l. Note that such a HΦI group might form one or more hydrogen bonds with groups on P, but this gain in the binding energy due to hydrogen bonds will be considered here as part of the *direct* interaction U_{PL} between L and P, and not a part of the solvent-induced effect.

There is an important intermediate effect due to FG's in class J, i.e. in the region where the FG's do not interact directly, but do interact indirectly through the solvent. Consider the following replacement R → R′ in L (Figure 5), where A is a fixed (unchanged) group on P and R is replaced by R′. We assume that there is no direct interaction between R (or R′) and A, so that in the absence of solvent, this replacement will not affect the free energy of binding. (If it does affect, then, by definition these FG's should belong to the class I, i.e. to the binding domain.)

For the replacement R → R′ (in Figure 5) we have eq 2.18 where now both ratios are affected by this replacement; thus we can write[10]

$$\begin{aligned}\delta\delta G &= \delta G' - \delta G \\ &= (\Delta G^*_{R',A/BB} - \Delta G^*_{R'/BB} - \Delta G^*_{A/BB}) - \\ &\qquad (\Delta G^*_{R,A/BB} - \Delta G^*_{R/BB} - \Delta G^*_{A/BB}) \\ &= COR(R',A) - COR(R,A) \end{aligned} \quad (2.22)$$

where we added and subtracted $\Delta G^*_{A/BB}$ to rewrite this expression as the difference in the (conditional) correlation free energies between R′, A and R, A. This difference depends on the type of groups R, R′, and A.

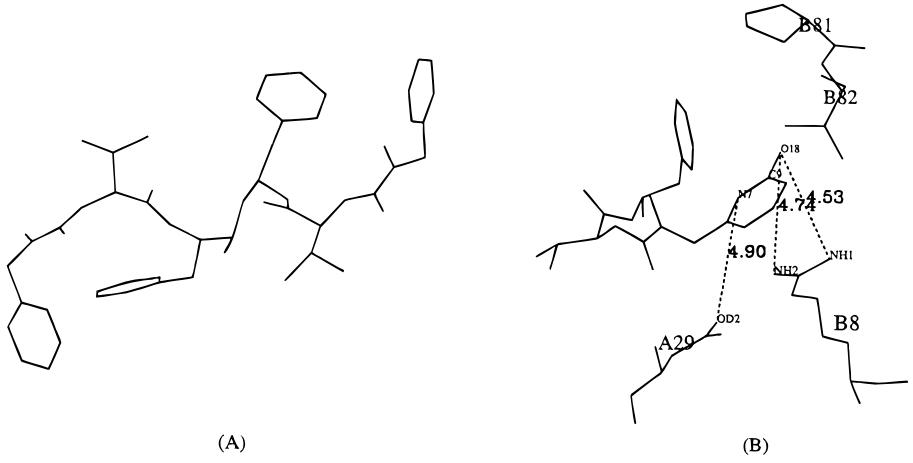


Figure 6. (A) Structure of pseudo C_2 symmetric HIV-1 protease inhibitor. (B) Structure of HIV-1 protease−inhibitor complex. Two pyridine rings are exposed to the solvent at two sides of the active site. Residues at one side of the active site around the pyridyl group are shown. Half of the inhibitor is shown with C_7 replaced by nitrogen and a hydrogen attached to C_9 replaced by a hydroxyl group. The distance between these groups and some hydrophilic groups on the surface of HIV-1 protease are indicated (dashed line).

Table 1. Distances (Å) between Atoms of HIV-1 Protease and Atoms on the Inhibitor[a]

	C_7[b]	N[c]	C_{11}	N	H_{C7}[d]	O[e]	N_f	H_{C9}	O	N	H_{C11}	O	N
N_{H1} (A8)	4.34	4.34	4.30	4.30	4.70	4.86	4.88				4.23	4.34	4.35
N_{H2} (A8)			4.06	4.06				4.25	4.43	4.46			
	$C_{7'}$	N	$C_{11'}$	N	$H_{C7'}$	O	N	$H_{C9'}$	O	N	$H_{C11'}$	O	N
N_{H1} (B8)			4.40	4.40	4.79	5.02	5.06	4.34	4.53	4.56	4.10	4.27	4.29
N_{H2} (B8)	4.19	4.19						4.52	4.74	4.74			
O_{D2} (A29)	4.90	4.90											

