Calculation of Substituent Effects on pK_a Values for Pyrone and Dihydropyrone Inhibitors of HIV-1 Protease

Igor A. Topol,*,† Stanley K. Burt,† Alexander A. Rashin,‡ and John W. Erickson§

Advanced Biomedical Computing Center, Structural Biochemistry Program, SAIC Frederick, NCI Frederick Cancer Research and Development Center, P.O. Box B, Frederick, Maryland 21702-1201, and BioChemComp, Inc., 543 Sagamore Avenue, Teaneck, New Jersey 07666

Received: July 29, 1999; In Final Form: November 23, 1999

We have investigated the influence of different substitutions and solvent effects on the pK_a values of the hydroxy group for a set of pyrone and dihydropyrone HIV-1 protease inhibitors. Absolute and relative pK_a values were calculated for model compounds using a combination of density functional theory (DFT) and continuum solvation methods and were compared to experimental data for related compounds. The theoretical results shed light on the unusual pH dependence of the inhibition constants for these compounds and may lead to the design of new, improved pyrone and dihydropyrone inhibitors of HIV-1 protease.

Introduction

Ionization processes are of fundamental importance in many areas of chemistry and biochemistry. The determination of pK_a values is of special interest for studying the role of ionizable groups in enzymatic reaction mechanisms, in proton-transfer processes, and for understanding the binding of inhibitors to enzymes. Experimental determination of individual pK_a values can be quite complicated in complex systems such as proteins, or inhibitor-enzyme complexes. Quantum chemical methods can, in principle, provide a reliable and accurate means of calculating relative and/or absolute pK_a values. Moreover, these types of calculations should allow a better understanding of the different factors that influence pK_a values and are essential for interpretation of experimental pK_a values in proteins and other complex systems. However, until recently, such applications have been scarce.

A series of achiral pyran-2-one analogues that possess a 3-phenylthio group have been shown to be high-affinity inhibitors of HIV-1 protease.[1] The general structures of these pyrone and dihydropyrone compounds are shown in Figure 1. X-ray crystallographic studies have revealed the mode of binding of these nonpeptidic inhibitors in the active site of HIV-1 protease[1] (Figure 2). These crystal structures were useful in the design and optimization of inhibitors with high affinity for the enzyme. However, structure−activity relationships for these compounds revealed that the antiviral potency of many compounds was much lower than expected on the basis of enzyme inhibition.[2] For example, some inhibitors with K_i values in the nanomolar range at pH 4.5 exhibited no measurable antiviral potency at 100 μM.[2] Unlike most other classes of HIV protease inhibitors that possess an alcohol-containing transition state analogue, the pyrones contain an acidic enolic hydroxyl group. Thus, many of the substituted pyrones and dihydropyrones may exist in anionic form at physiological pH, whereas they would be protonated at acidic pH. X-ray crystallographic studies show

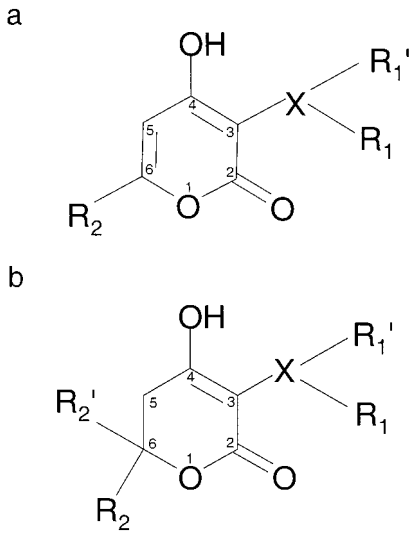


Figure 1. Chemical structures of (a) pyran-2-one and (b) dihydropyran-2-one inhibitors of HIV-1 protease.

that when the inhibitor is bound to the enzyme, the enolic hydroxyl group is within hydrogen bonding distance of the active site aspartates (Figure 2).[1] If inhibition of HIV protease requires the protonated enol form of the inhibitor, this could explain the poor antiviral potency of many of these inhibitors. In support of this idea, many of these compounds exhibit marked pH dependencies in their enzyme inhibition properties in contrast to other classes of HIV protease inhibitors.[2,3] Structure−activity studies showed that substituents at the 3- and 6-positions can influence the pK_a of the hydroxyl group[1−3] in a manner consistent with the hypothesis that the anionic form of the inhibitor is less potent (Table 1). In addition, negatively charged compounds are also known to have generally poorer cell permeability than their neutral analogues.

These observations prompted us to undertake a theoretical investigation to calculate the observed pK_a changes. An important goal of this study was to develop a theory-based method to aid in identifying for model pyrones and dihydropyrones substituents at the 3- and 6-positions that could enhance

† Advanced Biomedical Computing Center, NCI Frederick Cancer Research and Development Center.

‡ BioChemComp, Inc.

§ Structural Biochemistry Program, NCI Frederick Cancer Research and Development Center.

10.1021/jp992691v CCC: $19.00
Published on Web 01/04/2000

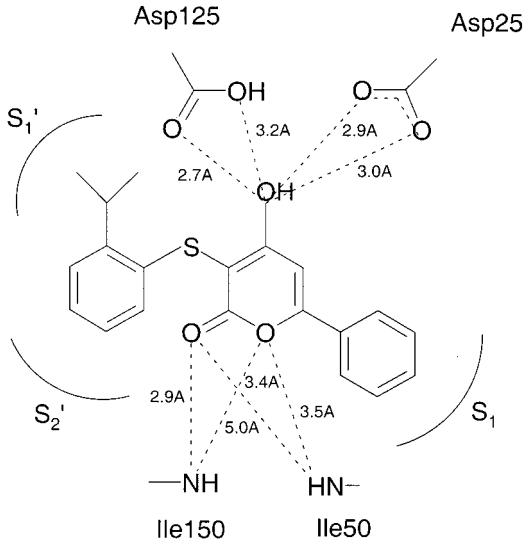
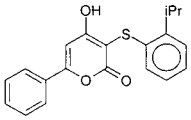
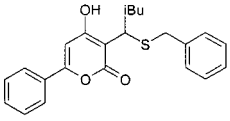
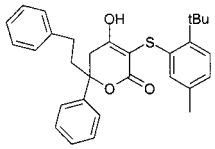


Figure 2. Schematic illustration of binding mode of an inhibitor PD 154906 in the active site of HIV-1 protease. Possible hydrogen bonds are drawn as dashed lines. S_1 and $S_{2'}$ represent enzyme subsites for binding functional groups on inhibitor as defined in ref 52. Based on ref 1.

