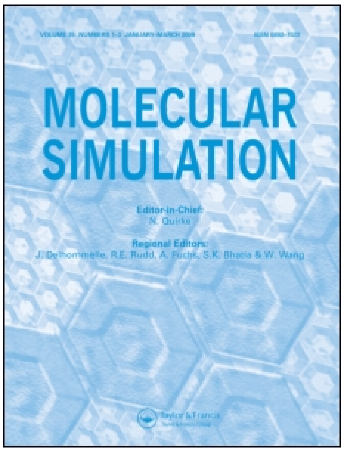


Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Free Energy Perturbations in Ribonuclease T_1 Substrate Binding. A Study of the Influence of Simulation Length, Internal Degrees of Freedom and Structure in Free Energy Perturbations

Arne Elofsson[a]; Lennart Nilsson[b]
[a] Department of Medical Biophysics, Karolinska Institutet, Stockholm, Sweden [b] Center for Structural Biochemistry Novum Research Park, Karolinska Institutet, Huddinge, Sweden

To cite this Article Elofsson, Arne and Nilsson, Lennart(1993) 'Free Energy Perturbations in Ribonuclease T_1 Substrate Binding. A Study of the Influence of Simulation Length, Internal Degrees of Freedom and Structure in Free Energy Perturbations', Molecular Simulation, 10: 2, 255 — 276

To link to this Article: DOI: 10.1080/08927029308022168

URL: http://dx.doi.org/10.1080/08927029308022168

Molecular Simulation, 1993, Vol. 10(2-6), pp. 255-276
Reprints available directly from the publisher
Photocopying permitted by license only

FREE ENERGY PERTURBATIONS IN RIBONUCLEASE T_1 SUBSTRATE BINDING. A STUDY OF THE INFLUENCE OF SIMULATION LENGTH, INTERNAL DEGREES OF FREEDOM AND STRUCTURE IN FREE ENERGY PERTURBATIONS

ARNE ELOFSSON[1] and LENNART NILSSON[2*]

[1]*Department of Medical Biophysics Karolinska Institutet S-104 01 Stockholm, Sweden*
[2]*Karolinska Institutet Center for Structural Biochemistry Novum Research Park S-141 57 Huddinge, Sweden*

(Received January 1993, accepted January 1993)

We have studied the reliability of free energy perturbation calculations with respect to simulation protocol and simulation length in a real biological system, the binding of two different ligands to wildtype Ribonuclease T_1 (RNT1) and to a mutant of RNT1 with Glu-46 replaced by Gln (RNT1-Gln46). The binding of the natural substrate 3′GMP has been compared with the binding of a fluorescent probe, 2-aminopurine 3′mono phosphate (2AP3′MP). These simulations predict that the mutant binds 2AP3′ MP better than 3′GMP. Four complete free energy perturbations were performed that form a closed loop of four free energy differences, which should sum up to zero. This could be used as a tool for searching for systematic errors that are not detected by standard forward ↔ backward perturbations. The perturbation between 2AP3′MP and 3′GMP is quite straightforward and similar to what has been done by other groups. The perturbation between Glu46 and Gln46 is much more complex, involving as many as twelve atoms and a change of charge. This perturbation needs much longer simulation time, 500-600 ps, than used in free energy perturbations before. The increased simulation time is needed both to reach an equilibrium and to include several phases of fluctuations of the observed parameters in the production run. The extremely long simulation time is not such a severe problem as much of the work might be done on several different machines in parallel and cheap workstations are excellent for these calculations. Problems may also occur with values of the coupling parameter λ close to 0 or 1, due to the high mobility of atoms as well as insertion/deletion in a previously unoccupied space involved in the perturbation.

KEY WORDS: Ribonuclease T_1, free energy perturbations, molecular dynamics, protein-ligand interaction, drug design, computer simulations

INTRODUCTION

Ribonuclease T_1 from the fungus *Aspergillus oryzae*, which has been crystallised in the presence of r2′GMP, r3′GMP (coordinates not available, when this study started) and several other substrates, offers an ideal system for the study of protein action [1-14]. RNT1 specifically cleaves single-stranded RNA on the 3′ side of guanosine residues via a cyclic guanosine 2′,3′-phosphate intermediate that is hydrolysed to yield a terminal 3′guanosine. It is a small enzyme with a rather good selectivity, it is well studied, has been cloned and point mutants are available. Biophysical studies of the kinetics and thermodynamics of the cleavage reaction would be greatly facilitated if the enzyme could work on a fluorescent substrate, such as 2-aminopurine (2AP3′MP). 2AP3′MP is a fluorescent ribonucleotide, very useful for the study of DNA and RNA dynamics by time resolved fluorescence spectroscopy [15].

Free energy perturbation calculations, as reviewed by Beveridge & DiCapua [16], have been successfully performed on small systems, like the difference in free energy of solvation between ethanol and methanol [17, 18] as well as on some biologically relevant systems. Some examples of the latter include studies of: RNT1 bound to either 2′AMP or 2′GMP [1, 2]; netropsin bound to DNA [19]; different metallo enzymes [20].

The difficulties inherent in these calculations have also been discussed previously. Mark *et al.* have described the problem of sampling the full conformational space [21] with a normal MD simulation, and studies of the properties of inter- and intramolecular energy terms by Pearlman [22] indicate that problems may occur when simulating free energy differences that are due to structural rather than electrostatic differences. Even though there are some cases where the molecular dynamics simulation, used for the free energy perturbations, seems to describe the conformational space accurately enough even for such a complicated system as a protein, one of the major problems with free energy perturbation studies on protein systems is that the simulations can not sample the total conformational space. By increasing the simulation time a larger part of the conformational space can be sampled [23]. A major problem also seems to occur when the perturbed atoms are too flexible, i.e. not placed in a rigid environment. We show below that the change from 3′GMP to 2AP3′MP is such a benign case, while the perturbation Glu to Gln is much more complicated, and some different ways of treating the problems are tested. Two different techniques of describing the free energy terms of the simulation and describing the ensemble used for sampling have been used in this study. These approaches were examined by using a closed loop of four free energy simulations that should yield a total free energy change of zero (figure 1). We performed two such closed loops of free energy simulations with different simulation approaches and further performed several tests to examine the importance of starting structure and simulation method on one or several of these four steps. First the difference between 3′GMP and 2AP3′MP binding to RNT1-Glu46 was calculated, and it was predicted that 2AP3′MP would bind less well than 3′GMP due to the loss of two hydrogen bonds between RNT1 and 3′GMP. In the second step the binding of 2AP3′MP to RNT1-Glu46 was compared with RNT1-Gln46. The third step compared the binding of 3′GMP and 2AP3′MP to RNT1-Gln46. We predicted that a hydrogen bond would be formed between Gln46 and 2AP3′MP, favouring the binding of 2AP3′MP over 3′GMP. The last simulation compared the binding of 3′GMP to RNT1-Gln46 with the binding of 3′GMP to RNT1-Gln46.

Free energy calculations were performed on two different kinds of “alchemical” reactions. In steps 1 and 3 a guanosine base was changed into a 2-aminopurine, i.e. an oxygen double bonded to carbon 6 (standard nomenclature) was deleted, resulting in a double bond between carbon 6 and nitrogen 1 and the movement of a hydrogen from nitrogen 1 to carbon 6 (figure 3). The typical free energy difference for this transition is about 6 kcal/mol in this study and it should be noted that this is the free energy difference and not the relative free energy difference. In