[a] When the atoms (groups) on the inhibitor are replaced by HΦI groups, water bridges may form between the inhibitor and the protein. [b] The atom numbering is the same as in the PDB files. [c] The previous carbon atom is replaced by a nitrogen atom. This meaning holds for the same symbol in this table and the following tables. [d] Hydrogen atom on C_7, the coordinates of hydrogen atoms in this table are presented in the PDB file. Hydrogen coordinates in the following tables are generated by molecular modeling software InsightII. [e] Replace the previous hydrogen atom by a hydroxyl group. The listed distance is between the oxygen atom of hydroxyl group and the corresponding atom on the protein. This meaning holds for the same symbol in this table and the following tables. [f] Replace the previous hydrogen atom by an amine group. The listed distance is between the nitrogen atom of amine group and the corresponding atom on the protein. This meaning holds for the same symbol in this table and the following tables.

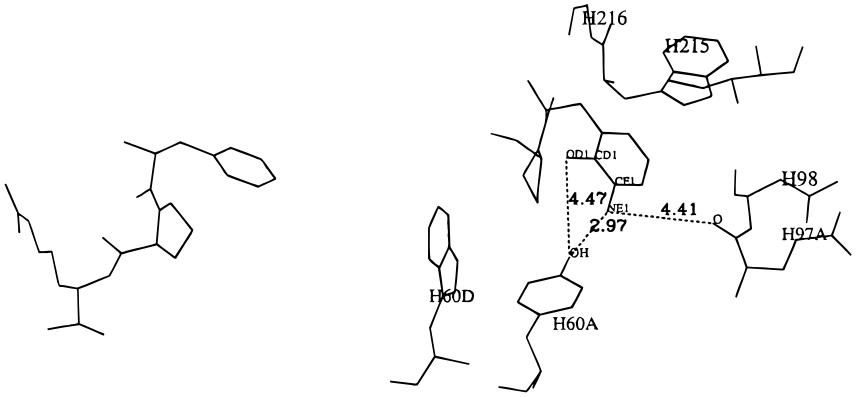


Figure 7. (A) Structure of thrombin inhibitor PPACK with chloromethyl ketone replaced by a methylene group. (B) Structure of thrombin−inhibitor complex. The phenyl ring of the inhibitor is exposed to the solvent. Residues of thrombin around this phenyl group are shown. Only a part of the inhibitor is shown for clarity. Hydrogens attached to C_{D1} and C_{E1} are replaced by hydroxyl and amine group, respectively. The distance between these groups and some hydrophilic groups on the surface of thrombin are indicated (dashed lines).

Table 2. Distances (Å) between Atoms of Thrombin and Atoms on the Inhibitor[a]

	C_{D1}	N	C_{E1}	N	H_{CD1}	O	N	H_{CE1}	O	N
O_H (H60A)	4.33	4.33			4.40	4.47	4.49	3.09	2.98[b]	2.97[b]
O (H97A)			4.40	4.40				4.36	4.40	4.41

[a] When the atoms (groups) on the inhibitor are replaced by HΦI groups, water bridges may form between the inhibitor and the protein. [b] A direct hydrogen bond may be formed here.

We now consider the following four cases:

If A is a hydrophobic group, then

(i) Replacing a hydrophilic R by a hydrophobic R′ will be *favorable*, and the magnitude of the effect is small, about −0.3 to −0.5 kcal/mol.[10]

(ii) Replacing a hydrophobic R by a hydrophilic R′ will be *unfavorable*.