TABLE 1: pH Dependence of Inhibitor Activity of Selected Pyrone and Dihydropyrone HIV-1 Protease Inhibitors[a]

	Enzyme inhibition IC_{50}, μM			Anti-HIV activity in cell culture IC_{50}, μM
	pH 4.7	pH 6.2	pK_a	
	0.037	0.64	4.1	> 69
	0.014	0.097	4.0	69
	0.26	0.30	6.0	28

[a] Data taken from ref 2.

the antiviral activity of these compounds by elevating the pK_a value of the enolic hydroxy group.

Theory

Theoretically, the absolute or relative pK_a values can be determined from the thermodynamic cycle shown in Figure 3. Such pK_a calculations require reliable and highly accurate gas-phase protonation/deprotonation energy calculations and solvation energy calculations for both products and reactants of the gas-phase proton abstraction reaction. Inaccuracies in the quantum mechanical calculations of proton affinities in the Hartree−Fock (HF) approximation using modest basis sets of level 6-31G(d) or 6-31G(d,p) can lead to errors in relative pK_a values by several pK_a units.[4−7] Inaccuracies in solvation energy calculations can lead to additional sources of error.

Recently, however, the possibilities for accurate pK_a calculations have improved. One area of improvement that is relevant to small molecules is in the ab initio quantum chemical methods,

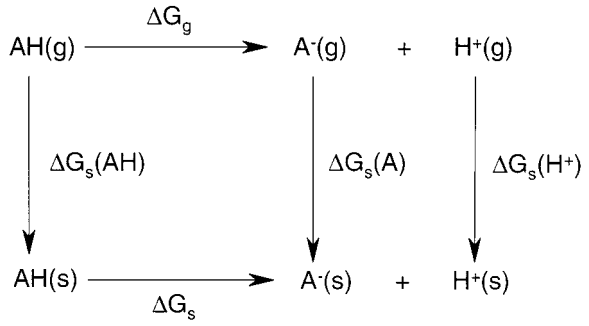


Figure 3. Thermodynamic cycle for proton abstraction reaction in the gas phase (g) and in solvent (s).

e.g., density functional[8−13] and G2[14−16] type theories. These methods consistently yield proton affinities and proton-transfer enthalpies within 1−4 kcal/mol of experimental values. Another area of improvement has been in the development of self-consistent reaction field procedures that combine ab initio quantum mechanics with dielectric continuum solvation theory.[17−28] These procedures can give a remarkably accurate representation of the properties of molecules in aqueous environments.

We have recently shown using a set of substituted imidazoles that both absolute and relative pK_a values for the deprotonation of nitrogen on the imidazole ring can be calculated with an average absolute deviation less than 0.8 units from experimental values.[29] This degree of accuracy is possible only if the solutes are treated at the correlated level using either G2 type or density functional theory (DFT). A similar accuracy of about 1 pK_a unit has been found by other authors in most calculated absolute pK_a values using combinations of DFT and continuum solvation theory.[30] Taking into account all information obtained from the calculations of pK_a values for small molecules, we can formulate some requirements for the computational tools necessary for pK_a calculations in much larger systems. The gas-phase part of the reaction in Figure 3 should be calculated at a post-HF level, including electron correlation effects and using extended basis sets. The solvation part of the calculation requires an accuracy of better than 1 kcal/mol for relative solvation energy values when compared with experimental data. Current DFT methods combined with a continuum dielectric solvation model fulfill these requirements in most cases and should allow the calculation of pK_a values for complex systems. Recently, the combination of these two methods was applied to the calculation of pK_a values for several small molecules[29,30] and was shown to be more accurate than most previous approaches used to calculate absolute pK_a values. Thus, while the approach of combining DFT and continuum solvation theory has been shown to yield larger errors in calculated relative to experimental values in a few particular cases, it works very well in most cases and can be expected to yield meaningful and insightful results with a high probability.

Computational Methods

Given the thermodynamic cycle shown in Figure 3, the pK_a value for compound AH with the ionizable group can be calculated from the following formula:

$$pK_a = 0.434(RT)^{-1}\{\Delta G_g + \Delta G_s(A^-) - \Delta G_s(AH) + \Delta G_s(H^+)\} \quad (1)$$

In Figure 3 and formula 1, ΔG_g and ΔG_s are free energies of ionization in the gas phase and solvent, respectively. $\Delta G_s(A^-)$,

$\Delta G_s(AH)$, and $\Delta G_s(H^+)$ are solvation free energies of anion A^-, protonated compound AH, and proton H^+, respectively.