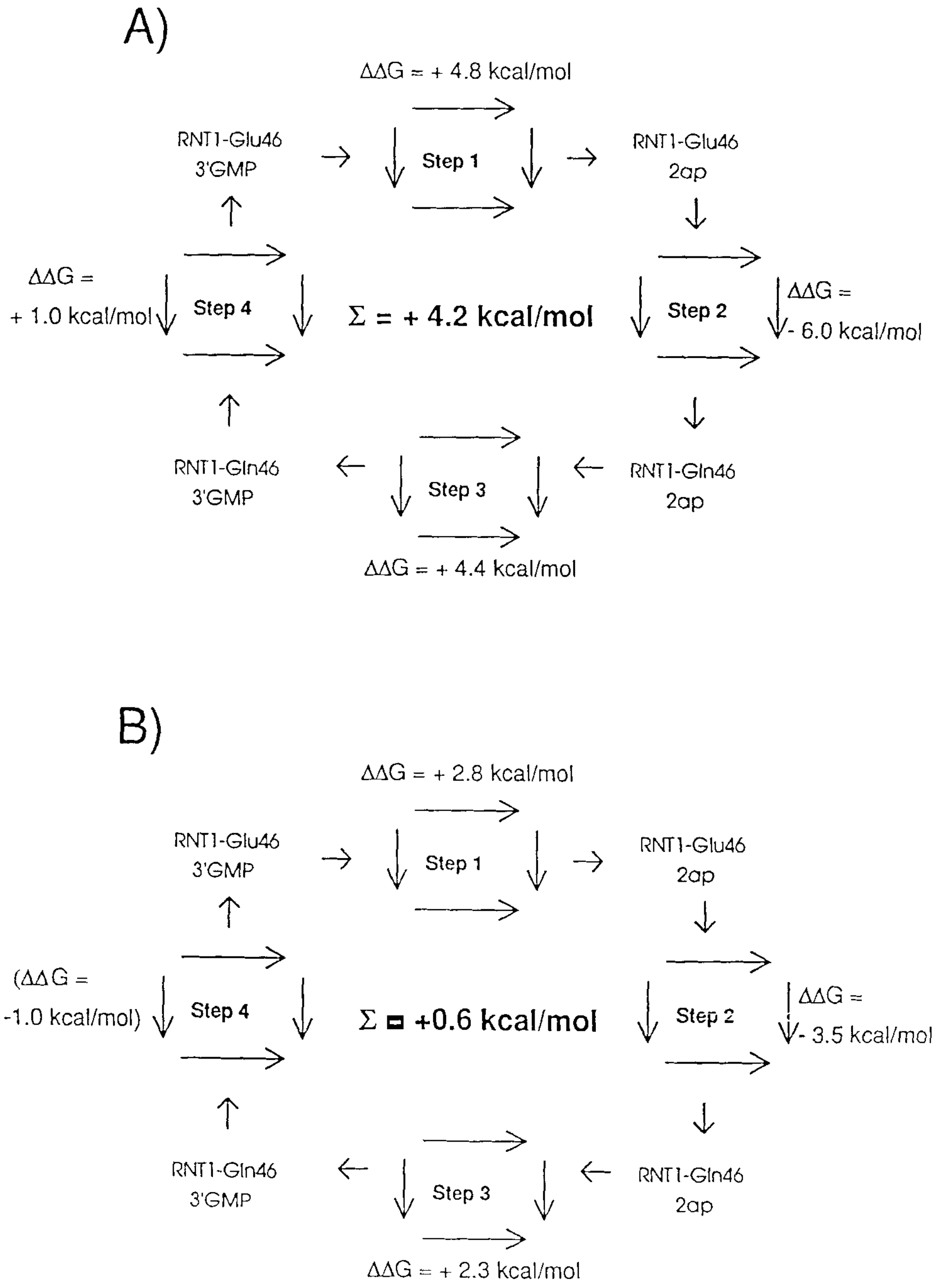


Figure 1a and b

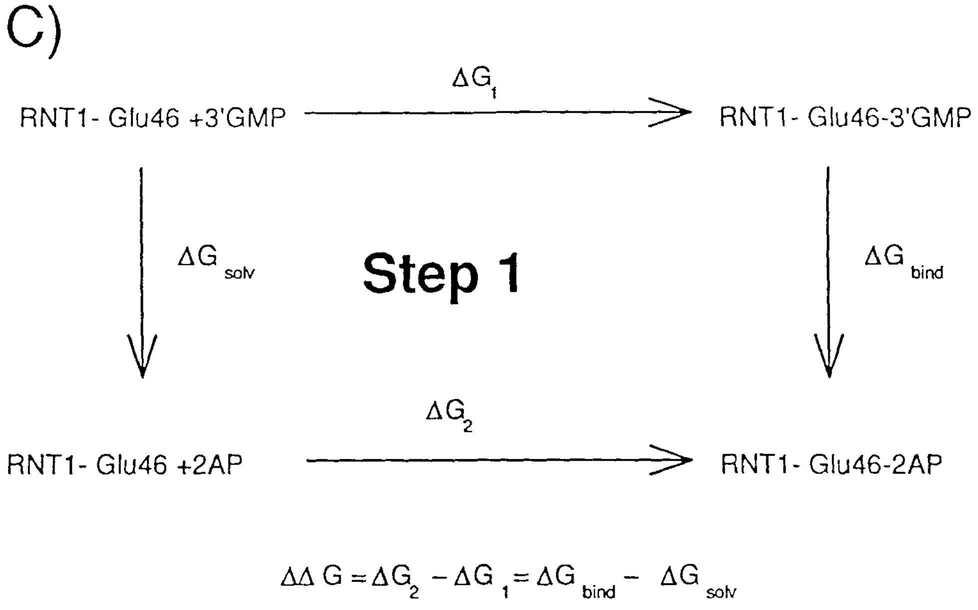


Figure 1 Summary of $\Delta\Delta G$ values obtained with A) the four perturbations with the FULL method, and with B) the DONT method. C) The relationship between $\Delta\Delta G$, ΔG_{bind}, ΔG_{solv}, and ΔG_2 for step 1.

steps 2 and 4 we calculated the free energy difference for the change of a glutamate (Glu) to a glutamine (Gln), i.e. we change an oxygen to an amine group. The free energy difference, about 40 kcal/mol in this study, of this change is much larger than in the first transition. The higher free energy difference creates problems that will be discussed below. The change from Glu to Gln also includes a change of charge, and involves many atoms, further complicating the situation.

METHODS

Theory

The theory ot free energy simulations has been fully described in several books and articles so we, will only cover it briefly [16, 22, 24]. Provided that we have an energy function depending on all coordinates $\bar{R}$, and on a coupling parameter λ:

$$H = H(\bar{R}, \lambda) \tag{1}$$

The free energy in an ensemble is:

$$G(\lambda) = -k_B T \ln \left\{ \frac{\int e^{-H(\bar{R},\lambda)/k_B T} d\bar{R}}{N!\Lambda^{3N}} \right\} \tag{2}$$

where k_B is Boltzmann's constant, T is the absolute temperature, N is the number of particles, and Λ is a function of the temperature. From equation 2 the following equation is obtained:

$$G(\lambda) - G(\lambda + \Delta\lambda) = -k_B T \ln \left\langle e^{\frac{H(\bar{R},\lambda) - H(\bar{R},\lambda + \Delta\lambda)}{k_B T}} \right\rangle_{\lambda} \tag{3}$$

here $\langle \ldots \rangle_{\lambda}$ indicates the ensemble or time average at an intermediate position along the conversion pathway characterised by the coupling parameter λ. The interaction Hamiltonian H is $H' + (1 - \lambda)H_0 + \lambda H_1$, where H_0 is the Hamiltonian for the "reactant" atoms, the atoms that are present in the starting structure but not in the final structure, and H_1 is the Hamiltonian for the "product" atoms, the atoms present in the final structure but not in the starting structure; H' is the Hamiltonian for atoms present which do not change during the conversion.

What kind of free energy are we calculating? This depends on the type of ensemble used in the simulation. If the volume and temperature are kept constant it is the Helmholtz free energy difference ($\Delta\Delta A$) that we are calculating. If instead the pressure and temperature are fixed, we are calculating the Gibbs free energy ($\Delta\Delta G$). During a stochastic deformable boundary simulation [25] volume and temperature are approximately constant [26], i.e. they are allowed to vary within a certain range. In practice, however, if the system is large enough it does not matter what kind of ensemble is used.

From the difference $\Delta\Delta G = \Delta G_{\mathrm{bind}} - \Delta G_{\mathrm{solv}}$ we can calculate the difference in binding strength between the two ligands (see figure 2 for a complete description of our simulated systems).

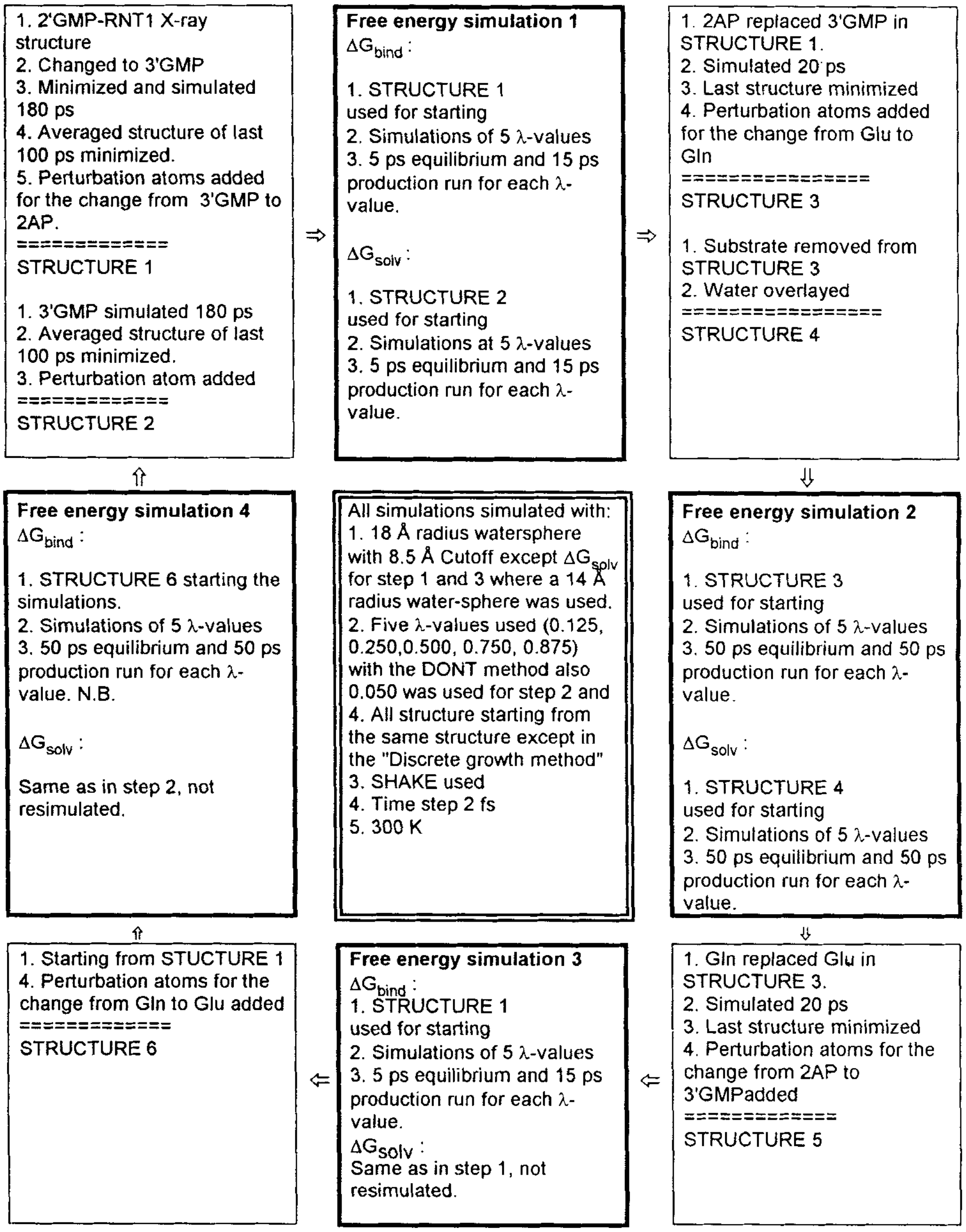
Calculation of free energy changes

When simulating the free energy difference between two ligands binding to a protein, one calculates the difference between the two states by changing from one of the substrates to the other in a stepwise manner, gradually adding new atoms and deleting atoms. The strength with which these atoms are allowed to interact with the surroundings depends on a parameter λ. We call the atoms that are added to the system product atoms and the ones that are deleted reactant atoms.