If A is a hydrophilic group, then

(iii) Replacing a hydrophobic R by a hydrophilic R′ will be *very favorable*, if the distance and orientation are the optimal to form a water bridge between R′ and A. The magnitude of this effect was estimated to be large, about −2.5 to −3.0 kcal/mol.[10]

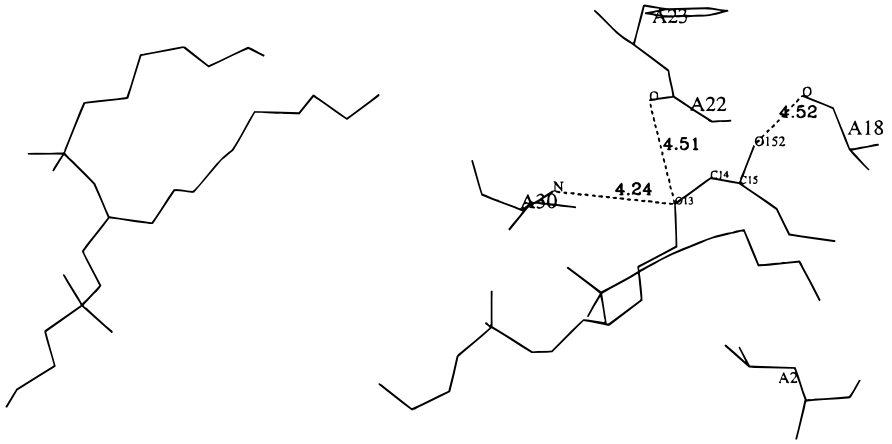
(iv) Replacing a hydrophilic R by a hydrophobic R′ will in general be *unfavorable* (presumably because of interference of the hydrophobic group with the solvation of the hydrophilic group at A).[10]

We see that out of the four possibilities considered above, only two effects lead to favorable contribution to δG. The most important one is the case iii. Therefore in the next section we shall try to search for candidates of ligands L where replacements of R → R′ in the J region will cause a favorable change, i.e. lowering the Gibbs energy of the binding, i.e. $\delta\delta G < 0$. We expect that the largest effect will occur when A is HΦI and the replacement R → R′ brings a HΦI group at R′ at the right distance and orientation to form a hydrogen-bonded bridge. The HΦO effect in the above sense is negative as expected, but the HΦI effect of this kind is also negative and possibly much larger than the corresponding HΦO effect. This is in sharp contrast to HΦO and HΦI effects that arise from the loss of the conditional solvation of such groups in the I region.

We have given above only one example of a solvent-induced HΦI effect between a pair of HΦI groups. We note that in principle higher order correlation, three and four HΦI groups might also contribute significantly to δG.[11] We shall not consider such correlation in this article.

3. Examples of Actual Systems Where Solvent Effects Is Expected To Increase the Binding Affinity of the Drugs

A. Protein-Binding Drugs. Here we select three examples of protein-binding drugs. After studying the structural features outside the binding domain, we suggest some simple modifications on the drugs. These modifications are designed to enable the formation of water bridges between a HΦI group on the drug and a



(A) (B)

Figure 8. (A) Structure of a transition-state analog inhibitor of PLA2. (B) Structure of PLA2−inhibitor complex. The terminal of 1-octyl is exposed to the solvent. Residues of PLA2 around 1-octyl group are shown. C_{13} of the inhibitor is replaced by oxygen, a hydrogen attached to C_{15} is replaced by a hydroxyl group. The distance between these groups and some hydrophilic groups on the surface of PLA2 are indicated (dashed lines).

Table 3. Distances (Å) between Atoms of PLA_2 and Atoms on the Inhibitor[a]

	C_{13}	O	N	C_{14}	O	N	H_{C18}	O	N
O (A18)				4.79	4.79	4.79	4.62	4.52	4.51
O (A22)	4.51	4.51	4.51	4.23	4.23	4.23			
N (A30)	4.24	4.24	4.24						

[a] When the atoms (groups) on the inhibitor are replaced by HΦI groups, water bridges may form between the inhibitor and the protein.

HΦI group on the target protein. These modifications will increase the binding affinity and possibly also the selectivity of the site on which the drug binds.