The gas-phase free energy of ionization can be expressed in terms of the enthalpy and entropy contributions:

$$\Delta G_g = \Delta H_g - T\Delta S_g \quad (2)$$

The enthalpy of ionization in the gas phase, ΔH_g, is

$$\Delta H_g = E_{SCF}(A^-) - E_{SCF}(AH) + E_{vib}(A^-) - E_{vib}(AH) + \frac{5}{2}RT \quad (3)$$

where E_{SCF} represents the energy obtained from the SCF calculations, E_{vib} includes zero point energy and temperature corrections to the vibrational enthalpy, and the term (5/2)*RT* includes the translational energy of the proton and the $\Delta(PV)$ term. The entropy contribution is given by

$$-T\Delta S_g = -T[S(A^-) + S(H^+) - S(AH)] \quad (4)$$

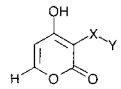
For $T = 298$ K, the second term $TS(H^+) = 7.76$ kcal/mol.[31] Thus,

$$\Delta G_g = \Delta H_g - T[S(A^-) - S(AH)] - 7.76 \quad (5)$$

Actually, the translational and rotational contributions to the ΔH_g should be included in expression 3; however, for the anionic and the neutral system in the ideal gas approximation, these contributions cancel each other. The translational enthalpy of the proton is included in the (5/2)*RT* term.

E_{SCF} and vibrational contributions to ΔH_g were calculated using the DFT program DGauss.[32] The geometries for the neutral (AH) and anionic (A^-) conformations were optimized at the NLSCF (nonlocal, gradient-corrected) level using DZVP, DZVP2,[33] and DZVPD[34] basis sets. The lowest energy conformations were used for all subsequent pK_a calculations. The DZVPD basis set in addition to the p polarization functions on all atoms includes diffuse d functions on heavy atoms. It has been found that including d functions on hydrogens has a negligible impact on the results but requires a significant increase in the computational time.[34] We have recently shown that the DFT/DZVPD calculations yield the best agreement between calculated and experimental vacuum dipole moments with an unsigned mean error of 0.07 D.[35] This is over 3 times more accurate than conventional HF calculations with a 6-31G(d) basis set and over 2 times more accurate than DFT calculations with DZVP and DZVP2 basis sets.[35] At the stage of geometry optimization as a first step in pK_a calculations the gradient-corrected Becke's exchange functional,[36] the Perdew correlational functional (BP86),[36] and DZVPD basis sets were used in the framework of the DGauss program package.[32] The hybrid three-parameter Becke−Perdew exchange correlation functionals (B3P86)[37,38] were used for more accurate calculations of the E_{SCF} energies using the 6-311+G(d,p) basis set (see Table 2) and Gaussian 94 program.[39] In the last case we investigated the basis set dependence on the energy difference between neutral and anionic states (electron affinity) by adding diffuse functions on the hydrogens (6-311++G(d,p)) to the 6-311+G(d,p) basis set and additional polarization functions on all atoms (6-311++G(2df,2p)) and using an extended correlation consistent aug-cc-pVTZ basis set. To study basis set and electron correlation dependencies, the E_{SCF} values were also calculated for some model compounds with complete optimizations in the Hartree−Fock (HF) and second-order Moller−Plesset (MP2) approximations using the Gaussian 94 program and 6-31G(d,p), 6-31+G(d,p), and 6-311+G(d,p) basis sets.

TABLE 2: Energy Differences [$E_{scf}(A^-) - E_{scf}(AH)$] (in kcal/mol) between Neutral and Anionic States for Various Model Pyrones Calculated Using Different Methods and Approximations



Method[a]	X=H	X=CH_2 Y=CH_3	X=S Y=CH_3	X=NH Y=CH_3	X=O Y=CH_3
BP86/DZVP//DZVP	331.7	332.1	333.8	338.1	337.6
BP86/DZVPD//DZVP	334.5	335.0	338.5	339.5	340.4
BP86/TZVP//DZVP	333.1	333.5	337.0	338.6	339.2
BP86/DZVPD//DZVPD	334.3	335.0	336.9	340.3	340.1
B3P86/6-311+G(d,p)[b]	334.3	335.0	337.0	340.4	340.1
B3P86/6-311++G(d,p)[b]	334.4	335.1	337.0	340.4	340.1
B3P86/6-311++G(2df,2p)[b]	335.8	336.5	338.6	340.8	341.2
B3P86/aug-cc-pVTZ[b]	336.4	337.1	339.1	341.5	341.6
HF/6-31G(d,p) //6-31G(d,p)	349.5	349.2	348.0	357.8	353.7
HF/6-31+G(d,p) //6-31+G(d,p)	340.8	341.3	341.2	349.2	345.4
HF/6-311+G(d,p) //6-311+G(d,p)	340.9	341.5	341.2	349.0	345.1
MP2/6-31G(d,p) //6-31G(d,p)	347.5	346.6	346.9	354.3	351.9
MP2/6-31+G(d,p) //6-31+G(d,p)	333.9	334.1	336.1	340.6	339.0
MP2/6-311+G(d,p) //6-311+G(d,p)	334.9	335.1	336.6	341.5	339.2

[a] The general abbreviation for the calculation protocol is A/B/C, where A is the particular method used [DFT(BP86 or B3P86), HF, or MP2], B is the basis set used for the single point calculation, and C is the basis set used for geometry optimization. [b] All B3P86 calculations with different basis sets were performed at the BP86/DZVPD optimized geometries.

The solvation free energies, $\Delta G_s(A^-)$ and $\Delta G_s(AH)$ in eq 1, were calculated using a reaction field approach based on the boundary element method (BEM).[40,41] In our version of the BEM for calculations of the solvation thermodynamics,[41] the reaction field of the solvent is represented by point charges on the analytically described surface of the solute-formed cavity, and the polarization of the solute is represented by additive dipolar polarizabilities on atoms. Polarization of both the solvent (expressed as polarization point charges) and the solute (expressed as induced dipoles on atoms) is described by a simultaneous system of linear equations that are iterated until convergence.[41]

The effective atomic charges used in the BEM calculation of solvation energies were obtained from fitting to the DFT/DZVPD electrostatic potentials for neutral and anionic systems in optimized gas-phase geometries. The atomic radii of Rashin et al.[42,43] were used for hydration energy calculations. These radii generally give accurate hydration free energies when the dielectric continuum approximation is used for the solvent.[44] We have recently shown[42] for 50 small molecules that a combined DFT/DZVPD-BEM approach provides agreement between experimental and theoretical solvation energies better than 1.5 kcal/mol. It should be noted that this value is often the result of compensation of larger errors on the order of 3 kcal/mol in individual enthalpies and entropies of hydration.[42] Similar accuracies, within 5% (corresponding to up to 3 kcal/mol for a charged molecule of the imidazole size), were found earlier in the calculations of hydration enthalpies of monatomic[45] and polyatomic ions.[41] For ionic species, "experimental" free energies of hydration are obtained from measurable quantities

relative to the free energy of solvation of the proton, and it further complicates evaluations of the agreement between the calculated and experimental values.