We calculate the free energy changes according to equation 3. Normally simulations have been performed at five λ-values (0.125, 0.250, 0.500, 0,750 and 0.875). From these simulations the free energies associated with perturbations to intermediate λ-values were calculated using "double-wide sampling", where each simulation is used to obtain two pieces of the total ΔG: from $\lambda = 0.125$ the free energy going to $\lambda = 0.000$ and to $\lambda = 0.1875$ is calculated, and so on. Molecular dynamics was used to calculate the time average in equation 3, from which ΔG between states A ($\lambda = 0$) and B ($\lambda = 1$) is evaluated by:

$$\Delta G = \sum_{\lambda=0}^{\lambda=1} (G(\lambda) - G(\lambda + \Delta\lambda)) \tag{4}$$

The calculations were performed using CHARMM [25] with free energy calculations implemented as described by Fleischman and Brooks [18]. We used standard CHARMM19 parameters not involving any specific energy terms for hydrogen bonds, and only polar hydrogens.



1. 2'GMP-RNT1 X-ray structure
2. Changed to 3'GMP
3. Minimized and simulated 180 ps
4. Averaged structure of last 100 ps minimized.
5. Perturbation atoms added for the change from 3'GMP to 2AP.

=============

STRUCTURE 1

1. 3'GMP simulated 180 ps
2. Averaged structure of last 100 ps minimized.
3. Perturbation atom added

=============

STRUCTURE 2

⇒

Free energy simulation 1

ΔG_{bind} :

1. STRUCTURE 1 used for starting
2. Simulations of 5 λ-values
3. 5 ps equilibrium and 15 ps production run for each λ-value.

ΔG_{solv} :

1. STRUCTURE 2 used for starting
2. Simulations at 5 λ-values
3. 5 ps equilibrium and 15 ps production run for each λ-value.

⇒

1. 2AP replaced 3'GMP in STRUCTURE 1.
2. Simulated 20 ps
3. Last structure minimized
4. Perturbation atoms added for the change from Glu to Gln

================

STRUCTURE 3

1. Substrate removed from STRUCTURE 3
2. Water overlayed

=================

STRUCTURE 4

⇑

Free energy simulation 4

ΔG_{bind} :

1. STRUCTURE 6 starting the simulations.
2. Simulations of 5 λ-values
3. 50 ps equilibrium and 50 ps production run for each λ-value. N.B.

ΔG_{solv} :

Same as in step 2, not resimulated.

All simulations simulated with:
1. 18 Å radius watersphere with 8.5 Å Cutoff except ΔG_{solv} for step 1 and 3 where a 14 Å radius water-sphere was used.
2. Five λ-values used (0.125, 0.250,0.500, 0.750, 0.875) with the DONT method also 0.050 was used for step 2 and
4. All structure starting from the same structure except in the "Discrete growth method"
3. SHAKE used
4. Time step 2 fs
5. 300 K

⇓

Free energy simulation 2

ΔG_{bind} :

1. STRUCTURE 3 used for starting
2. Simulations of 5 λ-values
3. 50 ps equilibrium and 50 ps production run for each λ-value.

ΔG_{solv} :

1. STRUCTURE 4 used for starting
2. Simulations of 5 λ-values
3. 50 ps equilibrium and 50 ps production run for each λ-value.

⇑

1. Starting from STUCTURE 1
4. Perturbation atoms for the change from Gln to Glu added

============

STRUCTURE 6

⇐

Free energy simulation 3

ΔG_{bind} :
1. STRUCTURE 1 used for starting
2. Simulations of 5 λ-values
3. 5 ps equilibrium and 15 ps production run for each λ-value.

ΔG_{solv} :
Same as in step 1, not resimulated.

⇓

⇐

1. Gln replaced Glu in STRUCTURE 3.
2. Simulated 20 ps
3. Last structure minimized
4. Perturbation atoms for the change from 2AP to 3'GMPadded

============

STRUCTURE 5

Figure 2 The stimulation described in more detail. The thin lined corner boxes symbolise structures and the thick line boxes between tham are descriptions of the free energy simulations. The double framed box shows some data about all simulations.

For each step in the cycle, the free energy difference, between state A and state B, for the complex of enzyme and ligand was calculated (ΔG_{bind}) as well as the free energy difference for the corresponding change of the enzyme/ligand in aqueous surroundings (ΔG_{solv}). The difference $\Delta\Delta G(\Delta G_{\text{bind}} - \Delta G_{\text{solv}})$ is equal to the difference in free energy of binding between the two forms of enzyme/ligand in aqueous solution (see figure 1). Only two calculations of ΔG_{solv} had to be performed, as the free energy difference for the change from 3′GMP to 2AP3′MP is equal to the difference between 2AP3′MP and 3′GMP but with opposite sign. All four combinations of ΔG_{bind} had to be calculated.

Ribonuclease T_I simulation method

The simulation procedure is shown in figure 2. The calculations started from the RNT1: 2′GMP inhibitor complex, in its initial conformation taken from the X-ray structure determined by Saenger to 1.9 Å resolution [6]. With the help of a graphical workstation the phosphate group was moved from the 2′ to the 3′ position and positioned without any bad van der Waals contacts with the protein. An 18 Å radius sphere of TIP3P [27] water was placed around Tyr-42 OH in this structure, enclosing the active site and the whole ligand. The simulations were performed in a similar way as previously described [28–31] with Langevin dynamics on all atoms within 2 Å from the surface of the water sphere and standard Verlet dynamics on the remaining waters and protein atoms. The histidines 27, 40 and 92 were designated to be in a protonated +1 charge state in accordance with experimental evidence [6, 7]. This structure that consists of about 2700 atoms were simulated during 180 ps, with a time step of 0.002 ps and with the SHAKE routine [32] applied to keep all bond lengths involving hydrogens fixed. A nonbonded cutoff at 8.0 Å was applied with a shifting function to smooth out the energies at 8.0 Å. The temperature was kept at 300 K by a allowing a temperature window of ±10 K. An averaged structure from the last 100 ps of the simulation was used as the start for the free energy simulations after 100 steps of ABNR minimization [25] to relax bad structures that are a result from the averaging.

When simulating the free energy perturbation for the "change" from 3′GMP to 2AP3′MP, the O6, C6, N1 and H1 atoms in 3′GMP are changed to the N1 and C6 atoms of 2AP3′MP. There were no hydrogens attached to these atoms as only polar hydrogens were used. The first free energy calculation, calculating ΔG_{bind} between 2AP3′MP and 3′GMP binding to Ribonuclease T_I-Glu-46 (1_{bind}) was simulated directly from this structure, using five 20 ps free energy runs with $\lambda = 0.125, 0.250, 0.500, 0.750, 0.875$. The first 5 ps were used to relax the structure for a specific λ-value. The last 15 ps of each simulation were used for calculating the free energies. We monitored the interaction energy for the product and reactant atoms separately and as long as there was no drift, we assumed that the system was in equilibrium. The error estimates were calculated by separating the trajectory into bins of 100 steps each and then calculating the standard deviation of the free energy change. In the next simulation (1_{solv}), 3′GMP in the same configuration as when in the active site, was placed in a water sphere of radius 14 Å where it was simulated for 180 ps, followed by energy minimisation of the average structure. Starting from this minimized average structure the free energy difference was calculated in the same way as described above yielding the ΔG_{solv}. The mutation from Glu46 to Gln46 changes all the atoms in residue 46 from C_β and out to the atoms of the Gln sidechain. This was the minimal number of atoms that had to be changed as C_β is not the same atom type in Glu and Gln, in the CHARMM parameter set PARAM19 [25]. To include more atoms would increase the possible structural difference between Glu46 and Gln46, but this would also increase the problems with sampling the configuration space. 2AP3'MP was built from the coordinates of 3'GMP in complex with RNT1 followed by a 20 ps simulation, of which the last coordinate set was used as a start for the free energy simulations (2_{bind}). This simulation, calculating ΔG for 2AP3'MP binding to RNT1-Glu46 compared to RNT1-Gln46 were performed as in step 1_{bind}, yielding ΔG_{bind}. When simulating this transition a problem occurred as the free energy change between $\lambda = 0$ and $\lambda = 0.125$ was larger, around 20 kcal/mol, than could be accepted, since it may lead to inefficient sampling when calculating the average in equation 3. When a λ value lower than 0.125 was used there were large motions that created errors in the SHAKE algorithm [32], due to the weak interaction between the amino group and its covalently bound neighbouring atoms. Several different approaches for solving this were examined and will be discussed below. The free energy simulation for the free energy difference between RNT1-Glu46 and RNT1-Gln46 without any ligand bound to it (2_{solv}) was performed from the same structure as 2_{bind} with water molecules replacing the ligand. The water was placed in a similar way as used for placing the water sphere around Ribonuclease T_1 with substrate bound to it, i.e. a water sphere of 18 Å was placed over the protein and all waters that were closer than 2.8 Å from a protein heavy atom were deleted. This procedure was repeated four times with the water sphere rotated in between.