A.1. HIV-1 Protease−Inhibitor Complex.[15] The crystal structure of HIV-1 protease complex with pseudo C_2 symmetric inhibitor A78791 has been solved (PDB[16] file for this structure is 1hvj). Figure 6A shows the structure of the inhibitor. Parts of the two terminal pyridine rings of the inhibitor are exposed to the solvent. As the enzyme and inhibitor have the C_2 symmetry, we show only one side of the complex structure in Figure 6B (atom numbering is the same as in the corresponding PDB file). In this figure C_7 of the inhibitor is replaced by a nitrogen atom, and a hydroxyl group is added to C_9. The distance between the groups that can form water bridges in aqueous solution are indicated. Similar modification on the other terminal pyridine will have the same effect. Table 1 summarizes some of the possible modifications and geometry parameters. These are among the most straightforward ones; other possible modifications are left to further investigation.

As is pointed out in the literature,[17,18] most HIV-1 protease inhibitors suffer from poor solubility and oral bioavailability. By introducing such HΦI groups on the inhibitor, we expect not only an increase in affinity but also better solubility and oral bioavailability for HIV-1 protease inhibitors.

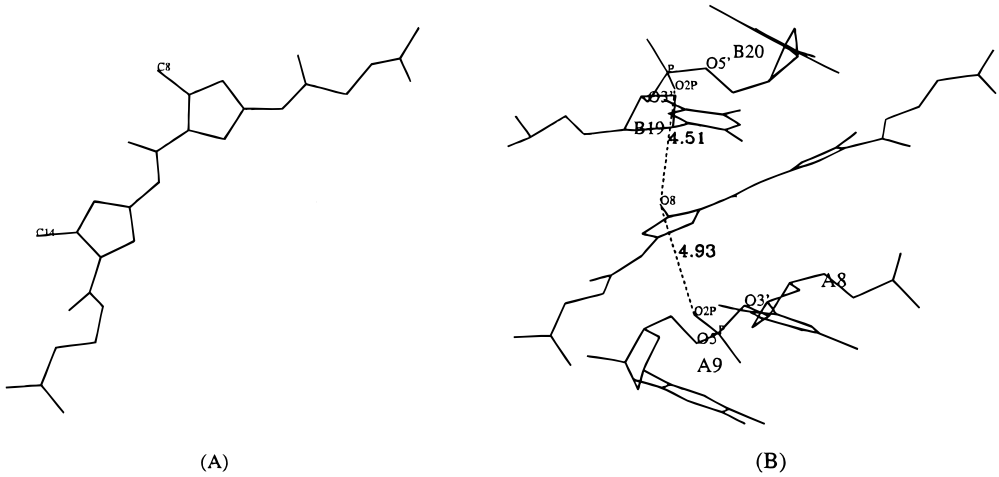
A.2. Thrombin−Inhibitor Complex. PPACK is an inhibitor of thrombin.[19] The crystal structure of PPACK with chloromethyl ketone replaced by a methyl group and thrombin complex has already been solved (PDB file for this structure is 1ppb).[20] The inhibitor binds to the active site of thrombin with part of its phenyl ring exposed to the solvent. Figure 7A shows the structure of the inhibitor. In Table 2 we indicate four positions for modifications in the inhibitor such that water bridges can form between the drug and the thrombin. Figure 7B shows one possible modification on the inhibitor. We added two HΦI groups to C_{D1} and C_{E1}. The distance between the hydroxyl O on C_{D1} and tyrosine hydroxyl group of residue H60A is 4.47 Å. The distance between amine N on C_{E1} and the backbone carbonyl O of residue H97A is 4.41 Å. Two water bridges may form after the suggested modification. Also in Figure 7B we see that there might be a direct hydrogen bond between the N on C_{E1} and the hydroxyl of H60A. Other possible modifications and the corresponding geometry parameters are indicated in Table 2.

