Journal of Computer-Aided Molecular Design, **15:** 129–144, 2001.
KLUWER/ESCOM

The configurational dependence of binding free energies: A Poisson–Boltzmann study of Neuraminidase inhibitors

Christopher J. Woods[a], Michael A. King[b] & Jonathan W. Essex[a,*]
[a]*Department of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK;* [b]*Celltech Chiroscience plc, 216 Bath Road, Slough, SL1 4EN, UK*

Received 24 March 2000; Accepted 11 August 2000

Key words: affinity, conformation, ensemble, implicit solvent, snapshot

Summary

The linear finite difference Poisson-Boltzmann (FDPB) equation is applied to the calculation of the electrostatic binding free energies of a group of inhibitors to the Neuraminidase enzyme. An ensemble of enzyme-inhibitor complex conformations was generated using Monte Carlo simulations and the electrostatic binding free energies of subtly different configurations of the enzyme-inhibitor complexes were calculated. It was seen that the binding free energies calculated using FDPB depend strongly on the configuration of the complex taken from the ensemble. This configurational dependence was investigated in detail in the electrostatic hydration free energies of the inhibitors. Differences in hydration energies of up to 7 kcal mol^{-1} were obtained for root mean square (RMS) structural deviations of only 0.5 Å. To verify the result, the grid size and parameter dependence of the calculated hydration free energies were systematically investigated. This showed that the absolute hydration free energies calculated using the FDPB equation were very sensitive to the values of key parameters, but that the configurational dependence of the free energies was independent of the parameters chosen. Thus just as molecular mechanics energies are very sensitive to configuration, and single-structure values are not typically used to score binding free energies, single FDPB energies should be treated with the same caution.

Abbreviations: FDPB, finite-difference Poisson Boltzmann; PB, Poisson Boltzmann; GB, generalised Born; MD, molecular dynamics; LIE, linear interaction energy; MC, Monte Carlo; RMSD, root mean square deviation; RMS, root mean square.

Introduction

The linear finite-difference Poisson–Boltzmann (FDPB) equation has proven a useful tool for the rapid evaluation of electrostatic binding and hydration free energies. Its use has included the calculation of the relative binding free energies of thermolysin inhibitors [1], the electrostatic contribution to the binding free energy of the λcl repressor to DNA [2], and calculating the solvation free energies of several small molecules [3]. In FDPB the molecular detail of the solvent is replaced by a dielectric continuum, and the solute immersed within a cavity in the continuum. Good descriptions of the FDPB equation and its applications have been published elsewhere [4, 5]. An assumption often implicitly used when calculating electrostatic binding or hydration free energies with the FDPB equation, is that the whole ensemble of configurations accessible to the solute can be represented by a single configuration. Thus relative hydration or binding free energies are estimated from static, single-structure FDPB calculations [6, 7].

Recent simulations that combine FDPB with molecular dynamics appear to challenge this assumption. In one study the relative stability of different conformations of DNA helices was investigated using a combination of FDPB and molecular dynamics [8]. Molecular dynamics simulations of different confor-

*To whom correspondence should be addressed. E-mail: j.w.essex@soton.ac.uk

mations of DNA and RNA were performed in explicit solvent. To allow a comparison of the relative stabilities of these conformations, 100 snap-shots were taken from each MD simulation, and the intra-molecular and FDPB hydration energy for each was determined. The standard deviation of the FDPB hydration energy across these 100 snapshots was around 3–5 kcal mol^{-1} and in the worst case it was 18 kcal mol^{-1}. Interestingly, the hydration free energies calculated using the Generalised-Born (GB) method [9] were also calculated for the 100 snapshots. The standard deviations for these GB hydration free energies were almost identical to the standard deviations for the FDPB energies.

A similar approach was used in another study [10] to calculate the binding free energies between proteins. Explicit-water MD simulations were performed separately on a protein, a peptide and their complex, and 100 snapshots were taken from each trajectory. The intramolecular energy and FDPB hydration free energy were calculated, along with other energy contributions, and these were then combined to give an estimate of the binding free energy between the protein and peptide. The standard deviations of the FDPB hydration free energies across the 100 snapshots were very large, typically around 30–40 kcal mol^{-1}. Such large standard deviations imply that many snapshot configurations would need to be averaged to obtain a reliable binding or hydration free energy, and that the use of just a single configuration to calculate or rank binding or hydration free energies may be flawed.

The question of using a single structure in a PB calculation has already been asked in the case of pK_a calculations [11]. A molecular dynamics simulation generated an ensemble of protein conformations. The titration curve for each snap-shot structure was calculated via FDPB, and different techniques for averaging the curves were investigated. The paper concluded that conformational effects did need to be accounted for in pK_a calculations, and it went on to suggest how this may be best done.

Thus there is a lot of circumstantial evidence in the literature that PB calculations of binding free energies may be configurationally dependent. However, to our knowledge, there have been no papers that have explicitly investigated it, or tried to quantify it. Thus the aim of this work is to investigate configurational dependence in detail, and to ask the question of whether binding and hydration free energies can, via FDPB, be calculated using a single conformation of the solute.

Recently, the results of a study of the binding free energies of fifteen inhibitors of the Neuraminidase enzyme were presented [12]. Its primary goal was to examine the Linear Interaction Energy (LIE) method as a fast route to relative binding free energies. The study involved several Monte Carlo (MC) simulations of the inhibitors within the enzyme's active site, resulting in a large number of snapshot structures for each enzyme-inhibitor complex. In this paper, these snapshot structures have been used to determine whether the electrostatic binding free energy calculated using the FDPB equation is stable with respect to the small configurational changes that occurred during the Monte Carlo simulation. Since the aim of this study is only to investigate the stability of the FDPB equation, it is not necessary to calculate the non-polar and entropic contributions to the binding free energy. Consequently we shall make no attempt to compare the results with the available experimental binding free energies. Indeed, we shall demonstrate that the configurational dependence of FDPB is such that a comparison would be meaningless.

Methods

Theory

The methods underlying the FDPB equation have been well described elsewhere [13]. A typical FDPB calculation places partial atomic charges onto the finite-difference grid using a tri-linear weighting function [14]. The dielectric constant is assigned to grid-line centres, the value used being based on whether the grid-line centre is inside or outside the solute. This assignment of charge and dielectric will henceforth be referred to as the *original protocol*. Through the study of the grid convergence characteristics of the hydration free energies of small molecules, it has been shown that this protocol has two sources of error [3]. There is a random error resulting from the position of the solute within the finite-difference grid, and a systematic error resulting from the representation of the solute on a cubic grid. The systematic error increases dramatically as the grid size increases [3]. Henceforth, the random error shall be referred to as *positional error*, and the systematic error shall be referred to as *grid size error.*

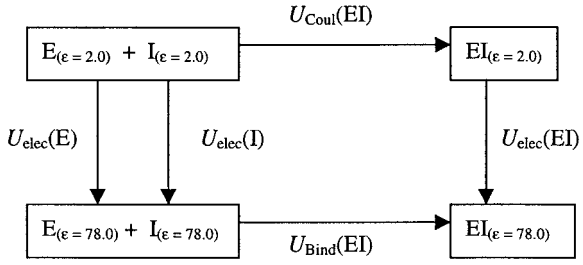
There are many methods to overcome positional error in an FDPB calculation. All of these methods involve moving the solute within the grid, calculating the electrostatic hydration free energy for each position, and calculating the mean [3, 14]. Most of these

methods involve translating or rotating the solute by systematic amounts. However, since the grid is cubic, many levels of symmetry exist. Thus systematic movement of the solute within the grid may not adequately sample the electrostatic hydration free energies, and produce reliable averages. In this study random rotation is used; the electrostatic hydration free energy of the solute is calculated for one solute orientation, then the solute is rotated at random around each of the three grid axes. These axes have their origin in the centre of the grid, and the centre of geometry of the solute is placed on this origin; any rotation will not move the solute any closer to the grid boundaries. A random number generator configured to give a mean rotation of 0.0°, and standard deviation of 90° generates the three angles of rotation.

The original protocol is very sensitive to the grid size, and the grid size error is therefore quite large. Methods to overcome this problem generally rely on making the grid spacing small, either by using a very high resolution finite-difference grid, or by using focussing to 'zoom in' on the area of interest in the solute [15]. Since the grid size error has its source in the unphysical cubic representation of the solute on the grid, a second solution is to make the charge and dielectric distribution better represent the solute's symmetry. The techniques of charge anti-aliasing and harmonic dielectric smoothing adopt this approach [16] and their use significantly reduces the grid size error as well as reducing the solute positional error of the calculation. Calculations using charge anti-aliasing and harmonic dielectric smoothing will henceforth be referred to as the *smoothed protocol.*

Poisson–Boltzmann methodology

Unless otherwise stated, all PB calculations were performed using an interior dielectric constant of 2.0, an exterior dielectric constant of 78.0 and an ionic strength of 0.0 mM. The dielectric boundary was defined by the probe accessible surface of the solute, with a probe sphere radius of 1.4 Å to represent a water molecule. Since the ionic strength was zero, the boundary potentials were obtained using Coulomb's law. The finite difference grid resolution and spacing were assigned so to keep the total grid volume roughly constant across a group of experiments, thereby minimising any influence the boundary of the grid has on the relative free energies. In addition, the solute was always centred in the finite difference grid.