The starting coordinates for the next free energy simulation (3_{bind}) were obtained from the starting coordinates for 2_{bind}, where RNT1-Glu46 was changed to RNT1 -Gln46. This complex was simulated for 20 ps and the last coordinate set was used for starting the free energy simulations. ΔG_{solv} in the third step (3_{solv}) is the same as in step 1_{solv} but with the opposite sign, and therefore was not resimulated. The fourth step of the free energy was started from the end structure of step 1_{bind}, with 3'GMP bound to RNT1-Glu46, which was then changed to RNT1-Gln46. The RNT1-Gln46 coordinate set had to be minimized (100 steps of ABNR [25]), due to a very close contact between Phe48 and the amide group of Gln46. The ΔG_{solv} (step 4_{solv}) is the same as in step 2_{solv} but with opposite sign and was not resimulated.

Different simulation methods

Different approaches were used for comparing the behaviour of the free energy perturbation in this system.

1. All four $\Delta\Delta G$ pairs were simulated with two simulation methods: one, referred to as the "FULL method" below, were the full Hamiltonian for the atoms involved was a linear combination of the Hamiltonians of reactant and product atoms; and one where we did not include any bond, angle or dihedral terms for the product or reactant atoms in the calculation of the free energy

---

Figure 3 Symbolic description of the goal with the point mutation with A) a part of the natural recognition site (RNT1-Glu46 + 3'GMP) and B) a model of the active site of RNT1-Gln46 + 2AP. The loss of hydrogen bonds when GUA is replaced with 2AP will be partly compensated by the mutation Glu46Gln. [40]

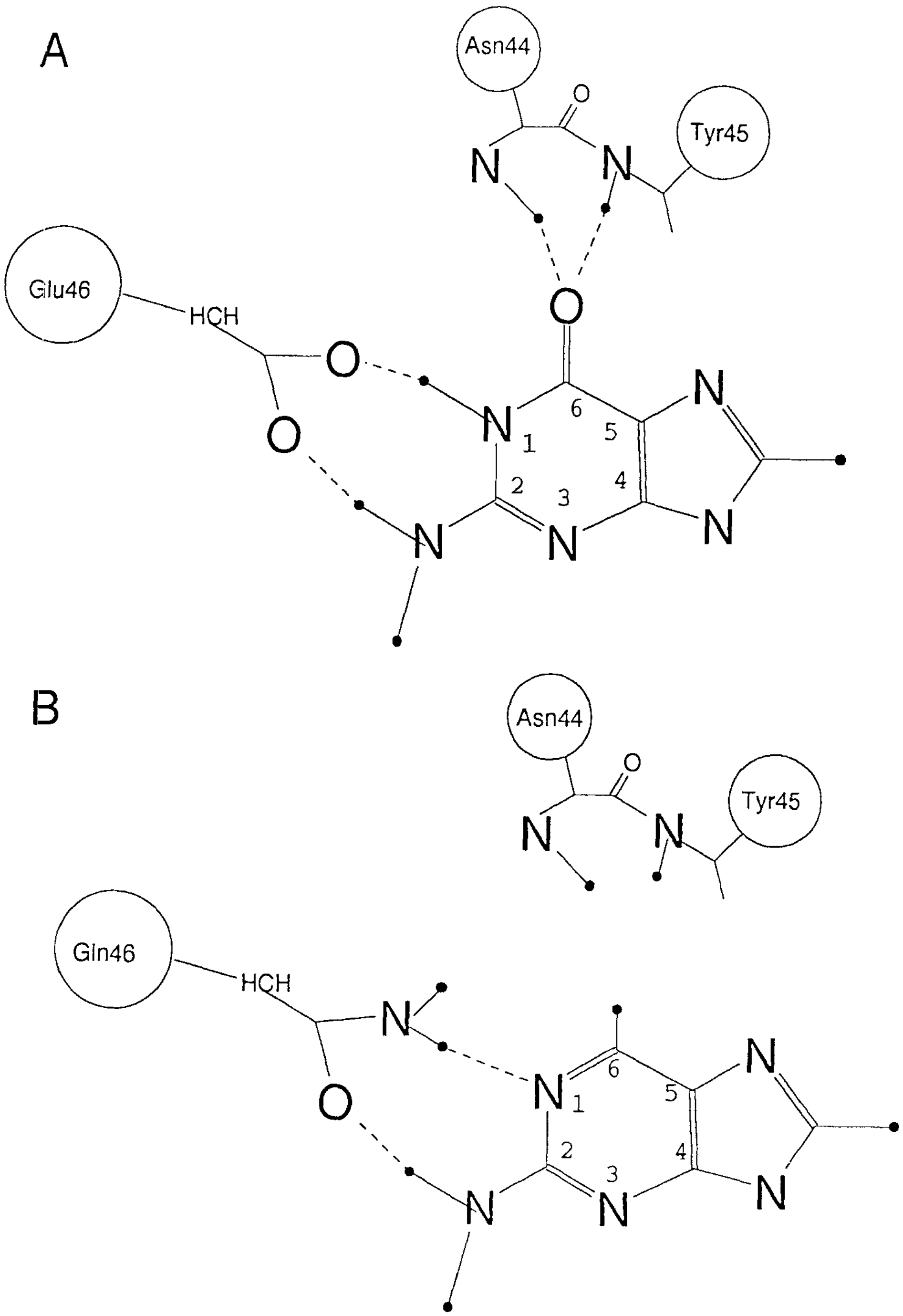

A
Asn44
O
Tyr45
N
N
Glu46
HCH
O
O
O
N
1
2
3
4
5
6
N
N
N
N
B
Asn44
O
Tyr45
N
N
Gln46
HCH
N
O
N
1
2
3
4
5
6
N
N
N
N

Table 1 General results.

		FULL method ΔG	$\Delta\Delta G$	DONT method ΔG	$\Delta\Delta G$	Experimental [35]* $\Delta\Delta G$
Step 1 (20 ps)	ΔG_{bind}	$+5.6 \pm 0.3$	$+4.8 \pm 0.8$	$+6.3 \pm 0.3$	$+2.8 \pm 0.7$	$+0.60$
	ΔG_{solv}	$+0.8 \pm 0.5$		$+3.4 \pm 0.4$		
Step 2 (100 ps)	ΔG_{bind}	$+38.6 \pm 0.6$	-6.0 ± 1.0	$+40.1 \pm 0.5$	-3.5 ± 0.8	-0.80
	ΔG_{solv}	$+44.6 \pm 0.4$		$+43.6 \pm 0.2$		
Step 3 (20 ps)	ΔG_{bind}	$+3.7 \pm 0.3$	$+4.4 \pm 0.8$	-1.1 ± 0.3	$+2.3 \pm 0.5$	-0.09
	ΔG_{solv}	-0.8 ± 0.5		-3.4 ± 0.4		
Step 4 (100 ps)	ΔG_{bind}	-43.7 ± 0.6	$+1.0 \pm 1.0$	-44.6 ± 0.2	-1.0 ± 0.5	$+0.29$
	ΔG_{solv}	-44.6 ± 0.4		-43.6 ± 0.2		
SUM			$+4.2 \pm 3.6$		$+0.6 \pm 2.5$	

* Experimental values are K_m from dinucleotides.

perturbation, i.e. these energy terms were present at normal strength during all simulations. In these simulations $\lambda = 0.050$ could be used in steps 2 and 4 without having problems with SHAKE. This approach, referred to as the "DONT method" below, is used in most free energy perturbations using the algorithm implemented by Fleischman and Brooks [18] in the CHARMM20 program. The standard method in the AMBER/GIBBS [33, 34] programs also does not include these internal energies; see discussion below. The assumption made is that bond stretching and bending terms are equal in ΔG_{bind} and ΔG_{solv} and therefore these terms can be neglected. It has, however, been shown by Pearlman [22] that this assumption is not always true.

2. The transitions 2_{solv} and 2_{bind} were also simulated in such a way that the last coordinates of a run with a given λ-value were used as the starting coordinates for the simulation with the nearest higher λ-value. This will be called the "DISCRETE GROWTH" method. It was also combined both with the "FULL" and the "DONT" methods. Only 20 ps of simulation for each λ-value was used.
3. The influence of starting structure on the free energy calculations was examined by simulating step 2_{bind} also from a structure obtained from the average coordinates from a 180 ps simulation of the complex RNT1-Glu46: 2AP3′MP. i.e. we are simulating the same step with structures that are closer to a "good" structure for either the product or reactant state.