A.3. Phospholipase A_2 (PLA2)−Inhibitor Complex.[21] L-1-Octyl-2-(heptylphosphonyl)-*sn*-glycero-3-phosphoethanolamine is a transition state analog inhibitor of PLA2 (Figure 8A). The PDB file for the crystal structure of PLA2 complex with this inhibitor is 1poe. From the crystal structure we can see that the 1-octyl chain of the inhibitor is mostly exposed to the solvent. This provides a number of possible modifications. For instance a change of C_{13} or a replacement of a hydrogen on C_{15} with a HΦI group are among the simplest that will not affect the structure of the inhibitor. We show one example of such a modification in Figure 8B.

In the modified inhibitor, C_{13} is replaced by oxygen, and a hydrogen on C_{15} is replaced by a hydroxyl group. These modifications enable the inhibitor to form three water bridges with the backbone nitrogen of residue A_{30} and the backbone oxygen of residue A_{18} and A_{22}. Table 3 provides a list of geometry parameters for possible modifications.

B. DNA Binding Drugs. In this section we provide two examples where replacement of a HΦO group on the drug by a HΦI group can enable water molecules to form hydrogen-bonded bridges between the drug molecule and the DNA. This is expected to increase the binding affinity of the drug to the DNA.

B.1. Netropsin−DNA Complex. Netropsin is a well-studied drug that binds to the minor groove of DNA.[22] The crystal structure of d(CGCAAATTTGCG) complexes with Netropsin has been determined (PDB file for the structure is 121d).[23] In the crystal structure the drug binds to the minor groove, selectivity to AT base pairs. The drug lies in the minor groove with two



(A) (B)

Figure 9. (A) Structure of DNA minor groove binding drug netropsin. (B) Structure of DNA dodecamer d(CGCAAATTTGCG) complex with netropsin. Only A_8, A_9, B_{19}, and B_{20} of DNA, which composed of the minor groove around methyl group C_8 of netropsin, are shown. In this figure C_8 is replaced by a hydroxyl group. The distances between the hydroxyl O and O_{2P} of A_8 and B_{20} are indicated (dashed lines).

Table 4. Distances (Å) between Atoms of DNA and Atoms on Netropsin[a]

	C_8	O	N	C_{14}	O	N	H_{C7}	O	N	H_{C13}	O	N
O_{2p}(A7)				4.84	4.84	4.84						
O_{2p}(A8)	4.93	4.93	4.93									
O_{2p}(B20)	4.50	4.51	4.50							5.16	5.01	4.99
$O_{3'}$(A8)	4.98	4.98	4.98									
$O_{3'}$(A9)							4.44	4.26	4.23			
$O_{3'}$(B18)							4.85	4.83	4.88			
$O_{3'}$(B19)	4.03	4.03	4.03									
$O_{3'}$(B20)										4.27	5.01	4.99

[a] When the atoms (groups) on the inhibitor are replaced by HΦI groups, water bridges may form between the inhibitor and the DNA.

methyl groups C_8 and C_{14} exposed to the solvent. In Figure 9A the structure of Netropsin is given, with C_8 and C_{14} indicated.

The distance between C_8 and O_{2P} of B_{20} is 4.50 Å, while the distance between C_8 and O_{2P} of A_8 on the other side of DNA minor groove is 4.93 Å. When we replace this methyl group by a hydroxyl group, the distances between O of hydroxyl and O_{2P} of A_8 and B_{20} are 4.93 and 4.51 Å, respectively (Figure 9B). In aqueous solution, this geometry is suitable for the formation of two water bridges simultaneously with the backbone of DNA. Replacing the methyl group numbered C_{14} by a hydroxyl group gives the distances between the hydroxyl O and O_{2P} of A_7 as 4.84 Å. We expect that one water bridge can be formed here between the hydroxyl group and DNA backbone, resulting in a further increase in binding free energy. Replacing the two methyl groups into amine NH_2 will have similar results. Further modifications such as addition of a HΦI group at C_7 or C_{13} (i.e. change the hydrogen atoms on these two carbon atoms into HΦI groups) can also provide suitable geometry for the formation of water bridges between DNA and the modified drug. Table 4 gives a list of geometry parameters before and after the modifications on Netropsin.