(3) $U_{Bind}(EI) = U_{Coul}(EI) + U_{elec}(EI) - U_{elec}(E) - U_{elec}(I)$

Figure 1. The free energy cycle used to calculate electrostatic binding free energies. E: enzyme; I: inhibitor; EI: enzyme-inhibitor complex. U_{elec} = electrostatic hydration free energy. U_{Coul} = Coulomb association energy. U_{Bind} = electrostatic binding free energy.

In general, sixty PB calculations were performed to obtain a single electrostatic hydration free energy of one solute configuration. A calculation with the exterior dielectric at 78.0 is first performed, giving a set of potentials, $\phi(78.0)$, at each grid point, *i.* A calculation is then performed with the exterior dielectric constant equal to the interior dielectric constant. This results in a set of potentials, $\phi(2.0)$, at each of the grid points. The electrostatic hydration free energy, U_{elec}, is obtained from these potentials, and the charge, q, at each grid point using;

$$U_{elec} = \frac{1}{2}\sum_i q_i \left[\phi_i\,(78.0) - \phi_i\,(2.0)\right] \qquad (1)$$

This calculation is performed for each of thirty random rotations of the solute in the finite difference grid, resulting in thirty electrostatic hydration energies. The arithmetic mean of the hydration free energy is calculated, $\langle U \rangle$, and then the error around the mean is obtained from the number of observations, n, the estimate of the standard deviation, s, and a *t-value*, which approximates the error in estimating the standard deviation. *t-values* appropriate for a 95% confidence level were used, Equation 2;

$$U_{elec} = \langle U \rangle \pm \left[(t \times s) / (n)^{1/2}\right] \qquad (2)$$

Calculating relative binding free energies using the FDPB equation has been well described and validated [14]. Electrostatic binding free energies were calculated for single structures of the enzyme-inhibitor complex using a free energy cycle (Figure 1). The electrostatic binding free energy, U_{elec}(EI), is composed of the snapshot electrostatic hydration free energies of the inhibitor, U_{elec}(I), the enzyme, U_{elec}(E), and the enzyme-inhibitor complex, U_{elec}(EI). It also

includes the Coulomb electrostatic association energy, U_{Coul}(EI), between the inhibitor atoms and the enzyme atoms, calculated with the dielectric constant equal to the interior dielectric constant of the solutes.

Total binding free energies can be obtained by including the effects of non-polar solvent-solute interactions and solvent entropy via a surface area based function [17]. This was not performed in this work as we are primarily interested in the artefacts present in PB calculations, and not in comparing the results of these calculations to experiment.

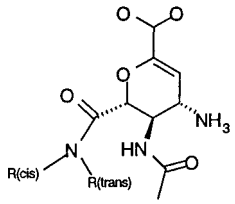
Structures and parameters

All structures were obtained from the Monte Carlo simulations performed for the LIE calculations [12]. To maintain consistency the parameters used in the FDPB calculations were the OPLS [18] parameters used in the Monte Carlo simulations. The radii used were the half-sigma values, in keeping with the Monte Carlo simulation and the recommendations of other workers [3]. Full details of the parameters and structures used are given in the study [12]. The only difference in parameters between this FDPB study and the Monte Carlo simulations is the radii of the polar hydrogens. The MC simulations assigned a zero radius to the polar hydrogens. The smoothed protocol does not allow charged atoms to have zero radius, and small, highly charged hydrogens can cause numerical instabilities in the FDPB method [15]. A polar hydrogen radius of 0.8 Å has been suggested [16] and was used in all calculations, unless otherwise specified.

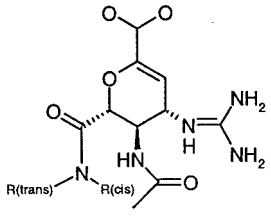
The Monte Carlo simulations were performed using the MCPRO package [19]. PDB structures of the enzyme-inhibitor complex solvated by explicit water molecules were produced every 0.1 million (M) MC configurations. Three of these snapshot structures were taken from each enzyme-inhibitor complex. The explicit waters were removed, and the structure was split into an enzyme and inhibitor. The three configurations taken were labelled A, B and C. These structures were separated from each other by 5 M MC steps; thus there were 10 M MC steps between the A and C structures. The structures of the inhibitors are illustrated in Figure 2.

Programs

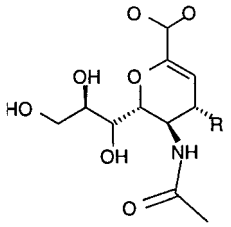
All electrostatic calculations were carried out using UHBD version 5.1 [20]. The code was modified to implement charge anti-aliasing [16], 15-point harmonic



Inhibitor	R(trans)	R(cis)
1	-Me	-H
2	-Et	-H
3	-Me	-Me
4	-Et	-Et
9	$-(CH_2)_2Ph$	-H
10	$-(CH_2)_2Ph$	-Pr



Inhibitor	R(trans)	R(cis)
5	-Me	-H
6	-Me	-Me
7	-Et	-Et
8	$-(CH_2)_2PH$	-Pr



Inhibitor	R
11	$-NH_3^+$
12	$-NH_3^+$(R)
13	$-N(CH_3)_3^+$
14	-OH
15	$-NH{=}C(NH_2)_2^+$

Figure 2. Structures of the fifteen inhibitors.

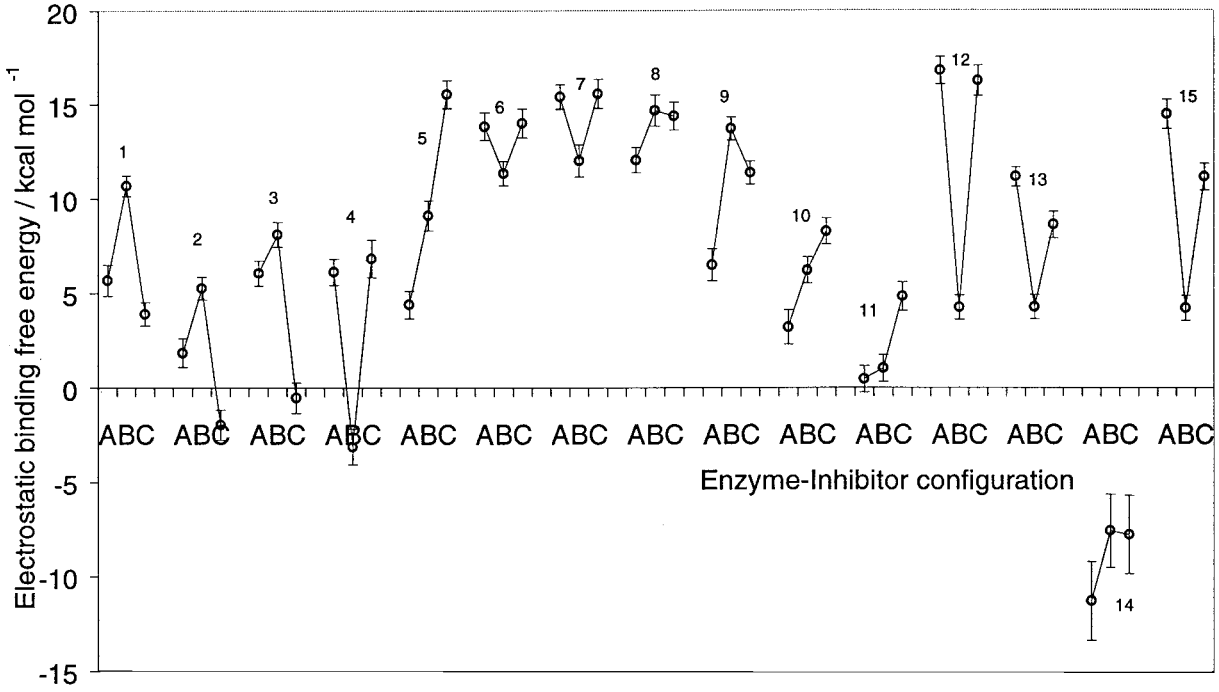


Figure 3. Calculated electrostatic binding free energies of the fifteen Neuraminidase-inhibitor complexes. Three configurations of each system were examined; A, B and C.

dielectric smoothing [16], and random rotation of the solute within the finite difference box. The angles for the rotation were generated using the random number generator already present in UHBD. The random number seed was the same for each experiment and so for the 30 random rotations normally employed, the same 90 random angles were used.