RESULTS AND DISCUSSION

General results and comparison with experiments

The main results from the free energy simulations with the DONT and the FULL method, shown in table 1 and figure 1, have been presented in detail elsewhere [35]. It should be noted that the experimental results are from experiments with dinucleotides and not with the 3′ mononucleotides used in the simulation ([35] and references therein). The reason for using dinucleotides in the experiment is that there are two binding sites [3] which could be accessed when mononucleotides are used. Only the wildtype binding site is studied in the simulations and therefore we think that the predicted binding strengths will agree better with studies on the real

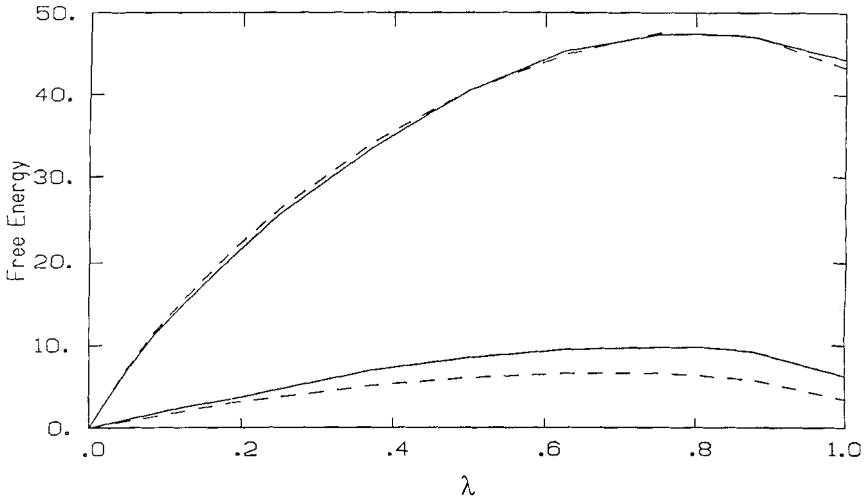


Figure 4 Description of the development of ΔG with λ. The two lines at top are from typical energy changes in steps 2 and 4. These particular curves are from calculations of step 4. The solid line is ΔG_{bind} and the dotted from ΔG_{solv}. Note that the values for step 4 are multiplied by a factor of -1, as the actual change is made from Glu and Gln. The two lower lines, from step 1 with DONT, are typical curves from steps 1 and 3.

substrates (single stranded oligonucleotides) than with studies on mononucleotides. It seems reasonable that the free energy difference from experiments and theory might vary in size, but that the relative size is kept between experiments and simulations. The experimental values are not from real binding constants but from K_m. It should also be noted that the experimental free energy difference for step 3 is calculated from the sum of all the other measurements, as all differences have to sum up to zero, and is thereby more uncertain. The error bars from these experiments are in the range of 0.3 kcal/mol [36].

ΔG in the perturbation of Glu to Gln is much larger (40 kcal/mol) than in the perturbation 2AP3′MP to 3′GMP (6 kcal/mol) (see figure 4). The change in free energy between $\lambda = 0.0$ and $\lambda = 0.125$ is about 15 kcal/mol, which may be too large for the method to work. It can be seen in figure 4 that the free energy change with λ is quite linear when λ is less than 0.5, i.e. it would not be necessary to include the $\lambda = 0.050$ in the calculations. This is also shown in table 2 as $\Delta\Delta G$ does not change significantly if $\lambda = 0.050$ is included or not in the calculations. Therefore it seems reasonable to assume that the values obtained with the FULL method are not influenced by the omission of simulations at $\lambda = 0.050$.

The free energy differences for the four steps both from experiments and simulations give a clear picture of the relative binding between the complexes. 2AP3′MP binds less well than 3′GMP to the wildtype enzyme, 2AP3′MP binds better to the mutant than to the wildtype enzyme, the mutant binds better to 2AP3′MP than to 3′GMP, and 3′GMP binds almost as well to the mutant as to the wildtype enzyme. It should be noted that binding strength is not linearly correlated with catalytic speed.

Table 2 Free energy calculations for step 2 and 4 with and without $\lambda = 0.050$ simulation used.

	ΔG_{bind}		ΔG_{solv}		$\Delta\Delta G$	
DONT	With 050	Without 050	With 050	Without 050	With 050	Without 050
Step 2	$+40.1 \pm 0.5$	$+39.1 + 0.6$	$+43.6 \pm 0.2$	$+42.6 \pm 0.3$	-3.5 ± 0.8	-3.5 ± 0.9
Step 4	-44.6 ± 0.2	-43.2 ± 0.6	-43.6 ± 0.2	$+42.6 \pm 0.3$	-1.1 ± 0.4	-0.8 ± 0.8

Small λ-values and internal degrees of freedom for perturbed atoms

If a free energy perturbation shall give an accurate result it is necessary that the difference in λ-values between different simulation step is small enough, that the averaging procedure $\langle \ldots \rangle_\lambda$, using $H(\lambda)$, gives a reasonable sampling also of states corresponding to $H(\lambda + \Delta\lambda)$. A rule of thumb is that the change in free energy (ΔG) only should be a few times $k_B T$ between two λ-values. When the change in ΔG is large it may be necessary to use λ-values close to 0 or 1, which might cause problems. The problems with using low λ-values in the FULL method, are due to that the perturbed atoms may be very loosely bound to the rest of the system, since in this approach the bonding strengths are scaled by λ or $(1 - \lambda)$, and these atoms thereby become more mobile. This is detected either by the SHAKE algorithm or by large fluctuations of the energy eventually resulting in unstable trajectories. One way to deal with this problem is to use the DONT method, which does not include "internal" terms for the perturbed atoms in the λ dependent part of the Hamiltonian, i.e. the "internal" energies are all on full force and do not depend on λ. There might be another fact contributing to the problems with low λ-values in steps 2 and 4. The changes in steps 1 and 3 only involve adding one atom and not a group of atoms. The additions of O6 and H1 to 3'GMP is not difficult as they are placed on a rigid structure, the nucleoside ring. The replacement of Glu by Gln is much more difficult as there are 12 (5 + 7) atoms that are included in the perturbation and they are much more mobile. It would be possible that the atoms belonging to Gln would form a hydrogen bond to 2AP3'MP but that the presence of the Glu pushes 2AP3'MP away so this bond not can be formed. Is this accurate or not? One problem is that we have no intuitive knowledge about how this ensemble should behave, as this is a non realistic system. We have also found that using harmonical constraints to keep the perturbed atoms in place (data not shown) disturbs the free energy difference less than to allow the unrealistic sampling occuring at low λ-values. Another way that might be used is to include $\lambda = 0.000$ in the calculations. Both these methods seem to work in some cases, while in others they create unstable systems, therefore we choose to not include them in our calculations.

We have examined the importance of using $\lambda = 0.050$ in steps 2 and 4 with the DONT method (table 2). As mentioned above the inclusion of $\lambda = 0.050$ hardly affects $\Delta\Delta G$ at all. It seems as the changes in ΔG_{solv} and ΔG_{bind} are equal and therefore cancel. This is encouraging for calculations with the FULL method, where $\lambda = 0.050$ could not be used.

The "DONT" method compared with the "FULL" method

The results from the DONT method together with the FULL method are shown in table 1 and figure 1. It is seen that the sum over the closed circle of four free energy simulations is closer to zero in the DONT method than in the FULL method, seemingly an indication that the DONT method performs better. It could, however, also be a result of all $\Delta\Delta G$ values being roughly half the corresponding values with the FULL method, possibly because something is missing. It was shown in a study by Pearlman [22] that in the AMBER/GIBBS algorithm for calculating ΔG, internal energies in the perturbed group are not always negligible and are difficult to describe correctly.

If we use the DONT method we leave out some terms that might influence the binding, while if we use the FULL method we are not able to use low λ-values. Unfortunately the experiments are not accurate enough (see discussion in [35]) to give values to calibrate the simulations with.

What λ-values should be used

Another problem is that systematic errors might occur as always a higher λ-value is used for sampling the ensemble used for calculating the lowest λ-values and similar for all parts of the trajectory. This can be compensated for, as in the so called slow growth method where λ is continuously changing during the simulation, by averaging simulations backwards and forwards in λ-space. This is, however, not possible in the simulation procedure we have used, since with "double-wide sampling" we normally use $\lambda = 0.125$ for calculating the change from $\lambda = 0.000$ to first $\lambda = 0.125$ and then to $\lambda = 0.1875$, and so on, so even if we do this simulation backwards the same λ-value will be used for the description of the same λ-area. It has been shown by Straatsma [37] that thermodynamical integration can overcome eventual systematic errors of this kind.

In the double-wide sampling used here, simulations over a given set of λ-values λ_i, $i = 1, \ldots N$ are used such that the free energy change from $\lambda_i - (\lambda_i - \lambda_{i-1})/2$ to $\lambda_i + (\lambda_{i+1} - \lambda_i)/2$ is calculated from one run at λ_i. If the λ-values used are close enough it should not matter if the next λ-value was used for calculating the change over the same λ-area. This was tested by using λ_i to calculate the change of free energy from λ_i to λ_{i+1}, except for λ_1 which was used to cover the range from $\lambda = 0$ to λ_2. We call this the high λ-area as all λ values (except $\lambda_1 = 0.125$) are used to calculate the change for a λ-area higher than the value. Similarly the low area was tested. This is testing the same kind of systematic errors as can be detected by forwards-backwards calculations in AMBER/GIBBS. The results are presented in table 3. The high and the low areas give mostly very similar results, with differences less than 1.0 kcal/mol. When using the DONT method, in step 2, a major difference between the low and the high areas is observed. This results in quite dramatic changes of $\Delta\Delta G$, from −7.5 to −3.0 kcal/mol. It is interesting that the high area gives similar results as the normal areas used. This should mean that the area around $\lambda = 0.875$ is more important than the area around $\lambda = 0.125$, although the slope at $\lambda = 0.125$ is much steeper. This could be explained by that the free energy change at low λ-values really is linear and therefore it matters less what λ-values are used.

How to introduce/remove the changed atoms into/from the system?

One methodological difference between the simulations by Hirono [1, 2] and our simulation is in the interactions of the added atoms with the environment when they are first put into the system. When the potential function gradually is changed in AMBER/GIBBS the sizes and bond lengths change gradually, while in CHARMM

Table 3 Calculations of free energies with different sets of λ-values covering the area between two successive λ-values.