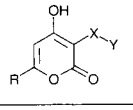
The proton solvation free energy $\Delta G_s(H^+)$, the last term in eq 1, unfortunately, is not known with high precision. The proposed values for $\Delta G_s(H^+)$ range from -259.5 [7] to -262.5 kcal/mol.[46] Our recent calculations on pK_a values of substituted imidazoles[29] indicated that the solvation free energy of the proton is closer to the lower end of this range. Moreover, our direct quantum mechanical calculations using a combination of explicit solvent molecules and dielectric continuum[47] also gave a value of the proton free energy of solvation close to -262.5 kcal/mol. This value was used in the present work for the calculation of absolute pK_a values. The current controversy regarding the magnitude of the proton free energy of solvation[47] causes uncertainty in any estimate of accuracy in predictions of absolute hydration energies of ions. However, most of the uncertainties cancel in computations of relative hydration thermodynamics in a series of closely related compounds, which is the interest in this work. While it is not guaranteed that the errors always cancel, it is highly probable as shown by our calculations[29,42] and calculations by others.[30] Thus, the determination of relative pK_a values for this series of structurally related HIV protease inhibitors can be expected to provide accurate values.

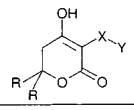
Results and Discussion

The crucial step in the absolute pK_a calculations, as mentioned above, is an accurate calculation of the energy difference (ED) between the neutral and deprotonated molecule (E_{SCF} (A^-) $-$ E_{SCF}(AH)). The results of those calculations for the HF, MP2, and DFT approximations, using different basis sets, are shown in Table 2. The ED is highly dependent on the quantum chemical approximation and basis set used. This strong dependence is primarily a general problem of the quantum chemical description of the negatively charged ion. At the HF/6-31G(d,p) level the EDs are highly overestimated. Inclusion of electron correlation effects and extension of the basis set by including diffuse functions (6-31+G(d,p)) decreases the ED to near the DFT values. In general, there is a much weaker basis set dependence of the ED values obtained in the DFT approximation compared to the HF and MP2 data. Let us consider the basis set dependence of the DFT results in more detail for the case of the B3P86 functional, mostly used for the gas phase and solvation energies calculations throughout this work. As can be seen from Table 2, the addition of diffuse s,p functions on hydrogens changes the ED values by less than 0.1 kcal/mol. The large increases in the basis set size (from 6-311++G(d,p) to aug-cc-pVTZ) lead to a consistent rise of the ED values by less than 2 kcal/mol with respect to the reference ED values obtained with the 6-311+G(d,p) basis set, constituting less than 0.5% of the total ED value. Considering the opposite sign of the vibrational energy contributions due to the basis set increase, the corresponding changes to the enthalpy of ionization (eq 3) will be even smaller. More importantly, the weak basis set dependence observed do not influence within 1 kcal/mol accuracy the relative changes of the ED values for different substituents (Table 3) among the set of pyrone compounds, and we can expect reliable prediction of corresponding relative pK_a values. Thus, the 6-311+G(d,p) basis set mostly used in this paper for energy evaluation is a reasonable compromise between accuracy and efficiency for large pK_a calculations.

The importance of the specific level of approximations used for the absolute pK_a calculations can also be seen by comparing the ED values obtained for the pyrone models with the S and CH_2 substituents (Table 2). In the HF/6-31G(d,p) approximation the ED for X = S is smaller than the ED with X = CH_2 by 1.2 kcal/mol. This result seemed contrary to chemical intuition. In the geometry optimization, the lowest energy conformer with X = S exhibited an internal hydrogen bond between OH and S in the neutral molecule. One would therefore expect that the stronger intramolecular hydrogen bond in the pyrone model with X = S would exhibit a larger ED than that with X = CH_2. Increasing the size of the basis set reduced the ED. Only at the MP2 level, using the extended 6-311+G(d,p) basis set, does the ED for X = S become larger than the corresponding ED for X = CH_2. As can be seen from Table 2, the ED values calculated by the DFT method using the TZVP, DZVPD, and 6-311+G(d,p) basis sets are quite similar and close to the ED values obtained at the MP2/6-31+G(d,p) level. This is another reason we used the 6-311+G(d,p) basis set in the B3P86 approximation to calculate ED values for all absolute pK_a determinations presented below.

TABLE 3: Absolute[a] and Relative[b] pK_a Values Calculated for the Pyrone and Dihydropyrone Models with Different Substituents X and Groups Y and R





X	Y	R=Cl	R=H	R=CH_3	R=H	R=CH_3
H		2.0 [-1.1]	3.4 [-1.4]	5.1 [1.0]	3.5 [-0.7]	4.5 [-1.3]
CH_2	CH_3	3.1 [0.0]	4.8 [0.0]	6.1 (5.9)[c] {0.0}(0.0)[c]	4.2 [0.0]	5.8 (5.5)[c] [0.0](0.0)[c]
S	CH_3	1.7 [-1.4]	3.2 [-1.6]	4.5 (4.1)[d] [-1.6](-1.8)[d]	2.7 [-1.5]	3.6 (3.6)[c] [-2.2](1.9)[c]
NH	CH_3	5.8 [2.7]	7.4 [2.6]	8.8 (8.6)[c] [2.7](2.7)[c]	7.4 [3.2]	7.7 [1.9]
O	CH_3	3.0 [-0.1]	4.9 [0.1]	6.2 [0.1]	3.9 [-0.3]	4.8 [-1.0]

[a] Absolute pK_a values were calculated using the proton solvation energy value 262.5 kcal/mol. Values in parentheses correspond to the experimentally measured pK_a values. [b] Relative pK_a values are in brackets. The relative pK_a values for the pyrone and dihydropyrone models with X = CH_2 were set equal to 0.0. Values in parentheses correspond to the experimentally measured pK_a values. [c] Tummino, P. J. Unpublished results. R = Ph and Y = CH_2Ph. [d] Reference 2. R = Ph and Y = Ph(2-^{i}Pr).