B.2. Hoechst 32528−DNA Complex. A similar effect can be obtained for Hoechst 32528, which is another DNA minor groove binding drug. The crystal structure for this drug and oligonucleotide d(CGCGAATTCGCG) complex has been solved (the PDB file for the structure is 127d).[24] The drug provides some possible positions for suitable modifications. For example we can add a HΦI group on C_{10} (change hydrogen on C_{10} into OH or NH_2). After the addition of a hydroxyl group at the position, the distances between the hydroxyl O and O_3' of B_{20} and O_{2P} of A_8 are 4.52 and 4.24 Å, respectively. Therefore, water bridges can be formed here between this HΦI groups on the drug and the DNA backbone. Figure 10 shows the structure of Hoechst 32528 and two of the possible modifications on the drug molecule. Table 5 gives a list of some possible modifications and the corresponding geometry parameters.

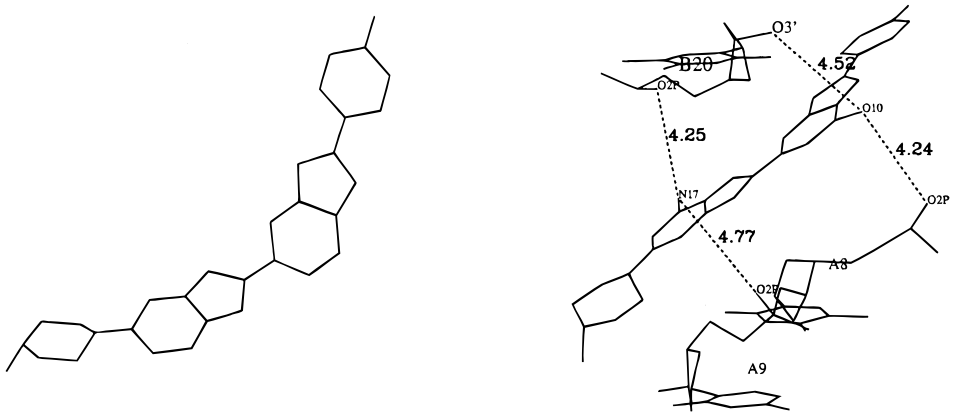
4. Discussion

In the previous section we have given a few examples where a modification in the drug can increase the binding Gibbs energy of a drug to the corresponding target.

We have focused on two types of changes, both in the region J, i.e. a region in which the FG's are sufficiently far apart so that direct interaction is negligible, but indirect or solvent-induced effect could be appreciable. The first type of change is to change a hydrogen atom by a HΦI group, in such a way that the newly HΦI group can form a hydrogen-bonded bridge with a HΦI on the target P. This change will cause a change in δG that was estimated to be of the order of −3 kcal/mol (for the most favorable configuration of the two HΦI groups). If this is the only change made to L, then the binding constant will change from K_{PL} to K'_{PL} such that (see eqs 2.15 and 2.22)

$$K'_{PL}/K_{PL} = \exp(-\beta\delta G' - \beta\delta G) \approx \exp(3.0/0.6) \approx e^5 = 148 \quad (4.1)$$

Here we assume that the pair H (on L) and HΦI group (on P) do not contribute to δG, but replacing the hydrogen atom by a HΦI group creates the possibility



(A) (B)

Figure 10. (A) Structure of DNA minor groove binding drug Hoechst 32528. (B) Structure of DNA dodecamer d(CGCGAATTCGCG) complex with Hoechst 32528. Only A_8, A_9, and B_{20} of DNA are shown. Hydrogen on C_{10} is replaced by a hydroxyl group. The distances between the hydroxyl O_{10} and O_{2P} of A_8, $O_{3'}$ of B_{20} are indicated (dashed lines). Hydrogen on C_{17} is replaced by an amine group. The distances between the nitrogen N_{17} and O_{2P} of A_9, O_{2P} of B_{20} are also indicated.