Results and discussion

Initial electrostatic binding free energy calculation

The electrostatic binding free energies of the A, B and C configurations of each of the 15 Neuraminidase-inhibitor complexes were calculated using Equation 3. The electrostatic hydration free energies of the inhibitors were determined on a finite difference grid of resolution $(38 \times 0.8\ \text{Å})^3$. The electrostatic hydration free energies of the enzyme and the enzyme-inhibitor complex were determined on a grid of resolution $(100 \times 0.8\ \text{Å})^3$. The same relative positions of the enzyme and enzyme-inhibitor complex within the grid were used. All calculations used the smoothed protocol. It will be demonstrated subsequently that the use of this protocol with these grid resolutions is sufficient to achieve converged free energies. The resulting electrostatic binding free energies (Figure 3) show strong configurational dependence; the Neuraminidase-inhibitor 5 complex has a 10 kcal mol^{-1} difference in electrostatic binding free energy between the A and C structures. The results show that it is difficult to qualitatively rank the inhibitor's binding strength, as choosing different structures results in a different ranking. These results are particularly significant as the enzyme-inhibitor complex structures were obtained from MC simulations where only minor structural changes occurred.

Electrostatic hydration free energies of the Neuraminidase inhibitors

The configurational dependence of the electrostatic binding free energy suggests that the electrostatic hydration energies of the inhibitors are also configurationally dependent. The assumption that a single snapshot configuration could represent the whole ensemble of configurations accessible to the inhibitor will be investigated. Since the inhibitor molecules were relatively small, and thus could fit onto grids with fewer grid points, it was possible to exhaustively investigate both the grid size dependence and

Table 1. Finite-difference grids used to calculate the electrostatic hydration energies of the inhibitors

Grid Spacing/Å	Grid Dimensions
0.3	$87 \times 87 \times 87$
0.4	$65 \times 65 \times 65$
0.6	$50 \times 50 \times 50$
0.8	$38 \times 38 \times 38$
1.0	$30 \times 30 \times 30$
1.2	$25 \times 25 \times 25$
1.4	$22 \times 22 \times 22$
1.6	$19 \times 19 \times 19$
1.8	$19 \times 19 \times 19$
2.0	$18 \times 18 \times 18$

the parameter dependence of the inhibitor hydration free energies. This allowed a thorough investigation of configurational dependence, to determine whether it is an artefact of the finite difference procedure, or whether it is a real effect in the PB equation.

To investigate configurational dependence, both the A and C configurations of each of the inhibitors were used, as illustrated in Figure 4. The figure shows the A and C configurations superimposed. These structures were not fitted, and thus they represent the inhibitors relative position and orientation within the enzyme's active site. The small differences between the configurations are illustrated in renderings of inhibitor 3 from multiple viewpoints (Figure 5).

The electrostatic hydration free energy of each structure was determined on finite difference grids of increasing resolution, giving the grid resolution dependence of the electrostatic hydration free energies. The calculations were performed using both the original and smoothed protocols to ensure that any artefact in either method did not compromise the results. The finite difference grids used in these calculations are detailed in Table 1. The resulting electrostatic hydration free energies of two representative inhibitors are shown in Figure 6. Inhibitor 3 is shown as an example of an inhibitor that displays strong configurational dependence, while inhibitor 12 is shown as an example of weak configurational dependence. The differences in free energy between the A and C configurations of each of the inhibitors, together with the root mean square deviations (RMSD) between the configurations, are tabulated in Table 2. These differences were calculated on the grid of highest resolution.

Table 2. Difference in electrostatic hydration free energy between configurations A and C of the fifteen Neuraminidase inhibitors

Inhibitor	RMSD / Å	Difference between the electrostatic hydration free energy of configuration C and configuration A/kcal mol^{-1}	
		Original protocol 0.3 Å grid spacing	Smoothed protocol 0.3 Å grid spacing
1	0.4	-0.73 ± 0.27	-0.73 ± 0.23
2	0.3	-1.04 ± 0.29	-1.02 ± 0.26
3	0.56	-6.96 ± 0.27	-6.89 ± 0.21
4	0.44	-0.71 ± 0.28	-0.81 ± 0.25
5	0.45	3.92 ± 0.33	4.05 ± 0.31
6	0.32	1.66 ± 0.41	1.64 ± 0.36
7	0.50	4.19 ± 0.40	4.13 ± 0.33
8	0.64	-0.50 ± 0.33	-0.62 ± 0.40
9	0.98	-2.32 ± 0.32	-2.20 ± 0.30
10	1.02	1.40 ± 0.35	1.38 ± 0.32
11	0.39	0.53 ± 0.37	0.40 ± 0.28
12	0.73	0.97 ± 0.34	0.86 ± 0.30
13	0.59	3.03 ± 0.26	2.99 ± 0.26
14	0.65	-0.21 ± 0.20	-0.23 ± 1.02
15	0.32	0.69 ± 0.32	0.75 ± 0.27

The results show a clear dependence of the electrostatic hydration free energy on the configuration of the inhibitor used. In some cases this configurational dependence is quite low, while in others it is as much as 7 kcal mol^{-1}. The typical configurational dependence is 1–2 kcal mol^{-1}, for a RMSD between configurations of 0.5–0.7 Å. Also, since the configurational dependence is consistent across the range of resolutions of the finite-difference grid, and across both protocols, it is not an artefact of the finite-difference procedure. Additionally, Figure 6 shows how much more stable the smoothed protocol is with respect to the grid resolution. A reliable estimate of the electrostatic hydration free energy may be obtained at grid spacings up to 1.0 Å using this methodology, while the original methodology has still not converged at 0.3 Å spacing.

Parameter dependence of the electrostatic hydration free energies

It may be argued that the source of the configurational dependence is not in the PB calculation itself, but rather in the choice of the parameters. The main parameters used in a PB calculation are the atomic

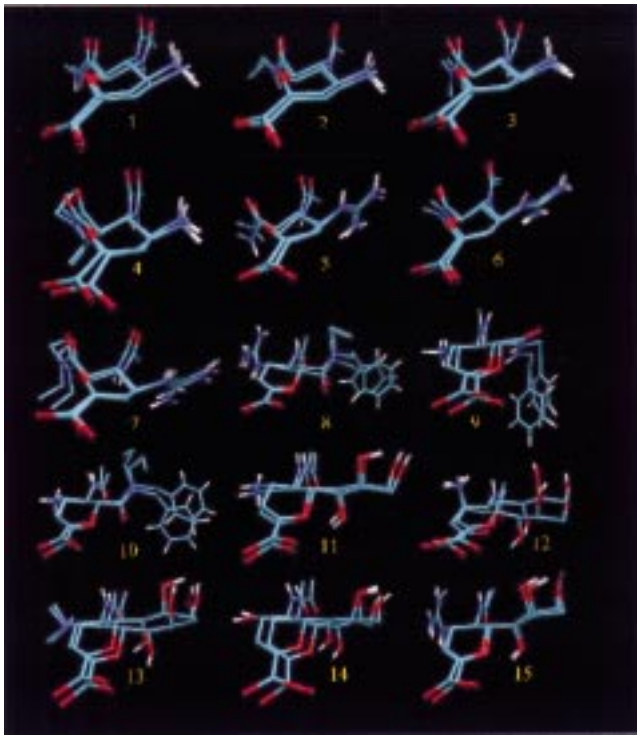


Figure 4. Configurations A and C of the fifteen Neuraminidase inhibitors.

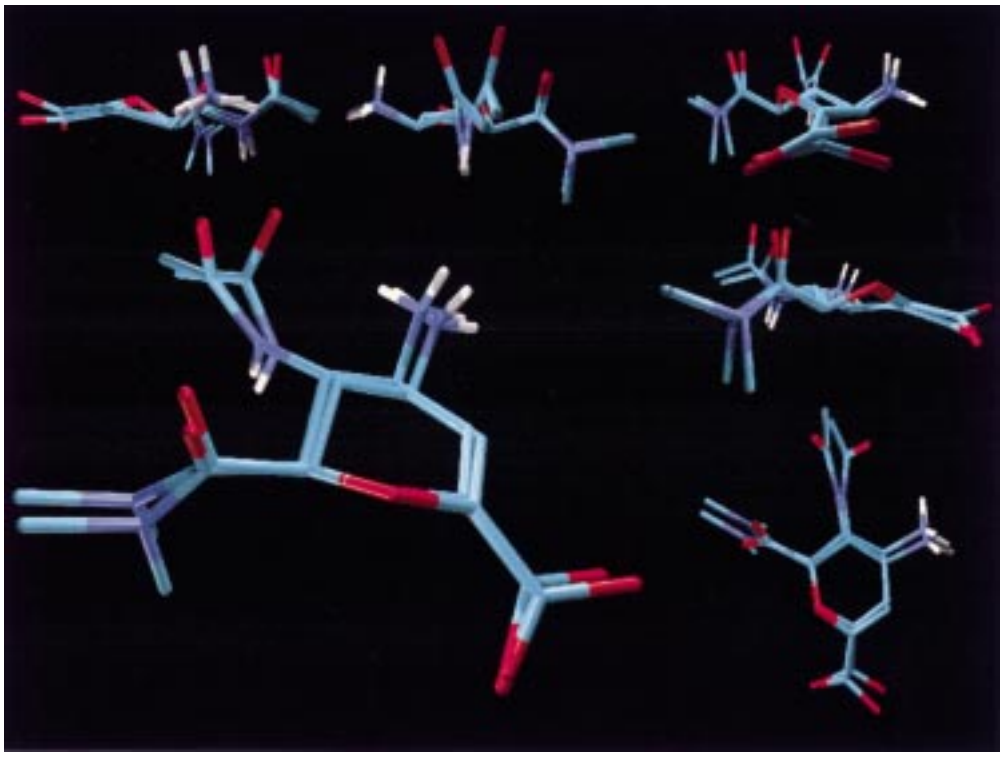


Figure 5. Configurations A and C of inhibitor 3 rendered from multiple viewpoints.