	ΔG_{bind}				ΔG_{solv}				$\Delta\Delta G$			
Standard	Low	High	Average	Middle	Low	High	Average	Middle	Low	High	Average	Middle
Step 1 (20 ps)	+5.4	+5.1	+5.2	+5.6	+0.8	−0.4	+0.2	+0.8	+4.6	+5.5	+5.0	+4.8
Step 2 (100 ps)	+37.6	+38.4	+38.0	+38.6	+43.9	+43.6	+43.8	+44.6	−6.3	−5.2	−5.6	−5.2
Step 3 (20 ps)	+3.8	+4.9	+4.4	+3.7	−0.8	+0.4	−0.2	−0.8	+4.6	+4.5	+4.6	+4.5
Step 4 (100 ps)	−41.8	−43.2	−42.4	−43.7	−43.9	−43.6	−43.8	−44.6	+2.1	+0.4	+1.4	+0.9
Sum									+5.0	+5.2	+5.1	+4.2
DONT												
Step 1 (20 ps)	+5.9	+6.5	+6.2	+6.3	+3.4	+2.8	+3.1	+3.4	+2.5	+3.7	+3.1	+2.8
Step 2 (100 ps)	+36.1	+39.8	+37.9	+39.1	+43.6	+42.8	+43.2	+42.6	−7.5	−3.0	−5.3	−3.5
Step 3 (20 ps)	−0.8	−0.7	−0.8	−1.1	−3.4	−2.8	−3.1	−3.4	+2.6	+2.1	+2.3	+2.3
Step 4 (100 ps)	−42.5	−42.6	−42.6	−43.2	−43.6	−42.8	−43.2	−42.6	+1.0	+0.2	+0.6	−0.8
Sum									−1.4	+3.0	+0.7	+0.8

the atoms are thrown in at full size directly. This causes the problems with the "FULL" method and low λ-values, described above. It should be no major problem to let the atoms grow in with a change in bond length. Thereby no van der Waals interactions would occur while λ is small, but there could still be a problem to describe changes in bond lengths when a group of atoms is involved. It might be useful to use similar methods as in X-plor [38], where several copies of one residue can be present at the same time.

Dependence of simulation length

Dependence, of simulation time for different alchemical steps

Each λ-value in step 2 was simulated for altogether 100 ps using both the "FULL" and the "DONT" methods. Several different time frames were used for calculating the ΔG. The trajectories were divided into parts of either 5, 10, 20 or 40 ps, and these different trajectories were used for the perturbation calculations. The calculations were started every 10 ps, or every 5 when a 5 ps time frame was used. Averages over the last 75 (or 70) and 50 ps of each time frame are given in table 4. In figures 5 and 6 the results from all these simulations are plotted against the middle of the period of the simulation used in the perturbation calculations. The $\Delta\Delta G$ is the difference between the line with the circles and the line with the squares.

In step 2_{bind} a free energy difference of approximately +38.5 ● 0.2 kcal/mol is obtained as an average, independent of the time intervals employed. The standard deviation between different parts of the trajectory decreases as a longer time frame is used for the calculations, but even with a 40 ps time frame the standard deviation more is than 1.5 kcal/mol. In step 2_{solv} the result was similar but we can detect a major change between 30 and 40 ps of the simulation. If this part of the trajectory is included in the free energy calculation a value of +40 kcal/mol is obtained compared with +44.8 kcal/mol if only later parts of the trajectory is used. The free energy difference ($\Delta\Delta G$) obtained for step 2 with this method and a suitably long simulation was approximately −6.4 kcal/mol (figure 5, table 1.4). With the DONT method, ΔG_{solv} was +39.2 and ΔG_{bind} was +42.4 kcal/mol respectively, yielding a $\Delta\Delta G$ of approximately −3.5 kcal (figure 6, table 1.4). In figures 5 and 6 it is shown that $\Delta\Delta G$ fluctuates from positive to negative values, both with the FULL method and the DONT method, and seems to stabilise at a negative value after about 50 ps of simulation. The fluctuations also seem to decrease as longer time spans are used.

Step 1_{bind} was also simulated for 100 ps per λ-value with the FULL method (table 4, figure 7). It is seen that the drift is much smaller than for step 2. The fluctuations also seem to be slightly smaller, i.e. step 1 (and 3) does not need as long simulation time as step 2 (and 4).

How long a simulation is required?

As we prolonged the simulation we studied two questions.

1. How long do we have to simulate to reach an equilibrium?
2. For how long do we have to sample to obtain a converged value?

The first question is probably very hard to answer. From a 1.2 ns simulation of RNT1-Glu46 with 3′GMP [39] it is shown that approximately 700 ps of simulations

Table 4 Free energy differences from different lengths of production run.

	Time frame	ΔG_{bind}	ΔG_{solv}	$\Delta\Delta G$
Step 1	5 ps**	4.7 ± 1.9		
Standard	10 ps**	4.6 ± 1.3		
50–100 ps	20 ps**	4.5 ± 1.1		
	40 ps**	4.8 ± 0.1		
	50 ps*	4.9 ± 0.3		
	Average**	4.7 ± 0.2		
25(30)–100 ps	5 ps**	4.5 ± 1.9		
	10 ps**	4.6 ± 1.4		
	20 ps**	4.8 ± 1.0		
	40 ps**	5.1 ± 0.3		
	75 ps*	4.7 ± 0.2		
	Average**	4.7 ± 0.2		
Step 2	5 ps**	38.3 ± 2.3	44.8 ± 2.5	−6.5 ± 4.8
Standard	10 ps**	38.7 ± 2.3	44.9 ± 2.3	−6.2 ± 4.5
50–100 ps	20 ps**	38.4 ± 1.8	44.7 ± 0.9	−6.3 ± 2.6
	40 ps**	38.2 ± 0.5	44.7 ± 0.3	−6.6 ± 0.8
	50 ps*	38.6 ± 0.6	44.6 ± 0.4	−6.0 ± 1.0
	Average**	38.4 ± 0.2	44.8 ± 0.1	−6.3 ± 0.3
25(30)–100 ps	5 ps**	38.4 ± 2.0	44.6 ± 2.4	−6.2 ± 4.4
	10 ps**	38.7 ± 1.9	44.1 ± 2.4	−5.4 ± 4.3
	20 ps**	38.9 ± 1.6	43.9 ± 1.9	−5.0 ± 3.5
	40 ps**	38.8 ± 1.0	43.6 ± 1.9	−4.9 ± 2.9
	75 ps*	38.6 ± 0.5	40.7 ± 0.5	−2.1 ± 1.0
	Average**	38.7 ± 0.2	43.4 ± 1.5	−4.7 ± 1.7
Step 2	5 ps**	39.4 ± 1.7	42.4 − 1.3	−2.9 ± 3.0
DONT	10 ps**	39.2 − 1.2	42.5 ± 1.2	−3.4 ± 2.3
50–100 ps	20 ps	39.2 ± 0.6	42.4 ± 0.8	−3.2 ± 1.4
	40 ps**	39.2 ± 0.0	42.4 ± 0.2	−3.2 ± 0.2
	50 ps*	39.1 ± 0.6	42.6 ± 0.3	−3.5 ± 0.9
	Average**	39.2 ± 0.1	42.5 ± 0.1	−3.3 ± 0.2
25(30)–100 ps	5 ps**	40.1 ± 1.8	41.3 ± 1.9	−1.3 ± 3.7
	10 ps**	39.8 ± 1.6	41.5 ± 2.0	−1.7 ± 3.6
	20 ps**	39.7 ± 1.1	41.6 ± 1.5	−1.9 ± 2.6
	40 ps**	39.3 ± 0.6	41.8 ± 0.7	−1.5 ± 1.3
	75 ps*	39.7 ± 0.5	41.6 ± 0.3	−2.0 ± 0.8
	Average**	39.7 ± 0.3	41.6 ± 0.2	−1.9 ± 0.4

* The standard deviation is calculated from one simulation, divided into 100 parts.

** The standard deviation is calculated from several simulations.

is needed before an equilibrium is reached as judged by the total RMS difference from the starting structure and by other factors. We do not know if the changes that occur at later stage in the simulation will influence the binding of the substrate. We can hardly detect any systematic drift in any of our five long simulations but the fluctuations are large and slow, which answers question number two: Long enough to include several phases of these fluctuations, i.e. several tens of ps for the transition from Glu to Gln. A guess would be that simulating 50 ps for equilibration and 50 ps for sampling might be enough for steps 2 and 4; if computer time is available 150 ps of sampling probably is good. It is noticeable that in the study of Hirono [1, 2] either 50 windows with $\Delta\lambda = 0.02$, with 0.3 ps equilibrium and 0.3 ps of production run or 20 windows with 1 ps equilibrium and 1 ps production are simulated, while we use 5 windows with 5 ps for equilibration and 15 ps of

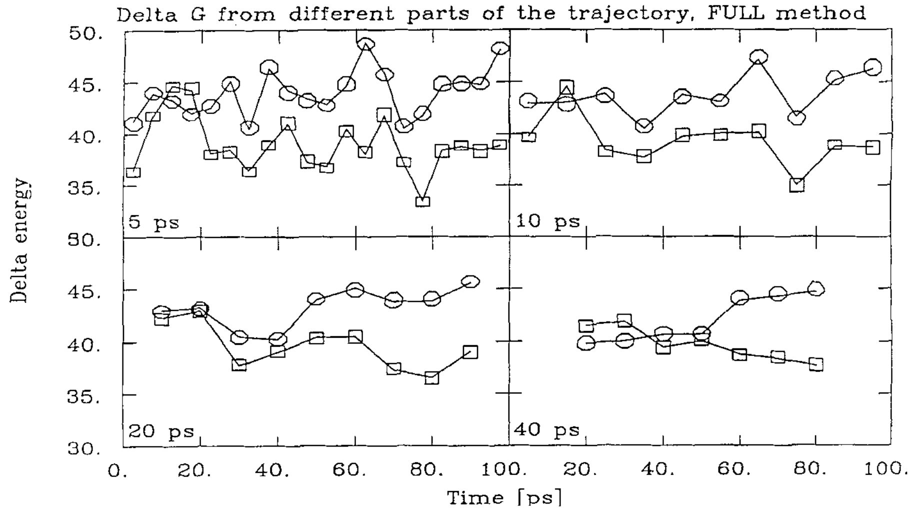


Figure 5 ΔG_{bind} (circles) and ΔG_{solv} (squares) from different parts of the 100 ps of simulation of step 2 with the FULL method and different lengths (5, 10, 20, 40 ps) of the production runs. The x-axis shows the middle time for each production run. $\Delta\Delta G$ for step 2 is calculated by subtracting ΔG_{solv} (the squares) from ΔG_{bind} (the circles).