Table 5. Distances (Å) between Atoms of DNA and Atoms on Hoechst 32528[a]

	H_{C5}	O	N	H_{C6}	O	N	H_{C10}	O	N	H_{C11}	O	N	H_{C17}	O	N	H_{C18}	O	N
O_{2p}(A8)							4.31	4.24	4.23	4.90	4.92	4.92						
O_{2p}(A9)													4.77	4.77	4.77			
O_{2p}(B20)													4.32	4.26	4.25			
O_{2p}(B22)	4.25	4.15	4.14	4.99	4.44	4.43												
$O_{3'}$(A6)	4.51	4.59	4.60															
$O_{3'}$(A7)							4.40	4.52	4.54	4.74	4.90	4.93						
$O_{3'}$(B19)													3.98	4.10	4.11	4.58	4.75	4.78
$O_{3'}$(B20)							4.37	4.52	4.54									

[a] When the atoms (groups) on the inhibitor are replaced by HΦI groups, water bridges may form between the inhibitor and the DNA.

of forming a hydrogen-bonded bridge. Such a HΦI effect will cause an increase of almost 2 orders of magnitude in the binding constant. If two or three water bridges can form, we expect an increase of 4 and 6 orders of magnitude in the binding constant.

The second type of replacement is when a HΦO group such as methyl is replaced by a HΦI group, such as a hydroxyl. In this case we start with a HΦO−HΦI correlation and replace it by a HΦI−HΦI correlation. It has been argued recently[7,10,12,25] that such a replacement can contribute more than −3 kcal/mol to δG. Qualitatively, the reason is that a HΦO group close to the HΦI group on P can interfere with the solvation of the latter. By replacing the HΦO group by a HΦI group we not only provide a favorable condition for the formation of hydrogen-bonded bridges but also eliminate the interference in the solvation of the two original groups. Both of these effects will contribute negatively to δG. We do not have an estimate for the change in the binding constant, but we expect that such a change will increase K_{PL} by a factor larger than that given in eq 4.1 (per hydrogen-bonded bridge).

It should be noted that in both types of changes suggested above a HΦI group is introduced in the drug. This should lead to an increase in the solubility of the newly designed drug. In contrast previous suggestions to place a HΦO grop in the I region were designed to enhance the binding affinity, but at the same time must have caused a reduction in the solubility of the new drug. We hope that by increasing both the binding affinity and the solubility of the drug, as expected in the modifications suggested in this paper, we can improve in the overall clinical efficiency of the drugs.

References

(1) Kuntz, I. D. Structure-based strategies for drug design and discovery. *Science* **1992**, *257*, 1078.
(2) Perun, T. J.; Propst, C. L. *Computer-Aided Drug Design*; Marcel Dekker, Inc.: New York, 1989.
(3) Propst, C. L.; Perun, T. J. *Nucleic Acid Targeted Drug Design*; Marcel Dekker, Inc.: New York, 1992.
(4) Greer, J.; Erickson, J. W.; Baldwin, J. J.; Varney, M. D. Application of the three dimensional structures of protein target molecules in structure-based drug design. *J. Med. Chem.* **1994**, *37*, 1035.
(5) Roerdink, F. H.; Kroon, A. M. *Drug Carrier Systems*; John Wiley and Sons: New York, 1989.
(6) Ben-Naim, A. Solvent induced interactions, hydrophobic hydrophilic phenomena. *J. Chem. Phys.* **1989**, *90*, 7412.
(7) Ben-Naim, A. Solvent effect on protein association and protein folding. *Biopolymer* **1990**, *29*, 567.
(8) Ben-Naim, A.; Ting, K. L.; Jernigan, R. L. Solvation thermodynamics of biopolymers. *Biopolymers* **1990**, *28*, 1309.
(9) Ben-Naim, A.; Ting, K. L.; Jernigan, R. L. Solvent effects on binding thermodynamics of biopolymers. *Biopolymer* **1990**, *29*, 901.
(10) Ben-Naim, A. *Statistical Thermodynamics for Chemists and Biochemists*; Plenum Press: New York, 1992.
(11) Ben-Naim, A. Strong forces between hydrophobic macromolecules. *J. Chem. Phys.* **1991**, *93*, 8196.
(12) Ben-Naim, A. *Solvation Thermodynamics*; Plenum Press: New York, 1987.
(13) Usually one defines $q'_{PL} = e^{-\beta U_{PL}} q_P q_L$. Hence instead of (2.6), we would have written $\Delta G^g = kT \ln (q_L q_P / q'_{PL})$, which is the same as (2.6). Here q_{PL} does not include the direct interaction energy between P and L.
(14) Wilf, J.; Ben-Naim, A. Direct measurement of intramolecular hydrophobic interactions. *J. Phys. Chem.* **1994**, *98*, 8594.
(15) Hosur, M. V.; Bhat, T. N.; Kempf, D.; Baldwin, E. T.; Liu, B.; Gulnik, S.; Wideburg, N. E.; Norbeck, D. W.; Appelt, K.; Erickson, J. W. Influence of stereochemistry on activity and binding modes for C2 symmetry-based diol inhibitors of HIV-1 protease. *J. Am. Chem. Soc.* **1994**, *116*, 847.
(16) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodger, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The protein data bank, a computer-based archival file for macromolecular structures. *Eur. J. Biochem.* **1977**, *80*, 319.

(17) Dorsey, B. D.; Levin, R. B.; McDaniel, S. L.; Vacca, J. P.; Guare, J. P.; Darke, P. L.; Zugay, J. A.; Emini, E. A.; Schlief, W. A.; Quintero, J. C.; Lin, J. H.; Chen, I. W.; Holloway, M. K.; Fitzgerald, P. M. D.; Axel, M. G.; Ostovic, D.; Anderson, P. S.; Huff, J. R. L-735,524: The design of a potent and orally bioavailable HIV protease inhibitor. *J. Med. Chem.* **1994**, *37*, 3443.

(18) Kempf, D. J.; Marsh, K. C.; Fino, L. C.; Bryant, P.; Kennard, A. C.; Sham, H. L.; Zhao, C.; Vasavanonda, S.; Kohlbrenner, W. E.; Wideburg, N. E.; Saldivar, A.; Green, B. E.; Herrin, T.; Norbeck, D. W. Design of orally bioavailable symmetry-based inhibitors of HIV protease. *Bioorg. Med. Chem.* **1994**, *2*, 847.

(19) Tapparelli, C.; Metternich, R.; Ehrhardt, C.; Cook, N. S. Synthetic low-molecular weight thrombin inhibitors: molecular design and pharmacological profile. *Trends Pharm. Sci.* **1993**, *14*, 366.

(20) Bode, W.; Mayr, I.; Baumann, U.; Huber, R.; Stone, S. R.; Hofsteenge, J. The refined 1.9 angstroms crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg Chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.* **1989**, *8*, 3467.

(21) Scott, D. L.; White, S. P.; Browning, J. L.; Rosa, J. J.; Gelb, M. H.; Sigler, P. B. Structures of free and inhibited human secretory phospholipase A2 from inflammatory exudate. *Science* **1991**, *254*, 1007.

(22) Krugh, T. R. Drug-DNA interactions. *Curr. Opin. Struct. Biol.* **1994**, *4*, 868.

(23) Tabernero, L.; Verdaguer, N.; Coll, M.; Fita, I.; Marel, G. A.; Boom, J. H.; Rich, A.; Aymami, J. Molecular structure of the A-tract DNA dodecamer d(CGCAAATTTGCG) complexed with the minor groove binding drug Netropsin. *Biochemistry* **1993**, *32*, 8403.

(24) Sriram, M.; Marel, G. A.; Roelen, H.; Boom, J. H.; Wang, A. J. Conformation of B-DNA containing O6-ethyl-G-C base pairs Stabilized by minor groove binding drugs: Molecular structure of d(CGC[E6G]AATTCGCG) complexed with Hoechst 33258 or Hoechst 33342. *EMBO J.* **1992**, *11*, 225.

(25) Mezei, M.; Ben-Naim, A. Calculation of the solvent contribution of the potential of mean force between two water molecules. *J. Chem. Phys.* **1990**, *92*, 1359.

JM950430D