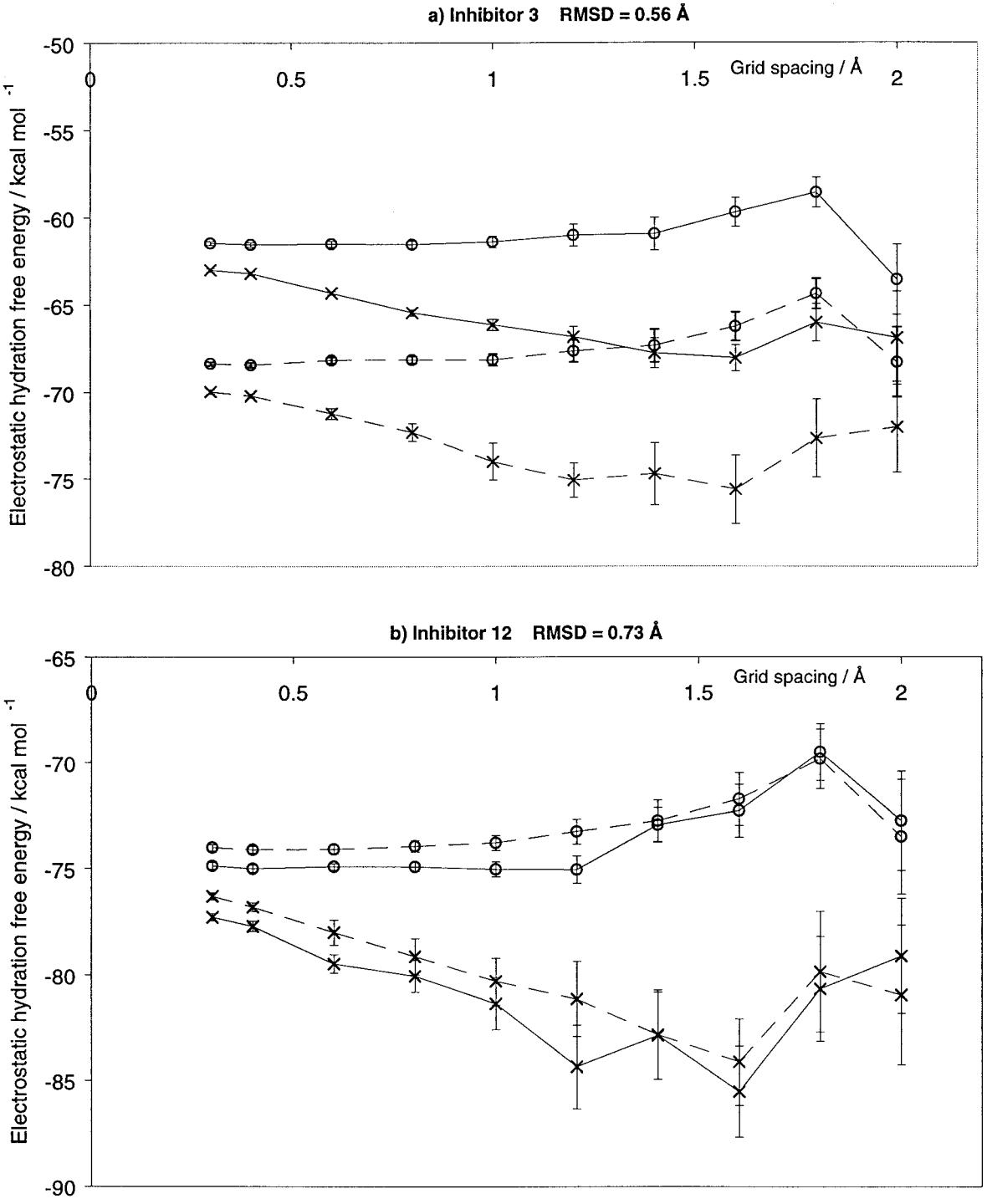


Figure 6. Grid resolution dependence of the electrostatic hydration free energies of inhibitors 3 and 12.
Solid line: configuration C; broken line: configuration A; circle: smoothed protocol; cross: original protocol. Error bars show the 95% confidence limits.

co-ordinates, charges and radii. The co-ordinates were obtained from simulation, while the charges and radii were from a standard, widely used force-field. The use of standard force-fields in FDPB calculations has been validated by previous authors [3]. However there are more subtle, yet equally critical, parameters that are used in the FDPB calculation. These include the radius of the probe sphere used to obtain the probe accessible surface, the radius of the polar hydrogen atoms and the value of the interior dielectric constant.

The 'correct' value of the probe sphere radius is subject to debate [14]. It is widely accepted that a value of 0 Å (corresponding to the van der waals surface of the solute) should not be used as this allows 'solvent' into all of the cracks and crevices in the solute. This is unphysical and increases the grid position sensitivity of the calculation.

To determine whether the configurational dependence is related to the probe sphere radius, the hydration free energy of the two configurations of inhibitor 3 were calculated for various values of the probe sphere radius. A finite difference grid resolution of $(65 \times 0.4$ Å$)^3$ was used. Calculations using both the original and smoothed protocols were performed (Figure 7). The results show that for both the original and smoothed protocols the configurational dependence is consistent across the range of probe sphere radii, demonstrating that the configurational dependence is not related to this parameter.

Polar hydrogen atoms can cause problems for PB calculations, as they represent highly concentrated charges and can thus cause numerical instabilities [15]. As for the probe radius, there is no agreement in the literature as to what the 'correct' polar hydrogen radius should be, and values from 0.8 Å [16] to 2.5 Å [3] have been suggested. To determine whether the configurational dependence is related to the choice of polar hydrogen radius, the electrostatic hydration free energies of the two configurations of inhibitor 3 were calculated for different polar hydrogen radii. Again, both the original and smoothed protocols were used (Figure 8). The results show that the configurational dependence is consistent across the range of hydrogen radii, and is thus independent of this parameter.

Since the PB equation is representing a molecular system with a macroscopic dielectric constant, there is the question as to what the dielectric constant of the solute should be. It is generally agreed that calculations on small molecules should use a value of 2.0. However there is disagreement in the literature as to what the value for proteins should be; values through 1.0 [8], 3.0 [1] to 4.0 [21] are typically used. Owing to the computational demands of calculating converged FDPB energies of large proteins, the effect of varying this parameter was investigated by calculating the electrostatic hydration free energy of the two configurations of inhibitor 3. Calculations were performed at values of the interior dielectric constant between 1.0 and 4.0, for both the original and smoothed protocols (Figure 9). The results show that the configurational dependence is unrelated to this parameter.

These results confirm that the configurational dependence of inhibitor 3 is not related to the values of these three important parameters. It is also seen how critically the 'absolute' electrostatic hydration free energies depend on the value of these parameters. This demonstrates that care must be taken when calculating an 'absolute' PB hydration free energy, and suggests that only relative electrostatic free energies can be reliably quoted.

Table 3. Finite-difference grids used to calculate the electrostatic hydration energies of the enzymes and enzyme-inhibitor complexes

Grid Spacing Å	Grid Dimensions
0.3	$250 \times 250 \times 250$
0.6	$134 \times 134 \times 134$
0.8	$100 \times 100 \times 100$
1.0	$80 \times 80 \times 80$
1.2	$67 \times 67 \times 67$
1.6	$58 \times 58 \times 58$
1.6	$50 \times 50 \times 50$
1.8	$45 \times 45 \times 45$
2.0	$40 \times 40 \times 40$

Grid size dependence of the enzyme and complex

While conformational dependence is a big effect in the small inhibitor molecules, there still remains the question as to whether it is important in a larger molecule. To examine this, two configurations of the enzyme-inhibitor complex and enzyme from the inhibitor 3 simulations and two from the inhibitor 12 simulations were taken. The electrostatic hydration free energies of each of these structures were calculated on finite difference grids of decreasing resolution. Details of the finite difference grids used are shown in Table 3. Owing to the computational expense of using such large finite difference grids, calculations on the 0.3 Å

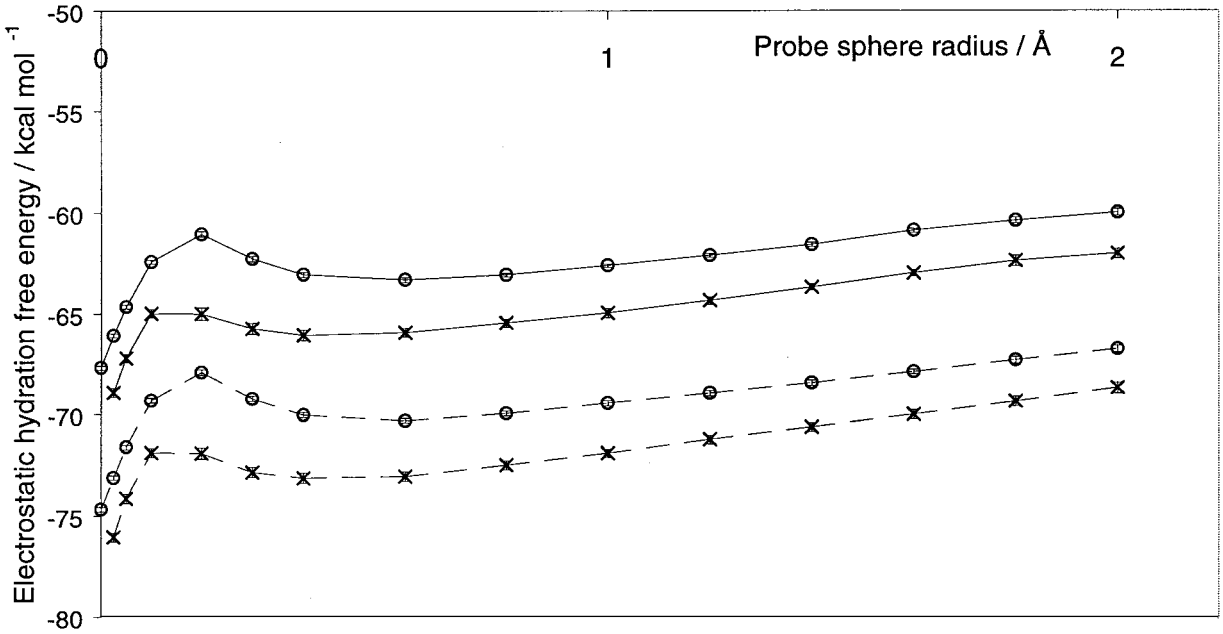


Figure 7. Dependence of the electrostatic hydration free energy of inhibitor 3 on the value of the probe sphere radius. Solid line: configuration C; broken line: configuration A; circle: smoothed protocol; cross: original protocol. Error bars show the 95% confidence limits.

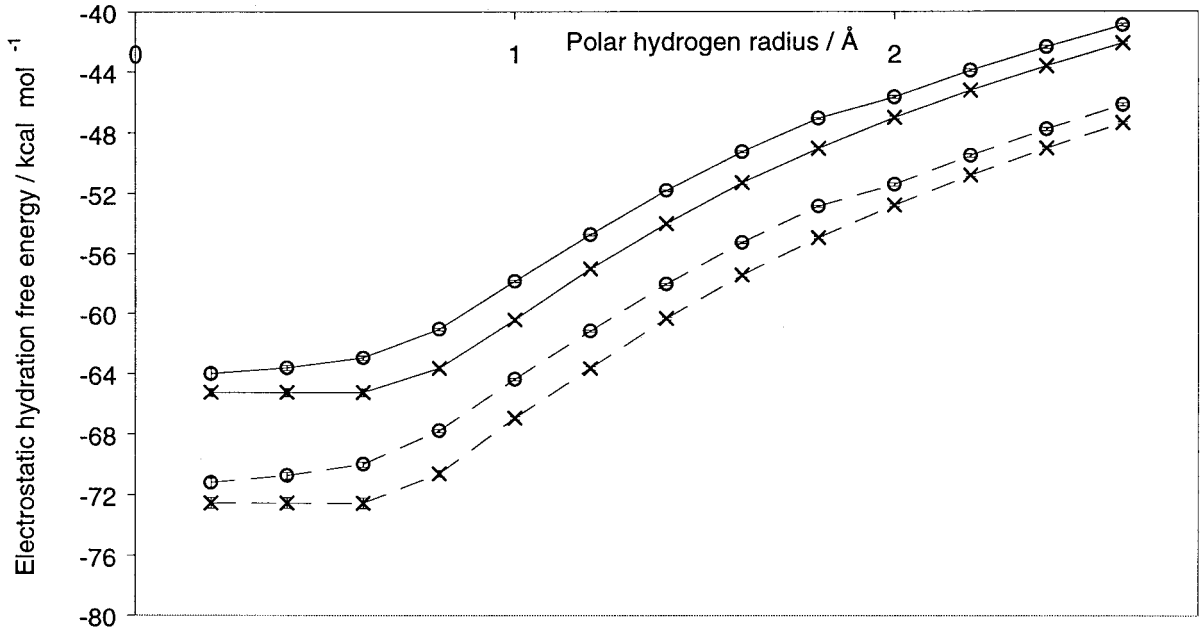


Figure 8. Dependence of the electrostatic hydration free energy of inhibitor 3 on the value of the polar hydrogen radius. Solid line: configuration C; broken line: configuration A; circle: smoothed protocol; cross: original protocol. Error bars show the 95% confidence limits.

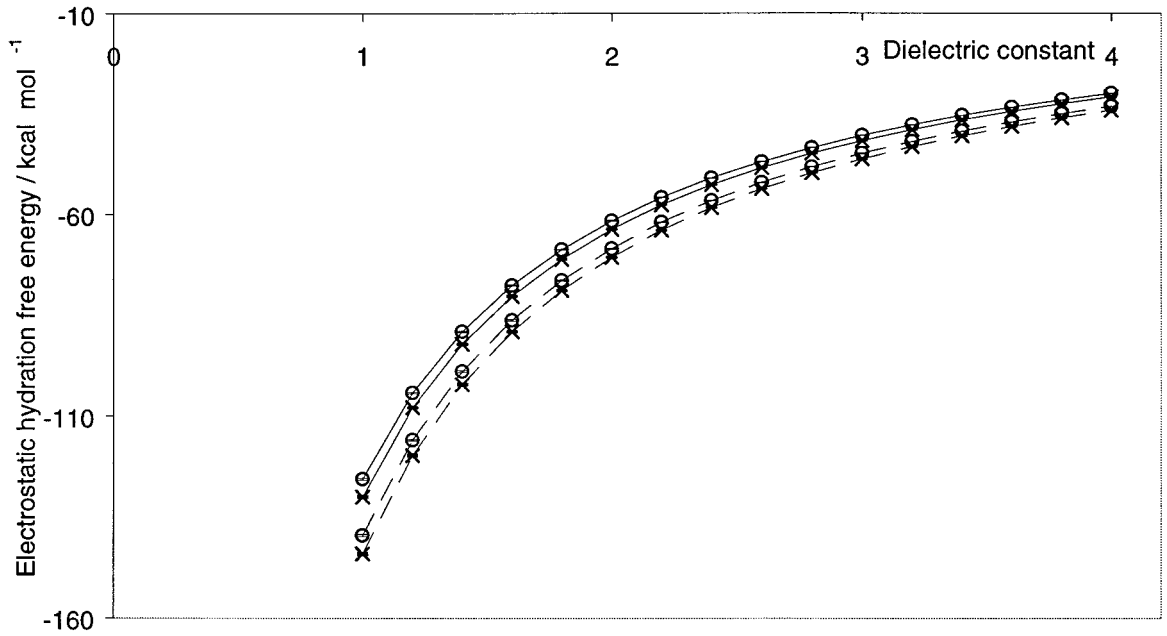


Figure 9. Dependence of the electrostatic hydration free energy of inhibitor 3 on the value of the interior dielectric constant. Solid line: configuration C; broken line: configuration A; circle: smoothed protocol; cross: original protocol. Error bars show the 95% confidence limits.

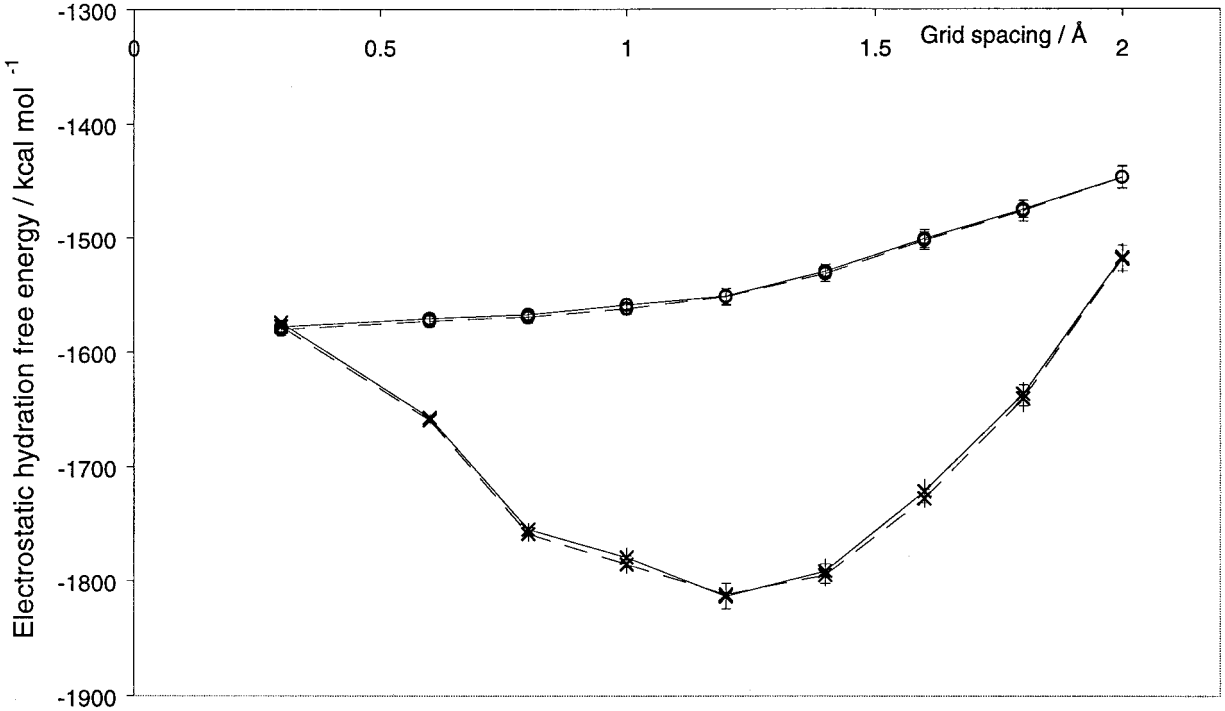


Figure 10. Grid resolution dependence of the electrostatic hydration free energy of configurations A and C of enzyme 3. Solid line: configuration C; broken line: configuration A; circle: smoothed protocol; cross: original protocol. Error bars show the 95% confidence limits.

Table 4. Difference in electrostatic hydration free energy between configurations A and C of the enzyme, and the enzyme-inhibitor complex.

System	Difference between the electrostatic hydration free energy of configuration C and configuration A/kcal mol^{-1}	
	Original protocol 0.3 Å grid spacing	Smoothed protocol 0.3 Å grid spacing
Enzyme from inhibitor 3 system	-3.1 ± 2.4	-2.5 ± 1.8
Enzyme from inhibitor 12 system	-3.2 ± 1.9	-1.1 ± 1.1
Enzyme-Inhibitor 3 complex	2.0 ± 2.3	1.8 ± 2.1
Enzyme-Inhibitor 12 complex	-2.8 ± 1.9	-1.9 ± 1.4

Table 5. Configurational difference of the difference in hydration free energy between the complex and the enzyme

System	Difference between the electrostatic hydration free energy of configuration C and configuration A/kcal mol^{-1}	
	Original protocol 0.3 Å grid spacing	Smoothed protocol 0.3 Å grid spacing
Difference between complex 3 and enzyme	5.2 ± 2.8	4.3 ± 1.7
Difference between complex 12 and enzyme	0.4 ± 1.7	-0.7 ± 1.1

grids used only 10 random rotations as opposed to the 30 normally employed. Calculations using both the smoothed and original protocols were performed. The grid resolution dependence of the two configurations of enzyme 3 is shown in Figure 10. It is seen that the smoothed protocol converges much more rapidly than the original protocol, and that both protocols converge to the same hydration free energy. However, while the smoothed protocol produces hydration free energies that are stable up to 0.8 Å, the original protocol only just converges at 0.3 Å. At grid spacings that are larger than this, the original protocol produces absolute hydration free energies with significant systematic errors. A similar picture is seen in the grid dependence of hydration free energies of enzyme 12, and of the complexes (data not shown).

The converged difference in electrostatic hydration free energy between the two configurations of the enzymes and complexes are given in Table 4. This shows that the configurational difference is, to within error, the same regardless of the protocol used, provided that the grid spacing is small. The superiority of the smoothed protocol is also shown in the smaller errors in the configurational dependence, compared to those calculated by the original protocol.

The results show that the configurational dependence of these large molecules is relatively weak in percentage terms. However, the electrostatic hydration free energies are of the order of 1500 kcal mol^{-1}, so a small percentage configurational dependence may correspond to a large absolute dependence. This is best illustrated by calculating the difference between the electrostatic hydration free energy of the enzyme-inhibitor complex, and the enzyme. From Equation 3, this difference is required as part of the calculation of the electrostatic binding free energy. The difference is calculated by ensuring that the enzyme and enzyme-inhibitor complex have equivalent positions on the finite difference grids, and that the random rotations used on both systems are the same. The electrostatic hydration free energies of both systems are determined for each rotation, as is their difference. After all rotations are completed the average of the differences is taken, along with the associated standard deviation. This procedure greatly minimises the error in the calculation. The grid dependence of the differences of the enzyme-inhibitor 3 complex is shown in Figure 11. This again shows that the smoothed protocol converges much more rapidly than the original protocol, with a stable difference possible on a grid of 0.6 Å, and a reasonable approximation possible at 0.8 Å. This compares very favourably to the original protocol, where the difference varies widely with the finite difference grid spacing and the associated standard errors are much larger; it seems to converge only as the grid spacing drops below 0.4 Å. A similar picture is seen for the enzyme-inhibitor 12 complex (data not shown).

The converged configurational dependence of this difference is also given in Table 5. Again, within error, both protocols calculate the same configurational difference, provided small grid spacing is used, with smaller errors calculated using the smoothed protocol. Combining the enzyme and complex hydration free energy data with the inhibitor and Coulomb association energy data allows the grid resolution dependence

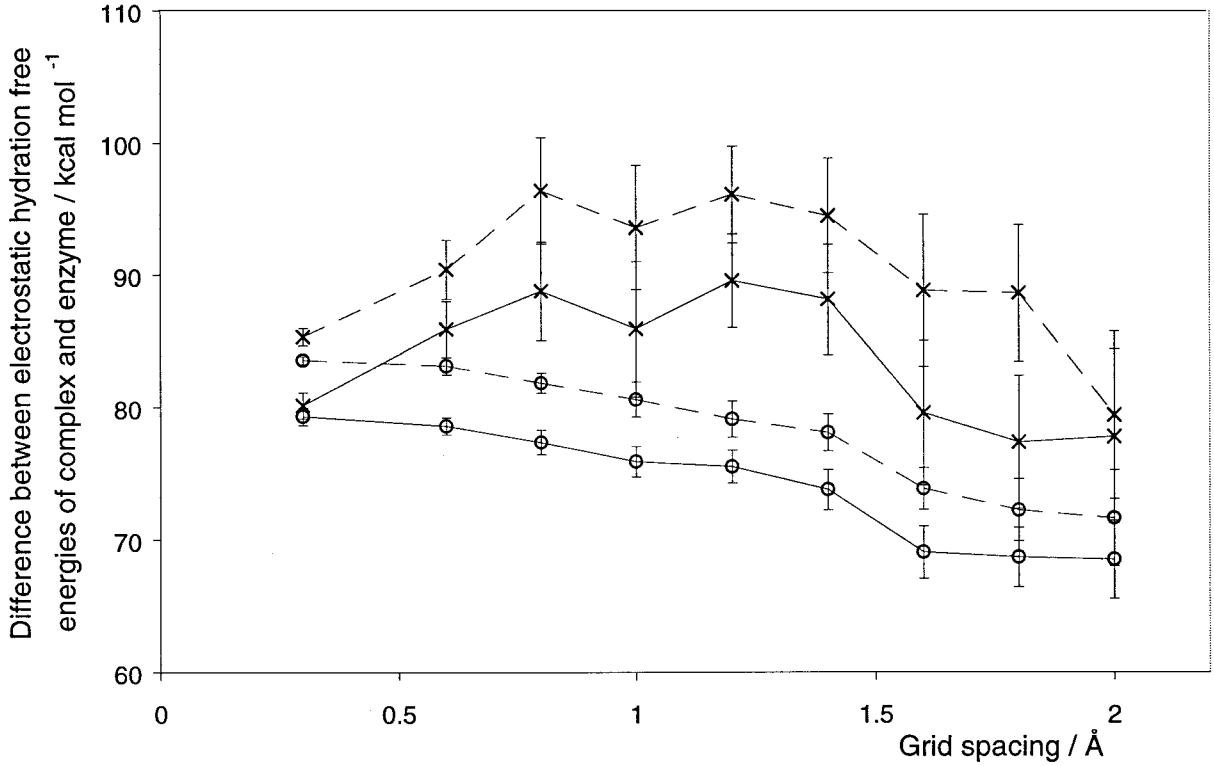


Figure 11. Grid resolution dependence of the difference between the electrostatic hydration free energies of the complex and enzyme, for configurations A and C of enzyme 3. Solid line: configuration C; broken line: configuration A; circle: smoothed protocol; cross: original protocol. Error bars show the 95% confidence limits.