We calculated absolute pK_a values for simplified model pyrone and dihydropyrone compounds with different substituents (Table 3). The experimental inhibitors, while structurally more complex, could all be represented by at least one of the model compounds for which computations were performed. For simplification we used methyl and vinyl groups to represent sp^3 and sp^2 hybridized carbon atoms, respectively, at X and R groups. As can be seen from Table 3, the pK_a values for pyrone models with X = S, CH_2, and NH are in good agreement with the available experimental data obtained for the analogous structures. The continuum-based solvation theory used here[41] is conceptually transparent, and its demonstrated successes indicate that it accurately captures the major features of the systems it has been applied to. Apparently some unaccounted features of a more comlex nature lead to errors that cancel out. Some possible sources of error have been discussed,[42] but they are far from being completely understood. Thus, very good agreement between the calculated and experimental absolute pK_a's (while it is achieved surprisingly often) might be fortituous. This, however, does not apply to the relative pK_a values.

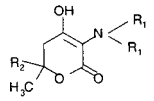
In contrast to the larger ED values calculated for X = S from the gas-phase structure, the calculated pK_a values for X = S were lower than those with X = CH_2. The difference in the relative ΔG_g and pK_a values was due to the inclusion of solvation effects (see terms $\Delta G_s(A^-)$, $\Delta G_s(AH)$, $\Delta G_s(H^+)$ in eq 1). Similarly, the increased differences in pK_a values in the model compounds with X = O with respect to those with X = NH were also due to the solvation energy differences. In addition to the good quantitative agreement between experimental and theoretical absolute pK_a values, our calculations reproduce well the pK_a differences between pyrones and their corresponding dihydropyrones. For all variations of X and R substituents, the dihydropyrone compounds have pK_a values lower than or equal to those of the corresponding pyrone compounds. In the pyrones, the replacement of an electron-donating group, R=CH_3, by an electron-accepting group, Cl, for example, leads to a large decrease in pK_a (Table 3). In the dihydropyrone models, similar substitutions also lead to low pK_a values, although these changes are not as large as seen in the case of the corresponding model pyrones (data not shown).

For a given substituent at the 6-position, different substituents at the 3-position all have a similar influence on pK_a values of the pyrones. This effect can be seen most clearly from a comparion of relative pK_a values where the pK_a for X = CH_2 was set to 0.0 for all studied R groups (Table 3). As discussed above, the relative pK_a values do not depend on the uncertainties in experimental and theoretical values of the proton solvation energy. As shown in Table 3, the relative pK_a values predicted by theory agree well with the available experimental data and reproduce the pK_a differences between pyrones and dihydropyrones.

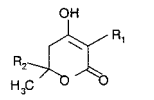
Our calculations suggested that protonated forms of pyrone and dihydropyrone are stabilized with X = NH relative to the X = CH_2 or X = S substitutions. Since most of the structure−activity data were obtained with X = CH_2 [48] or X = S,[2,3] we decided to investigate further N-substitution effects on the strong stabilization of the OH group in dihydropyrone models. The dependence of pK_a values for different dihydropyrone compounds with X = NH on the hybridization of carbon atoms is shown in Table 4. The highest pK_a values were obtained for compounds in which the hybridization of all carbon substituents was of the sp^3 type. An increase in the number of sp^2 carbons in any position led to a decrease in pK_a values, although the magnitude of the decrease depends on the locations of the sp^2 carbons. The opposite trend was observed for compounds where R_1 = piperazine, where there are two ring nitrogens (Table 4). However, when R_1 is an unsaturated analogue of piperazine, a dramatic reduction in pK_a values was obtained when an sp^2 carbon was used in the R_2 position.

Let us consider the reasons for the higher pK_a values of the OH group in the model compounds with X = NH as an example of a more general set of compounds with X = N. Compounds with X = NH are, in general, less acidic than compounds with X = S, CH_2, or O. The calculated free energies of proton abstraction in the gas phase for compounds with X = NH and X = O are similar (328.8 and 327.4 kcal/mol, respectively), suggesting that the corresponding pK_a values should be within 1 pK_a unit of each other. However, the difference between calculated pK_a values in water for compounds with X = NH and X = O was more than 2.6 pK_a units (see Table 3). Neutral pyrones with X = NH are better solvated than those with X = O by 0.6 kcal/mol. This small difference can be seen to represent the difference in interaction of NH and O groups with solvent (donation vs acceptance of H bond). In the anion of a compound

TABLE 4: Dependence of ΔG_g,[a] ΔG_s[b] (in kcal/mol), and pK_a Values of Dihydropyrone Model Compounds with X = N on the Hybridization State of the C Atoms in Groups R_1, R_1', and R_2



R_1 =	H	CH_3	CH_3	$CHCH_2$
R_1' =	CH_3	CH_3	$CHCH_2$	CH_3
R_2 =	CH_3	CH_3	CH_3	$CHCH_2$
ΔG_g	326.9	329.8	320.6	319.4
ΔG_s	-54.0	-53.8	-49.2	-48.7
pK_a	7.7	9.9	6.5	6.0