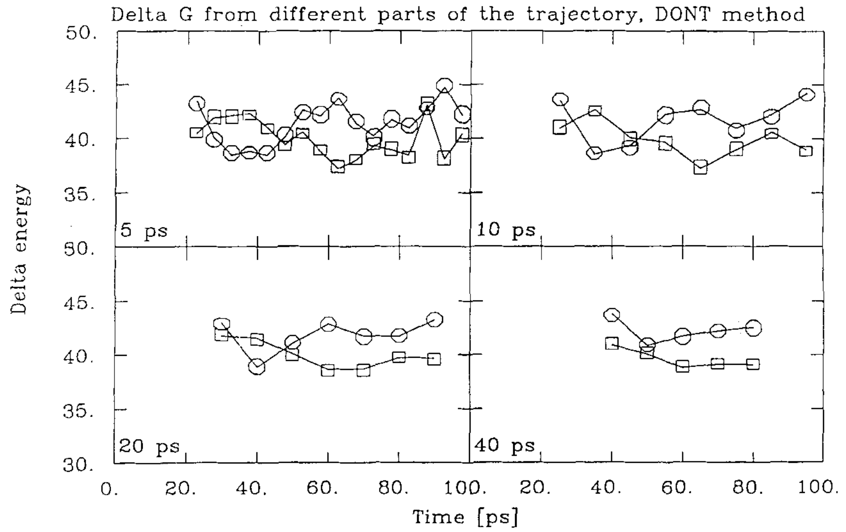


Figure 6 ΔG_{bind} (circles) and ΔG_{solv} (squares) from different parts of the 100 ps of simulation of step 2 with the DONT method and different lengths (5, 10, 20, 40 ps) of the production runs. The x-axis shows the middle time for each production run. $\Delta\Delta G$ for step 2 is calculated by substracting ΔG_{solv} (the squares) from ΔG_{bind} (the circles).

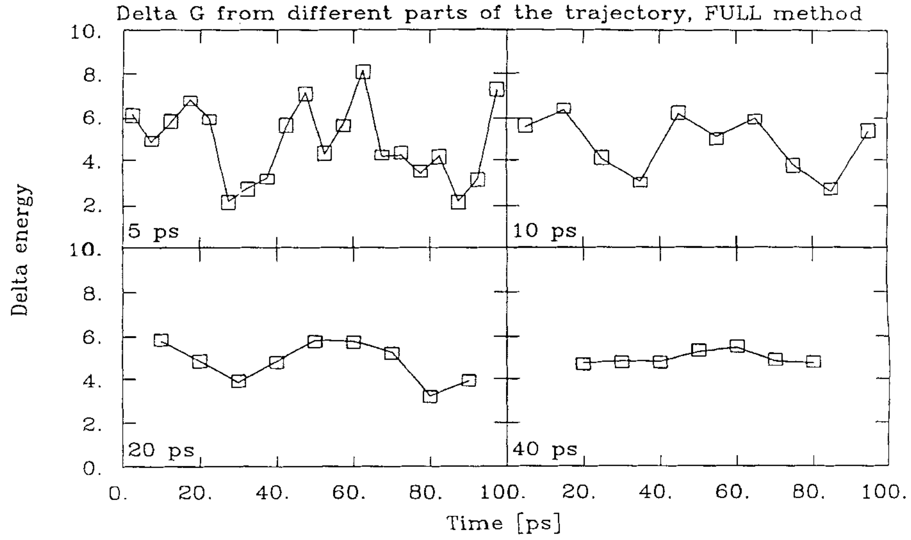


Figure 7 ΔG_{bind} from different parts of the 100 ps of simulation of step 1 with the FULL method and different lengths (5, 10, 20, 40 ps) of the production runs. The x-axis shows the middle time for each production run.

Table 5 Comparison of different simulations of step 2.

Simulation number	*Time*	ΔG_{bind}	ΔG_{solv}	$\Delta\Delta G$
1*	20 pc	+43.9 ± 0.5		
2*	20 ps	+42.3 ± 0.5		
3*	20 ps	+44.1 ± 0.5		
4*	20 ps	+41.9 ± 0.5		
5*	20 ps	+41.5 ± 0.4		
Average*	5*20 ps	+43.2 ± 1.2		
Long (FULL)	100 ps	+38.6 ± 0.6	+44.6 ± 0.4	−6.0 ± 1.0
Discrete Growth, (FULL)	20 ps	+43.9 ± 0.6	+46.0 ± 0.5	−2.1 ± 1.1
Discrete Growth, (DONT)	20 ps	+45.0 ± 0.4	+42.4 ± 0.5	−1.6 ± 1.1

λ-values (0.250, 0.500, 0.750, 0.875) are the same in all these simulations.

production. It seems reasonable that the smaller $\Delta\lambda$'s used by Hirono might decrease the total simulation time needed. This is seen as in their study 30–40 ps simulation time was needed for one backward/forward simulation while in our study 100 ps was used for the equivalent steps. In steps 2 and 4 where altogether 500 ps simulation time seems to be needed it seems unlikely that a similar scheme as used by Hirono would be long enough.

Trajectory dependence

Step 2_{bind} was simulated at $\lambda = 0.125$ with five different starting values for the random value generator giving five different trajectories. $\lambda = 0.125$ was chosen for these extra simulations because it is by far the most important factor for the perturbations, about 25 of the 40 kcal/mol are calculated with $\lambda = 0.125$. A ΔG

of +43.2 ± 1.2 kcal/mol was obtained (table 5). We can see that the error estimate from each simulation gives lower error bars than we get from the standard deviation of five different simulations. This shows that we examine a larger part of the conformational space than with only one simulation. It is also noteworthy that we have not obtained the same value as in the long (100 ps) simulation, showing the importance of (i) the higher λ-values and/or (ii) longer equilibration time.

Structure dependence

The importance of the starting structure was examined by also simulating step 2_{bind}, which gave, $\Delta G = 38.6$ kcal/mol (table 1), from the average structure of the last 100 ps of a 180 ps simulation of RNT1-Gln46 in complex with 3′GMP. We assumed that this structure would be more similar to the true first structure of the perturbation step. The value obtained from the new structure is +45.3 kcal/mol, which is very similar to the value obtained from calculating 2_{solv} (table 1). It seems as if the ligand in the average structure no longer interacts with Gln46 in the proposed way but in a more unspecific way, yielding the same result as in 2_{solv} [35].

The DISCRETE GROWTH method would allow us not to depend so much on the starting structure as it changes slowly with λ. This approach, but with many more windows (20 to 100) and shorter simulations at each λ-value, has mostly been used by the group of Kollman [1, 2, 22], notice though that the AMBER/GIBBS algorithm is not exactly the same as used in CHARMM in this study, see below for further discussion. It is probably much more important to use DISCRETE GROWTH when a smaller $\Delta\lambda$ is used between the different λ-values, since in this case one cannot afford as long equilibration period for each run. This method might have advantages for reaching the equilibrium, but as we can judge from the long simulation is not only reaching the equilibrium that is a problem but also sampling for long enough time, and it is easier to prolong a simulation with the other methods we have used here compared to the DISCRETE GROWTH method; if one knows how much computer time can be spent and what λ-values are best to use this method might be preferable. Probably it could be useful to let $\Delta\lambda$ depend on the slope for the last λ-value. This could be performed in a similar way as Pearlman has introduced dynamically modified windows [24]. One major disadvantage with DISCRETE GROWTH is that several computers can not be used in parallel, as a simulation with a particular λ-value starts from the simulation from the previous λ-value.

CONCLUSIONS

The binding of 3′GMP to RNT1-Glu46 is stabilised by several hydrogen bonds between the ligand and the enzyme. Two of these hydrogen bonds will not be present if 2AP is used as a ligand. It was predicted that the change of Glu46 to Gln46 would compensate for one of the lost hydrogen bonds. We have used several different methods for testing this hypothesis and found quite a clear answer: The mutant RNT1-Gln46 binds 2AP better than 3′GMP.

We have shown that even very complicated changes of a complicated system, such as the mutation from Glu46 to Gln46, seem to be within the reach of free energy perturbation methods. To make a complete examination of all different approaches is beyond the scope of this article. This would be an extremely computer expensive study and should therefore probably be performed on a smaller system. Such a system also should have good specificity, with a large number of mutants available and very accurate crystal structures.

The amount of computer time spent in this study, about 5 ns of simulations, is in the order of one year effective computing time on a Convex C210 or slightly less on an IBM RS/6000 320. As mentioned above a large number of computers could be used in parallel. To obtain ΔG_{solv} in step 2 or 4 would take roughly 100 CPU-hours on an IBM RS/6000 580. I.e. if ten machines could be used in parallel it would take less than four days to calculate $\Delta\Delta G$ for this step. Even with the methods and computer power available today realistic values for free energy differences can be obtained rather fast.

Although there are limitations, as shown in this study, it is probably still useful to perform free energy simulations as a first step in the redesigning of a binding site. The major problem that we discover is due to the large free energy difference between RNT1-Glu46 and RNT1-Gln46, which creates the need to simulate with low λ-values that cause problems. The exclusion of internal energies (the DONT method) seems to exclude some interactions as the $\Delta\Delta G$'s calculated with this method are about half the values obtained with the FULL method. This is in agreement with studies by Pearlman with AMBER/GIBBS.