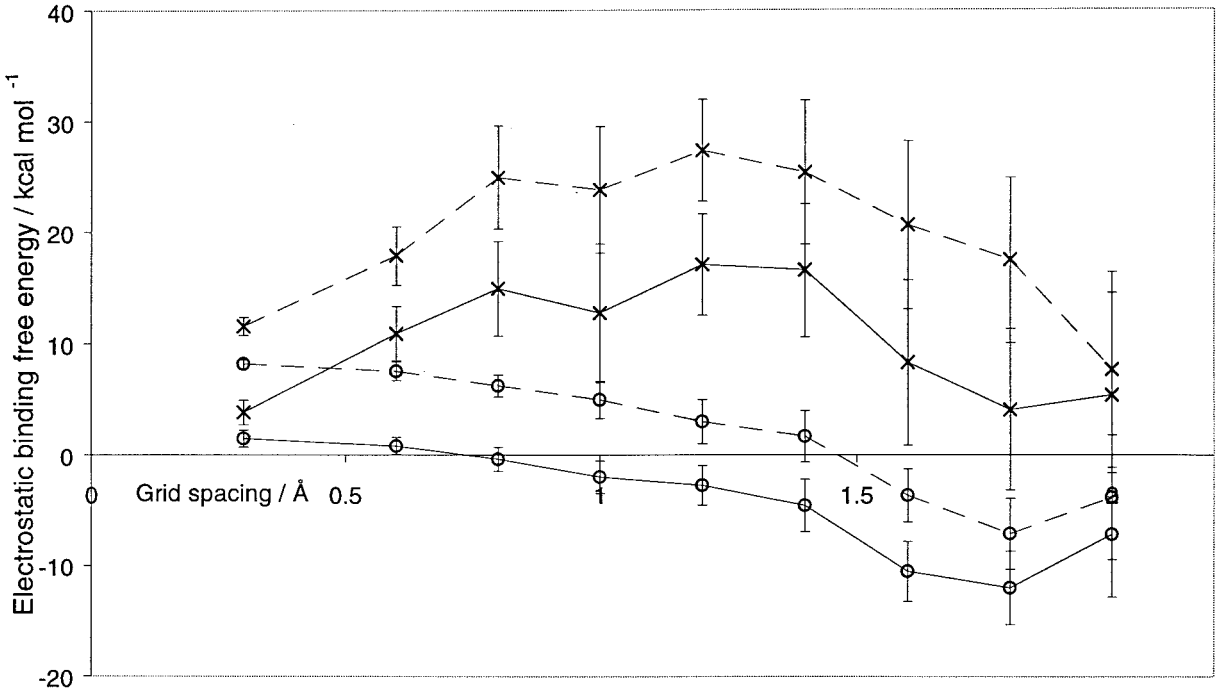


Figure 12. Grid resolution dependence of the electrostatic binding free energy of the enzyme-inhibitor 3 complex. Solid line: configuration C; broken line: configuration A; circle: smoothed protocol; cross: original protocol. Error bars show the 95% confidence limits.

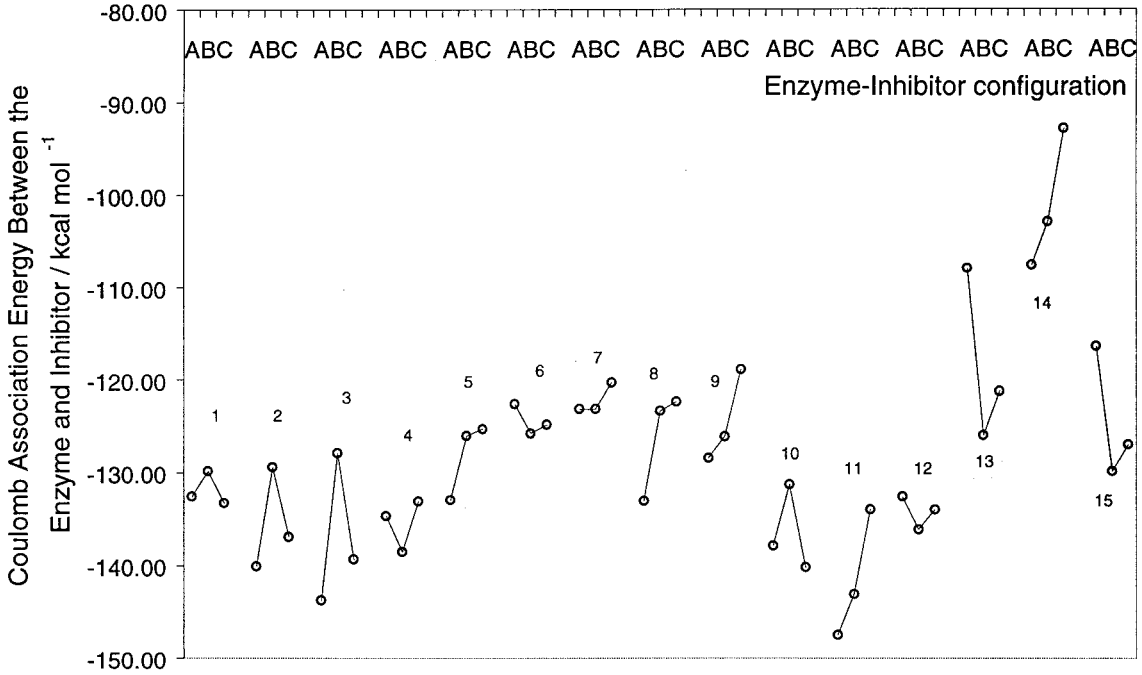


Figure 13. The Coulomb association energy of the fifteen Neuraminidase-inhibitor complexes.

of the electrostatic binding free energy to be calculated. This is shown for the enzyme-inhibitor 3 complex (Figure 12). This graph shows that the configurational dependence of the binding free energy is independent of the grid spacing or protocol of the calculation. Thus configurational dependence is not an artefact of the finite difference procedure, and thus must be an effect within the PB equation itself. This graph also shows that the electrostatic binding free energy calculated via the smoothed protocol is much more stable than that calculated via the original protocol. The smoothed protocol produces stable binding free energies at grid spacings up to 0.6 Å, and a reasonable approximation at spacings up to 0.8 Å. The original protocol only seems to converge by 0.3 Å, which raises doubts regarding its usefulness for calculating absolute binding or hydration free energies. Using random rotation with the smoothed protocol is also seen to produce highly precise electrostatic binding free energies, with errors of less than 1 kcal mol^{-1} calculated for grid spacings up to 0.8 Å. The errors calculated using rotational averaging with the original protocol for the same grid spacing are approximately 4 kcal mol^{-1}. These experiments show that a reasonably accurate and precise single-structure electrostatic binding free energy may be obtained via the FDPB equation, using 30 random rotations, the smoothed protocol, and grid spacings of at most 0.8 Å. If the original protocol were used, then accurate and precise binding free energies may only be obtained on grids with spacings less than 0.3 Å.

Configurational dependence of the Coulomb association energy

The previous calculations have shown that the electrostatic hydration free energy components of the binding free energy are all configurationally dependent. The only other component of the electrostatic binding free energy is the Coulomb association energy of the enzyme-inhibitor complex. The configurational dependence of the Coulomb association energy is large (Figure 13), with a mean unsigned dependence of 6.7 kcal mol^{-1}. Since this is quite strong, it may be argued that it could almost entirely account for the configurational dependence of the electrostatic binding free energies shown in Figure 3. To discover whether this is the case, the Coulomb association energy was subtracted from the electrostatic binding free energies. This left an energy that was entirely constructed from FDPB calculated electrostatic hydration free energies (Figure 14). It is seen that even this energy shows strong configurational dependence, showing that this

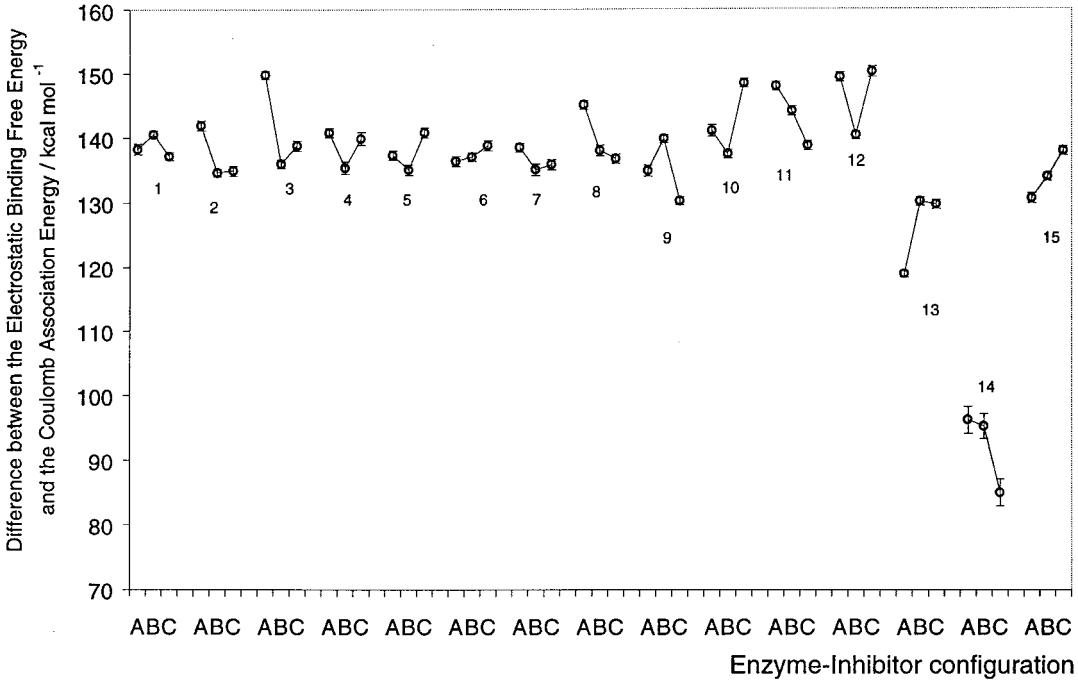


Figure 14. The difference between the electrostatic binding free energy and the Coulomb association energy of the fifteen Neuraminidase-inhibitor complexes.

effect features heavily in all components of the binding calculation.

Conclusion

In this paper, the assumption that an electrostatic binding or hydration free energy could be obtained from a single configuration of a solute molecule, using the FDPB method, has been investigated. From the experiments that were performed, six important conclusions can be drawn;

(1) For the Neuraminidase system, the electrostatic binding and hydration free energies calculated were highly sensitive to the exact configuration of the solute. Very small configurational changes, of the sort that occur due to thermal motion during a Monte Carlo simulation, were enough to change the hydration or binding free energies by typically 1–4 kcal mol^{-1}. In the worst case of inhibitor 3, a 0.56 Å RMS structural deviation was enough to affect the electrostatic hydration energy by 6.9 kcal mol^{-1}.

(2) This configurational dependence is not an artefact of the finite-difference method of solving the PB equation. This was shown by the fact that the configurational dependence was stable across all grid spacings, and across both protocols for assigning charge and dielectric to the finite-difference grid. Thus configurational dependence is a real effect within the PB equation. This is supported by the high configurational dependence observed in the Coulomb association energy between the inhibitors and the enzyme; since such a high configurational dependence was observed in this Coulomb's law calculation, it is unsurprising that it is also observed in the PB calculation.

(3) The configurational dependence is not a consequence of the parameters in the FDPB equation; several key parameters were varied and the configurational dependence was shown to be stable. Additionally, the absolute FDPB hydration free energies are very sensitive to the exact values of these parameters. By changing the interior dielectric constant, the probe sphere radius, or the polar hydrogen radius, the absolute hydration free energy of inhibitor 3 varies widely. However the relative hydration free energy between the two configurations of inhibitor 3 stayed roughly constant. This demonstrates that only relative FDPB energies may be reliably quoted.

(4) The use of the smoothed protocol, with random rotational averaging, dramatically reduces the grid effects in the FDPB calculation. Use of these two techniques has been shown to allow reasonable estimates of electrostatic hydration or binding free energies at grid spacings of up to 0.8 Å. The results also discourage the use of the original protocol, as this needed very fine grids to produced converged free energies.

(5) The configurational dependence makes the ranking of the affinities of the fifteen inhibitors for Neuraminidase even qualitatively difficult. All components of the calculation show strong configurational dependence, and these combine to give configurational dependencies in the electrostatic binding free energies of typically 5–10 kcal mol^{-1}. Since this is the same order of magnitude as the binding free energies, it becomes almost impossible to rank the inhibitors. This result suggests that the only route to reliable relative binding free energies is to generate and average FDPB free energies over many configurations. It is unlikely that any other terms that determine binding, such as solvent entropy, would counter out the configurational dependence, as the difference in surface area of different conformations of the solutes is quite small (of the order of 6 $Å^2$). Assuming a linear dependence of the non-polar and entropic free energies, with a surface area parameter of 0.025 kcal mol^{-1} $Å^{-2}$ [22], this corresponds to configurational dependences of around 0.15 kcal mol^{-1}.

(6) This work has shown that the electrostatic hydration free energies are strongly configurationally dependent. Thus, arguably all single-structure FDPB based calculations should include a check to see whether configurational dependence is present in the system and property that is being investigated.

In conclusion, this paper has demonstrated that configurational dependence is a real effect within the PB equation, as applied to the Neuraminidase-inhibitor system, and its influence on binding and hydration free energies can be quite strong. In the light of this finding, applications of the PB equation should ideally include a check for configurational dependence. Thus just as molecular mechanics energies are very sensitive to configuration, and single-structure values are not typically used to score binding free energies, single FDPB energies should be treated with the same caution.

Acknowledgements

The authors would like to thank the BBSRC, Celltech Chiroscience, and the Royal Society for supporting this work, and Dr I. D. Wall for the provision of the enzyme and inhibitor structures.
Dr J.W. Essex is a Royal Society University Research Fellow.

References

1. Shen, J., J. Med. Chem., 40 (1997) 2953.
2. Misra, V.K., Hecht, J.L., Yang, A.S. and Honig, B., Biophys. J., 75 (1998) 2262.
3. Mohan, V., Davis, M.E., McCammon, J.A. and Pettitt, B.M., J. Phys. Chem., 96 (1992) 6428.
4. Davis, M.E. and McCammon, J.A., Chem. Rev., 90 (1990) 509.
5. Honig, B., Sharp, K. and Yang, A.S., J. Phys. Chem., 97 (1993) 1101.
6. Eriksson, M.A.L., Pitera, J. and Kollman, P.A., J. Med. Chem., 42 (1999) 868.
7. Sitkoff, D., Sharp, K.A. and Honig, B., J. Phys. Chem., 98 (1994) 1978.
8. Srinivasan, J., Cheatham, T.E., Cieplak, P., Kollman, P.A. and Case, D.A., J. Am. Chem. Soc., 120 (1998) 9401.
9. Still, W.C., Tempczyk, A., Hawley, R.C. and Hendrickson, T., J. Am. Chem. Soc., 112 (1990) 6127.
10. Massova, I. and Kollman, P.A., J. Am. Chem. Soc. 121 (1999) 8133.
11. Vlijmen, H.W.T., Schaefer, M. and Karplus, M., Proteins, 33 (1998) 145.
12. Wall, I.D., Leach, A.R., Salt, D.W., Ford, M.G., Essex, J.W., J. Med. Chem., 42 (1999) 5142.
13. Madura, J.D., Davis, M.D., Gilson, M.K., Wade, R.C., Luty, B.A. and McCammon, J.A., In Lipkowitz, K.B., Boyd, D.B. (Eds.), Reviews in Computational Chemistry V, VCH Publishers Inc., New York, NY, 1994, p. 229.
14. Shen, J. and Wendoloski, J., J. Comp. Chem., 17 (1996) 350.
15. Gilson, M.K., Sharp, K.A. and Honig, B.H., J. Comp. Chem., 9 (1998) 327.
16. Bruccoleri, R.E., Novotny, J. and Davis, M.E., J. Comp. Chem., 18 (1997) 268.
17. Simonson, T. and Brünger, A.T., J. Phys. Chem., 98 (1994) 4683.
18. Jorgensen, W.L., Maxwell, D.S. and Tirado-Rives, J., J. Am. Chem. Soc., 118 (1996) 11225.
19. Jorgensen, W.L. MCPRO 1.4. Yale University, New Haven, CT (1996).
20. Davis, M.E., Madura, J.D., Luty, B.A. and McCammon, J.A., Comp. Phys. Comm., 62 (1991) 187.
21. Fogolari, F., Zuccato, P., Esposito, G. and Viglino, P., Biophys. J., 76 (1999) 1.
22. Hunenberger, P.H., Helms, V., Narayana, N., Taylor, S.S. and McCammon, J.A., Biochemistry, 38 (1999) 2358.