R_1 =				
R_2 =	CH_3	$CHCH_2$	CH_3	$CHCH_2$
ΔG_g	324.4	321.5	325.9	327.0
ΔG_s	-48.1	-49.2	-51.9	-51.9
pK_a	10.1	7.2	8.5	9.2

[a] ΔG_g: free energy for the deprotnation reaction in the gas phase (see Figure 3). [b] ΔG_s: difference of free energies of solvation between neutral and deprotonated compounds.

with X = NH, a strong internal H bond is formed between the NH group and 2-CO in the optimized structure. This H bond weakens the interaction of the NH group with solvent. As result, the solvation energy of the anion with X = NH (−62.6 kcal/mol) is higher than the corresponding solvation energy of the anion with X = O (−64.2 kcal/mol). This effect leads to an overall increase in the free energy of proton abstraction in the solvent for compounds with X = NH relative to that for compounds with X = O, S, or CH_2. The important question is whether this pattern of the internal hydrogen bond between X = NH and 2-CO will remain in the "real" solvent and will not be destroyed because of the external hydrogen bond interaction among enolate, the NH group, and water molecules. To investigate this possible effect, we performed a series of anion (X = NH) geometry optimizations in the aqueous solvent using the program Jaguar,[49] BLYP functional, and a 6-31++G(d,p) basis set. These calculations were performed with zero, one, and two explicit water molecules at different positions with respect to the enolate and NH groups. Figure 4 displays some of these structures and clearly demonstrates that the pattern of a strong internal hydrogen bond between X = NH and 2-CO atom is preserved for all structures optimized in solvent with and without explicit water molecules. Thus, the strong increase in pK_a value for pyrones with X = NH is mainly determined by structurally modified solvation effects.

The dependence of pK_a on sp^2 and sp^3 carbon atoms was also calculated for pyrone and dihydropyrone models with X = S and X = CH (Tables 5 and 6, respectively). The general tendency in the change of pK_a values for both substituents is the same as that observed for X = NH. The addition of sp^2 carbon substituents results in compounds with progressively lower pK_a values, with the exception of the X = S pyrone with R_1 = $CHCH_2$. This tendency is more pronounced in the dihydropyrones.

The contributions of free energy changes due to the gas-phase reaction and solvation (ΔG_g and ΔG_s, respectively; see also

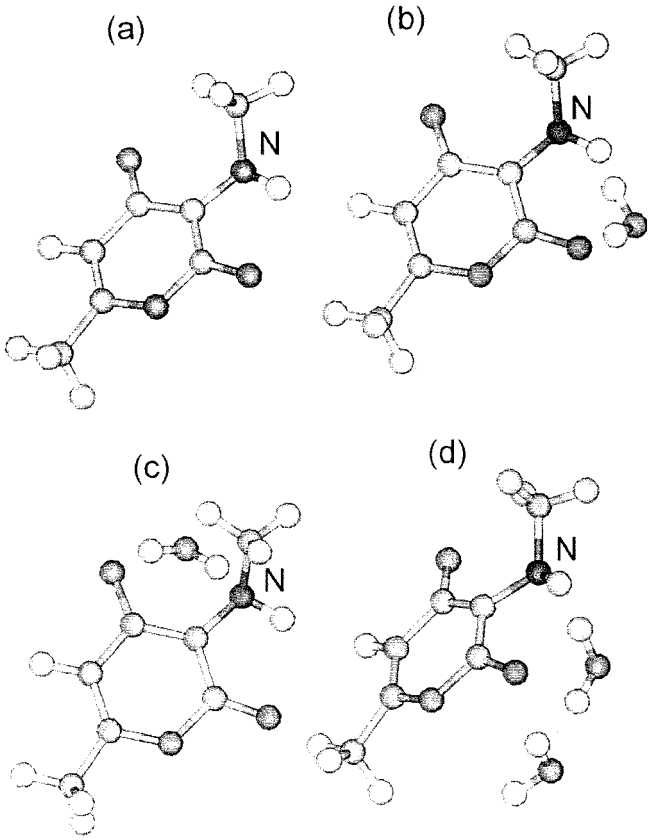
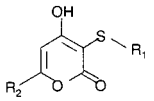


Figure 4. Structures of deprotonated pyrone models with X = NH and Y = CH_3: (a) optimized in aqueous solvent without explicit water molecules; (b, c) optimized in aqueous solvent with different positions of explicit single water molecule; (d) optimized in aqueous solvent with two explicit water molecules. In all cases the internal hydrogen bond between the NH group and 2-CO is evident.

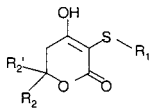
Figure 3 and eq 1) to the pK_a values are shown in Tables 4 and 5. One can see that for the studied set of model compounds the changes in ΔG_s are in general smaller than the corresponding changes in ΔG_g. In general, a decrease in ΔG_g value correlates with an increase in the ΔG_s. The pK_a changes are determined by the combined effect of those changes, which have different signs, and the resulting values do not necessarily follow exactly the tendencies in the changes of either ΔG_g or ΔG_s values.

Our calculations may help explain the pH dependence of the enzyme inhibition constants for pyrone and dihydropyrone compounds with HIV-1 protease.[2,3] At pH values above the pK_a's of inhibitor hydroxy and enzyme carboxylate groups, drug binding to the enzyme will be unfavorable due to electrostatic repulsion between ionized groups. Biochemical and NMR studies showed that the two active site carboxylates of HIV protease titrate with a single pK_a of 4.5.[50,51] Complex formation (inhibitor + active site) should favor the protonated form of the inhibitors to satisfy the overall charge requirements (−1 or neutral) of the active site based on the experimental data, implying that one or two aspartates at the active site are protonated.[53,54] Thus, the lack of antiviral activity observed for many potent pyrone- and dihydropyrone-based protease inhibitors with X = S and X = CH[1,2] may be explained by the fact that cell-based antiviral assays are performed at physiological pH conditions near pH 7.2, where these compounds should be ionized. However, some dihydropyrones with X = CH, such as PNU-140690,[55] exhibit potent antiviral activity, suggesting that either long-range substituent effects may also be important and that other factors besides the enolic equilibrium can influence enzyme binding and antiviral potency. Our calculations

TABLE 5: Dependence of ΔG_g,[a] ΔG_s[b] (in kcal/mol), and pK_a Values of Pyrone and Dihydropyrone Model Compounds with X = S on the Hybridization State of the C Atoms in Groups R_1, R_2, and R_2'



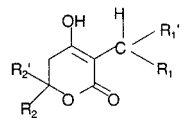
R_1=	CH_3	$CHCH_2$	CH_3	$CHCH_2$
R_2=	CH_3	CH_3	$CHCH_2$	$CHCH_2$
ΔG_g	324.3	321.4	321.9	317.8
ΔG_s	-55.6	-51.9	-53.3	-49.4
pK_a	4.5	5.1	4.5	4.3



R_1=	CH_3	$CHCH_2$	$CHCH_2$	$CHCH_2$
R_2=	CH_3	CH_3	CH_3	$CHCH_2$
R_2'=	CH_3	CH_3	$CHCH_2$	$CHCH_2$
ΔG_g	322.4	318.4	317.6	315.2
ΔG_s	-55.0	-52.2	-52.1	-50.2
pK_a	3.6	2.8	2.2	1.9

[a] ΔG_g: free energy for the deprotonation reaction in the gas phase (see Figure 3). [b] ΔG_s: difference of free energy of solvation between neutral and deprotonated compounds.

TABLE 6: Dependence of pK_a Values of Dihydropyrone Model Compounds with X = CH on the Hybridization State of the C Atoms in Groups R_1, R_1', R_2, and R_2'



	R_1=H R_1'=CH_3	R_1=H R_1'=$CHCH_2$	R_1=CH_3 R_1'=CH_3	R_1=CH_3 R_1'=$CHCH_2$
R_2=CH_3 R_2'=CH_3	6.1	5.0	6.2	5.3
R_2=CH_3 R_2'=$CHCH_2$	5.6	4.6	6.1	5.1

with model compounds predict that the inclusion of sp^3 substitutents at C-6 and at X = N should result in less acidic inhibitors, which should not exhibit pH-dependent protease inhibition over the pH range 4.0−7.0, and may lead to more favorable antiviral activity in cells.

Conclusions

In summary, we have shown that DFT methods combined with contemporary hydration continuum models can be successfully used in calculations of absolute and relative pK_a values for compounds that model the cores of synthetic pyrone and dihydropyrone inhibitors of HIV-1 protease. The agreement between our calculations and the experimental results leads us to believe that this methodology can be important for use in drug design when pK_a effects are important. The ability to accurately predict the effects of substituents at a variety of ring positions on pK_a values for pyrone and dihydropyrone com-

pounds has led to efforts to redesign and synthesize inhibitors with pH-independent protease inhibitory activity and improved antiviral potency.

Acknowledgment. We thank Drs. Vara Prasad, Brad Tait, Beth Lunney, and Mike Eissenstat for critical comments. We also acknowledge all our collaborators in the HIV project at Parke-Davis who have stimulated these studies. We thank the staff and administration of the Advanced Biomedical Computing Center for their support of this project. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. NO1-CO-56000. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services nor does mention of trade names, commercial products, or organization imply endorsement by the U.S. government.

References and Notes

(1) Prasad, J. V. N. V.; Lunney, E. A.; Ferguson, D.; Tummino, P. J.; Rubin, J. R.; Reyner, E. L.; Stewart, B. H.; Guttendorf, R. J.; Domagala, J. M.; Suvorov, L. I.; Gulnik, S. V.; Topol, I. A.; Bhat, T. N.; Erickson, J. W. *J. Am. Chem. Soc*. **1995**, *117*, 11070.

(2) Tummino, P. J.; Prasad, J. V. N. V.; Ferguson, D.; Nouhan, C.; Graham, N.; Domagala, J. M.; Ellsworth, E.; Gajda, C.; Hagen, S. E.; Lunney, E. A.; Para, K. S.; Tait, B. D.; Pavlovsky, A.; Erickson, J. W.; Gracheck, S.; McQuade, T. J.; Hupe, D. J. *Bioorg. Med. Chem*. **1996**, *4*, 1401.

(3) Tait, B. D.; Hagen, S.; Domagala, J.; Ellsworth, E. L.; Gajda, C.; Hamilton, H. W.; Vara Prasad, J. V. N.; Ferguson, D.; Graham, N.; Hupe, D.; Nouhan, C.; Tummino, P. J.; Humblet, C.; Lunney, E. A.; Pavlovsky, A.; Rubin, J.; Gracheck, S. J.; Baldwin, E. T.; Bhat, T. N.; Erickson, J. W.; Gulnik, S. V.; Liu, B. *J. Med. Chem*. **1997**, *40*, 3781.

(4) Rashin, A. A.; Rabinowitz, J. R.; Banfelder, J. R. *J. Am. Chem. Soc*. **1990**, *112*, 4133.

(5) Jorgensen, W. L.; Briggs, J. M.; Gao, J. *J. Am. Chem. Soc*. **1987**, *109*, 6857.

(6) Jorgensen, W. L.; Briggs, J. *J. Am. Chem. Soc*. **1987**, *111*, 4190.

(7) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem*. **1991**, *95*, 5610.

(8) Fitzgerald, G.; Andzelm, J. *J. Phys. Chem*. **1991**, *95*, 10531.

(9) (a) Becke, A. D. *J. Chem. Phys*. **1992**, *96*, 2155. (b) Becke, A. D. *J. Chem. Phys*. **1992**, *97*, 9173.

(10) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett*. **1992**, *197*, 499.

(11) Smith, B. J.; Radom, L. *Chem. Phys. Lett*. **1994**, *231*, 345.

(12) Schmiedekamp, A. M.; Topol, I. A.; Michejda, C. J. *Theor. Chim. Acta* **1995**, *92*, 83.

(13) Chandra, A. K.; Goursot, A. *J. Phys. Chem*. **1996**, *100*, 11596.

(14) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys*. **1991**, *94*, 7221.

(15) Curtiss, L. A.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys*. **1993**, *98*, 1293.

(16) Smith, B. J.; Radom, L. *J. Am. Chem. Soc*. **1993**, *115*, 4885.

(17) Miertus, S.; Tomasi, J. *Chem. Phys*. **1982**, *65*, 239.

(18) Hoshi, H.; Sakurai, M.; Inoue, Y.; Chujo, R. *J. Mol. Struct*. **1988**, *180*, 267.

(19) Mikkelsen, K. V.; Agren, H.; Jensen, H. J. A.; Helgaker, T. *J. Chem. Phys*. **1988**, *89*, 3086.

(20) Rashin, A. A.; Bukatin, M. A.; Andzelm, J.; Hagler, A. T. *Biophys. Chem*. **1994**, *51*, 375.

(21) Chudinov, G. E.; Napolov, D. V.; Basilesvsky, M. V. *Chem. Phys*. **1992**, *160*, 41.

(22) Tomasi, J.; Persico, M. *Chem. Rev*. **1994**, *94*, 2027.

(23) Rashin, A. A.; Young, L.; Topol, I. A. *Biophys. Chem*. **1994**, *51*, 359.

(24) Stefanovich, E. V.; Truong, T. N. *Chem. Phys. Lett*. **1995**, *244*, 65.

(25) Bachs, M.; Luque, F. J.; Orozco, M. *J. Comput. Chem*. **1994**, *15*, 446.

(26) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A., III.; Honig, B. *J. Am. Chem. Soc*. **1994**, *116*, 11875.

(27) Tawa, G. J.; Martin, R. L.; Pratt, L. R.; Russo, T. V. *J. Phys. Chem*. **1996**, *100*, 1515.

(28) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem*. **1996**, *100*, 11775.

(29) Topol, I. A.; Tawa, G. J.; Burt, S. K.; Rashin, A. A. *J. Phys. Chem. A* **1997**, *101*, 10075.

(30) Chen, J. L.; Noodleman, L.; Case, D. A.; Bashford, D. *J. Phys. Chem*. **1994**, *98*, 11059.

(31) Levin I. N. *Physical Chemistry*, 3rd ed.; McGraw-Hill: New York, 1988; Chapter 22.

(32) Andzelm, J.; Wimmer, E. J. *J. Chem. Phys*. **1992**, *96*, 1280. *UniChem*, version 4.0; Oxford Molecular, Medawar Centre, Oxford Science Park: Oxford, U.K.

(33) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem*. **1992**, *70*, 560.

(34) Rashin, A. A.; Bukatin, M. A.; Andzelm, J.; Hagler, A. T. *Biophys. Chem*. **1994**, *51*, 375.

(35) Rashin, A. A.; Young, L.; Topol, I. A.; Burt, S. K. *Chem. Phys. Lett*. **1994**, *230*, 182.

(36) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(37) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822; erratum in *Phys. Rev. B* **1986**, *34*, 7406.

(38) Becke, A. D. *J. Chem. Phys*. **1993**, *93*, 5648.

(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision C.3; Gaussian, Inc.: Pittsburgh, PA, 1995.

(40) Rashin, A. A. *Prog. Biophys. Mol. Biol*. **1993**, *70*, 73 and references therein.

(41) Rashin, A. A. *J. Phys. Chem*. **1990**, *94*, 1725.

(42) Rashin, A. A.; Young, L.; Topol, I. A. *Biophys. Chem*. **1994**, *51*, 359.

(43) Rashin, A. A.; Namboodiri, K. *J. Phys. Chem*. **1987**, *91*.

(44) Furuki, T.; Umeda, A.; Sakurai, M.; Inoue, Y.; Chujo, R.; Harata, K. *J. Comput. Chem*. **1994**, *15*, 90.

(45) Rashin, A. A.; Honig, B. *J. Phys. Chem*. **1985**, *89*, 5588.

(46) Gilson, M. K.; Honig, B. *Proteins: Struct. Funct. Genet*. **1988**, *4*, 7.

(47) Tawa, G. J.; Topol, I. A.; Burt, S. K.; Caldwell, R. A.; Rashin, A. A. *J. Chem. Phys*. **1998**, *109*, 4852.

(48) 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.; Waterpaugh, K. D. *J. Med. Chem*. **1996**, *39*, 4630.

(49) *Jaguar*, version 3.5; Schroedinger, Inc.: Portland, OR, 1998.

(50) Hogland, L. J.; Tomszek, T. A.; Meek, T. D. *Biochemistry* **1991**, *30*, 8454.

(51) Smith, M. R.; Brereton, I. M.; Chai, R. Y.; Kent, S. B. H. *Nature Struct. Biol*. **1996**, *3*, 946.

(52) Wlodawer, A.; Erickson, J. *Annu. Rev. Biochem*. **1993**, *62*, 543.

(53) Baldwin, E. T.; Bhat, T. N.; Liu, B.; Gulnik, S.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. *Structure* **1995**, *3*, 581.

(54) Wang, Y.-X.; Freedberg, D. I.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A. *Biochemistry* **1996**, *35*, 9945.

(55) Janakiraman, M. N.; Watenpaugh, K. D.; Timich, P.; Chong, K. T.; Turner, S. R.; Tommasi, R. A.; Thaisrivongs, S.; Strohbach, J. W. *Bioorg. Med. Chem. Lett*. **1998**, *8*, 1237.