Acknowledgements

We are very grateful for stimulating discussions with Prof. R. Rigler, Dr A.D. MacKerell, Dr U. Hahn and Prof. W. Saenger. We thank Dr T. Kulinski and Prof. R. Rigler for experimental results and Dr Ulf Skoglund at Karolinska Institutet for the use of a Convex C210. This study was supported by the Swedish board for technical development as a part of their program for Biotechnology, and by the Swedish Natural Science Research Council.

References

[1] S. Hirono and P. Kollman A, "Calculation of the Relative Binding Free Energy of 2′GMP and 2′AMP to Ribonuclease T1 Using Molecular Dynamics/Free Energy Perturbation Approaches." *J. Mol. Biol*, **212**, 197–209, (1990).

[2] S. Hirono and P.A. Kollman, "Relative binding free energy calculations of inhibitors to two mutants (Glu46 - Ala/Gln) of ribonuclease T1 using molecular dynamics/free energy perturbation approaches." *Prot. Eng*, **4**, 233–243, (1991).

[3] J. Granzin, R. Puras-Lutzke, O. Landt, H.P. Grunert, U. Heinemann, W. Saenger and U. Hahn, "RNase-T1 Mutant Glu46Gln Binds the Inhibitors 2′GMP and 2′AMP at the 3′ Subsite." *Journal of Molecular Biology*, **225**, 533–542, (1992).

[4] C.N. Pace, U. Heinemann, U. Hahn and W. Saenger, "Ribonuclease T1: Structure, Function, and Stability." *Angewandte Chemie*, **30**, 343–360, (1991).

[5] J. Martinez-Oyanedel, H.W. Choe, U. Heinemann and W. Saenger, "Ribonuclease T1 with Free Recognition and Catalytic Site: Crystal Structure Analysis at 1.5 Å Resolution." *J. Mol. Biol*, **222**, 335–352, (1991).

[6] U. Heinemann and W. Saenger, "Specific Protein-Nucleic Acid Recognition in Ribonuclease T1 - 2′-Guanylic Acid Complex: An X-Ray Study." *Nature*, **299**, 27–31, (1982).

[7] R. Arni, U. Heinemann, R. Tukuoka and W. Saenger, "Three-dimensional structure of the ribonuclease T1*2′-GMP complex at 1.9 Å resolution." *J. Biol. Chem.*, **263**, 15358–15368, (1988).

[8] S. Sugio, T. Amisaka, H. Ohishi and K.-I. Tomita, *FEBS Lett.*, **181**, 129, (1985).

[9] S. Sugio, K.-I. Oka, H. Ohishi, K.-I. Tomita and W. Saenger, "Three dimensional structure of

the ribonuclease T1*3′-guanylic acid complex at 2.6 Å resolution." *FEBS Lett.*, **183**, 115-118, (1985).

[10] J. Koepke, M. Maslowska, U. Heinemann and W. Saenger, "Three-dimensional structure of ribonuclease T1, complexed with guanylyl-2′,5′-guanosine at 1.8 A resolution." *J. Mol. Biol.*, **206**, 475-488, (1989).

[11] A. Lenz, F. Cordes, U. Heineman and W. Saenger, "Evidence for a substrate binding subsite in ribonuclease T1: crystal structure of the complex with two guanosines, and model building of the complex with the substrat GpG." *J. Biol. Chem.*, **266**, 7661-7667, (1991).

[12] A. Lenz, U. Heineman, M. Maslowska and W. Saenger, "X-ray analysis of cubic crystals of the complex formed between ribonuclease T1 and guanosine-3′,5′-bisphosphate." *Acta Crystallogr. Sect. B*, **47**, 521-527, (1991).

[13] D. Kostrewa, H.-W. Choe, U. Heineman and W. Saenger, "Crystal structure of guanisine-free ribonuclease T1, complexed with vanadate(V), suggests conformational change upon substrate binding." *Biochemisity*, **28**, 7592-7600, (1989).

[14] K. Takahashi and S. Moore, "Ribonuclease T1", *The Enzymes 3rd Ed.*, 1982, p. 435-467.

[15] T.M. Nordlund, S. Anderson, L. Nilsson, R. Rigler, A. Gräslund and L.W. McLaughlin, "Structure and Dynamics of a fluorescent DNA Oligomer Containing the Eco RI Recognition Sequence: Fluorescence, NMR and Molecular Dynamics Studies." *Biochemistry*, **28**, 9095-9103, (1989).

[16] D.L. Beveridge and F.M. DiCapua, "FREE ENERGY VIA MOLECULAR SIMULATION: Application to chemical and Biomolecular Systems." *Ann. Rev. Biophys. Biophys. Chem.*, **18**, 431-492, (1989).

[17] W. Jorgensen and C. Ravimohan, "Monte Carlo simulation of differences in free energies of hydration" *J. Chem. Phys.*, **83**, 3050, (1985).

[18] S.H. Fleischman and C.L. Brooks, "Thermodynamics of aqueous solvation: Solution properties of alcohols and alkanes" *J. Chem. Phys.*, **87**, 3029-3034, (1987).

[19] T. Härd and L. Nilsson, "Free Energy Calculations Predict Sequence Specificity in DNA-Drug Complexes." *Nucleosides & Nucleotides*, **11**, 167-173, (1992).

[20] J. Åqvist and A. Warshel, "Free Energy Relationships in Metalloenzyme-Catalyzed Reactions. Calculations of the Effects of Metal Ion Substitutions in Staphylococcal Nuclease." *J. Am Chem. Soc.*, **112**, 2860-2868, (1990).

[21] A.E. Mark, W.F. van Gunsteren and H.J.C. Berendsen, "Calculations of relative free energy via indirect pathways." *J. Chem. Phys.*, **94**, 3808-3816, (1991).

[22] D.A. Pearlman and P.A. Kollman, "The overlooked bond-stretching contribution in free energy perturtbation calculations." *J. Chem. Phys.*, **94**, 4532-4545, (1991).

[23] M.J. Mitchell and J.A. McCammon, "Free Energy Difference Calculations by Thermodynamic integratio: Difficulties in Obtaining a Precise Value." *J. Comp. Chem.*, **12**, 271-275, (1991).

[24] D.A. Pearlman and P.A. Kollman, "A new method for carrying out free energy perturbation calculations: Dynamically modified windows." *J. Chem Phys.*, **90**, 2460-2470, (1989).

[25] B.R. Brooks, R.E. Bruccoleri, B.D. Olafsen, D.J. States, S. Swaminathan and M. Karplus, "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations" *J. Comp. Chem.*, **4**, 187-217, (1983).

[26] C.L. Brooks, A. Brünger and M. Karplus, "Active site dynamics in protein molecules: A stochastic boundary molecular-dynamics approach" *Biopolymers*, **24**, 843, (1985).

[27] W.L. Jorgensen, J. Chandrasekhar, J. Madura, R.W. Impey and M.L. Klein, *J. Chem. Phys.*, **79**, 926-, (1983).

[28] A.D. MacKerell, R. Rigler, L. Nilsson, U. Hahn and W. Saenger, "Protein Dynamics: A Time-Resolved Fluorescence, Energetic and Molecular Dynamics Study of Ribonuclease T1." *Biophys. Chem.*, **26**, 247-261, (1987).

[29] A.D. MacKerell, L. Nilsson, R. Rigler and W. Saenger, "Molecular Dynamics Simulations of Ribonuclease T1: Analysis of the Effect of Solvent on the Structure, Fluctuations and Active Site of the Free Enzyme." *Biochemistry*, **27**, 4547-4556, (1988).

[30] A.D. MacKerell, R. Rigler, L. Nilsson, U. Heinemann and W. Saenger, "Molecular Dynamics of Ribonuclease T1: Effect of Solvent on the Interaction with 2′GMP." *Eur. Biophys. J.*, **16**, 287-297, (1988).

[31] A.D. MacKerell, L. Nilsson, R. Rigler, U. Heinemann and W. Saenger, "Molecular Dynamics of Ribonuclease T1: Comparison of the Free, Enzyme and the 2′GMP-Enzyme Complex." *PROTEINS: Structure, Function, and Genetics*, **6**, 20-31, (1989).

[32] J.P. Ryckaert, G. Ciccotti and H.J.C. Berendsen, "Numerical Integration of the Cartesian

Equations of Motion of a System with Constraints: Molecular Dynamics of n-alkanes." *J. Comp. Phys.*, **23**, 327-337, (1977).

[33] W.F. van Gunsteren, *Computer Simulation of Biomolecular Systems; Theoretical and Experimental Applications.*, W.F. van Gunsteren, Editor. 1989, Escom Science: Netherlands. p. 27.

[34] P.A. Bash, U.C. Singh, R. Langridge and P.A. Kollman, "Free energy calculations by computer simulations." *Science*, **236**, 564-568, (1987).

[35] A. Elofsson, T. Kulinski, R. Rigler and L. Nilsson, "Site Specific Point Mutations Changes Specificity. A Molecular Modeling Study by Free Energy Stimulations of the Thermodynamics in Ribonuclease T1 Substrate Interactions." In press in *PROTEINS*: *Structure*, *Function*, *and Genetics*, (1993).

[36] T. Kulinski, "Personal communication", 1992.

[37] T.P. Straatsma, "Free Energy Evaluation by Molecular Dynamics Techniques", PhD thesis, Groningen, The Netherlands, 1987.

[38] A. Brünger, "Personal communication", 1992.

[39] A. Elofsson and L. Nilsson, "A 1.2 ns Molecular Dynamic Simulation of Ribonuclease T1 in complex with Guanosin mono phosphate. Studies of Secondary Structure Stability, Time Development of RMS Differences, Substrate Binding, Correlation Functions and Interaction Between Water and Protein." *Manuscript in preparation*, (1993).

[40] Chemwin, "ChemWindows", 1991, Soft Shell Int'l ltd: