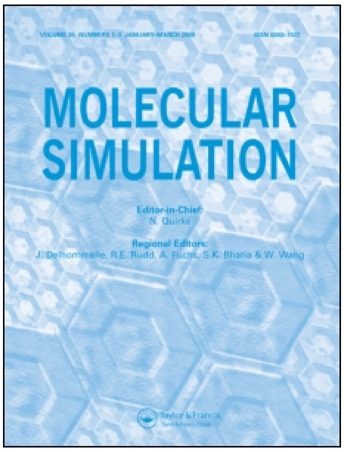

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

A Comparison of Seven Fast but Approximate Methods to Compute the Free Energy of Deprotonation for Amino Acids in Aqueous Solution

Roland Bürgi[a]; Florian Läng[a]; Wilfred F. Van Gunsteren[a]
[a] Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zürich, Zürich, Switzerland

To cite this Article Bürgi, Roland , Läng, Florian and Van Gunsteren, Wilfred F.(2001) 'A Comparison of Seven Fast but Approximate Methods to Compute the Free Energy of Deprotonation for Amino Acids in Aqueous Solution', Molecular Simulation, 27: 4, 215 — 236

To link to this Article: DOI: 10.1080/08927020108027948

URL: http://dx.doi.org/10.1080/08927020108027948

Molecular Simulation, 2001, Vol. 27, pp. 215–236
Reprints available directly from the publisher
Photocopying permitted by license only

A COMPARISON OF SEVEN FAST BUT APPROXIMATE METHODS TO COMPUTE THE FREE ENERGY OF DEPROTONATION FOR AMINO ACIDS IN AQUEOUS SOLUTION

ROLAND BÜRGI, FLORIAN LÄNG
and WILFRED F. VAN GUNSTEREN*

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology
Zürich, CH-8092 Zürich, Switzerland

(Received October 2000; accepted October 2000)

In recent years, a variety of methods based on statistical mechanics have been successfully applied to calculate free energy differences of chemical reactions from molecular simulation. The accuracy and computational efficiency vary strongly between these methods. Seven approximate but fast methods to calculate free energy differences are compared in terms of accuracy and efficiency with the accurate but expensive thermodynamic integration method as reference, using 28 protonation and deprotonation reactions of aspartic acid in aqueous solution as test cases. At least two simulations are required to obtain an accurate free energy difference between two states of the system. Both, the averaged one-step perturbation method and the linear response method yield the most accurate results, while the latter method shows the fastest convergence.

Keywords: Free energy; Protonation; Amino acid; Molecular dynamics

1. INTRODUCTION

Calculating free energy differences between two systems or between two states of a system or between the begin and end points of chemical reactions is a major topic in computational chemistry. For multi-dimensional condensed phase systems, free energy differences can be obtained from

---

*Corresponding author. e-mail: wfvgn@igc.phys.chem.ethz.ch

Monte Carlo or molecular dynamics (MD) simulations [1–7]. A range of methods have been developed [8–16] and successfully applied in recent years [17–20]. These methods differ in accuracy and computational expense, and most often it is not easy to predict which method will yield sufficient accuracy within a reasonable time spent on computing given a particular application. The problem of calculating a free energy difference using MD simulation becomes more difficult the bigger the difference between initial and final state is – most of the fast methods to calculate free energy differences assume that the initial and final state of the system are quite close in conformational space. Therefore, the accuracy of the fast methods will decrease enormously for big changes to the system. How much the accuracy is decreased is the subject of the current investigation.

Two major methods to obtain the exact free energy difference between two states A and B of a system or between two systems A and B are the thermodynamic integration method [21] and the multi-step perturbation method [22, 23]. Since they require a sizeable number of simulations at intermediate states I_i $(i = 1, 2, \ldots, M)$ that connect the end states $A = I_1$ and $B = I_M$ through a path characterised by small conformational differences between intermediate states I_i and I_{i+1} $(i = 1, 2, \ldots, M-1)$, these methods are computationally expensive. These methods are called M-step methods, since they require M separate simulations. The computational expense would be enormously reduced if it were possible to obtain accurate free energy differences through a one-step or one-point or one-simulation method. Amongst the one-step methods, two types can be distinguished (see methods section).

(1) The one-step perturbation formula [22]

$$\Delta F_{BA} := F_B - F_A = -k_B T \ln \langle e^{-(H_B - H_A)/k_B T} \rangle_A \tag{1}$$

requiring only a simulation of state A.

(2) The extrapolation formula of order N [10, 12]

$$\Delta F_{BA} = \sum_{n=1}^{N} \frac{1}{n!} F^{(n)}(A) \tag{2}$$

also requiring only a simulation of state A.

Since the one-step methods require only one simulation of state A for obtaining an approximate free energy difference ΔF_{BA}, they are computationally very efficient: a set of K free energy differences $\Delta F_{B_k A}$ with $k = 1, 2, \ldots, K$ can be calculated from a single simulation of state A. If the states B_k are very different from state A, one-step methods are not expected to yield accurate free energy differences ΔF_{B_kA}. Then one could use so-called two-step methods, which calculate the free energy difference ΔF_{BA} from two simulations, *e.g.*, of the end points A and B or of one end point A and an intermediate point I_2. Two-step perturbation and two-step extrapolation methods can again be distinguished. Two-step methods are more than twice as computationally expensive as their one-step equivalents in the case more than one free energy difference ΔF_{BA} is to be obtained: the calculation of K free energy differences ΔF_{B_kA} with $k = 1, 2, \ldots, K$ requires $K+1$ simulations (of states A and B_k) using two-step methods.

The different approximative methods have been separately tested by comparison of their results to exact ones for model systems or model reactions [8 – 16]. For example, Hummer and Szabo [12] considered the (de)charging of a water molecule in aqueous solution. Here, we consider a realistic model of the (de)protonation of the three titratable sites or atom groups of aspartic acid in aqueous solution. The results of seven fast but approximative one-step and two-step methods will be compared to the high-accuracy results obtained by (expensive) multi-step thermodynamic integration.

2. METHODS

In this section, we briefly describe the methods used to obtain the free energy differences ΔF_{BA}.

2.1. Thermodynamic Integration (TI)

In the thermodynamic integration (TI) method [1, 21], a reaction coordinate λ is chosen such that $\lambda = 0$ represents the initial state A and $\lambda = 1$ represents the final state B of the system under investigation. The simplest way to make the Hamiltonian λ-dependent is to choose a linear dependence on the initial and end states:

$$H(\lambda) = (1 - \lambda)H_A + \lambda H_B, \tag{3}$$

where H_A and H_B denote the Hamiltonians of the initial and final states respectively.

The free energy difference is then given by the TI formula:

$$\Delta F_{BA} = \int_0^1 \frac{dF(\lambda)}{d\lambda} d\lambda = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda, \tag{4}$$

where $\langle X \rangle_\lambda$ denotes the ensemble average of quantity X in state λ of the system. This relation between the free energy and the Hamiltonian follows directly by differentiation with respect to λ of the definition of the free energy

$$F = -k_B T \ln Z, \tag{5}$$

where k_B denotes the Boltzmann constant, T the temperature and Z the canonical partition function.

The accuracy of ΔF_{BA} obtained through TI is determined by the number and values of the intermediate λ-points chosen to approximate the integration in (4) and by the length of the simulation at each λ-point, which determines the accuracy of the ensemble average $\langle \cdots \rangle_\lambda$ in (4).

2.2. One-Step Perturbation (OSP)

Another approach for calculating free energy differences is the one-step perturbation (OSP) [10, 21]. From the definition of the free energy (5) follows directly the relation

$$\Delta F_{BA} = -k_B T \ln \int_\Xi d\xi p_A(\xi) e^{-\beta \Delta H_{BA}(\xi)}, \tag{6}$$

where ξ is the vector describing the state of the system within the phase space Ξ, $\Delta H_{BA} = H_B(\xi) - H_A(\xi)$, $\beta = 1/k_B T$, and $p_A(\xi)$ is the probability of finding the system in state ξ within ensemble A. Rewriting (6) as an ensemble average yields

$$\Delta F_{BA} = -k_B T \ln \langle e^{-\beta(H_B - H_A)} \rangle_A. \tag{7}$$

Only one MD simulation is needed to calculate the free energy difference. However, the OSP method suffers from the fact that states with a high probability in ensemble B are only well sampled if states A and B are not far apart in conformational space. Therefore, the OSP method converges very slowly if the differences between states A and B are large.

2.3. Averaged One-Step Perturbation (AOSP)

To improve the sampling, we can modify (6) by integrating twice over phase space, once over ensemble A and once over ensemble B and take the square root [12]:

$$\Delta F_{BA} = -k_B T \frac{1}{2}\left[\ln \int_{\Xi} d\xi p_A(\xi) e^{-\beta \Delta H_{BA}(\xi)} - \ln \int_{\Xi} d\xi p_B(\xi) e^{-\beta \Delta H_{AB}(\xi)}\right]$$
$$= -k_B T \frac{1}{2}[\ln \langle e^{-\beta \Delta H_{BA}} \rangle_A - \ln \langle e^{-\beta \Delta H_{AB}} \rangle_B]. \tag{8}$$

The AOSP method, however, requires two simulations, one over ensemble A and one over ensemble B, and is therefore computationally more expensive than the OSP method.

2.4. Half-Way Averaged One-Step Perturbation (HAOSP)

Another way of improving the OSP method is to extrapolate ensembles A ($\lambda = 0$) and B ($\lambda = 1$) to the intermediate point $I_{0.5}$ with $\lambda = 0.5$ and obtain the total free energy difference from the combination [12]:

$$\Delta F_{BA} = (F_{I_{0.5}} - F_A) - (F_{I_{0.5}} - F_B)$$
$$= -k_B T[\ln \langle e^{-\beta \Delta H_{BA}/2} \rangle_A - \ln \langle e^{-\beta \Delta H_{AB}/2} \rangle_B]. \tag{9}$$

The computational cost of this method is the same as for the AOSP method, provided that it has a similar convergence behaviour.

2.5. Taylor Expansion (TE)

From the definition of the free energy (5), the n-th derivative of the free energy with respect to λ can be calculated analytically [10, 22]:

$$F^{(n)}(\lambda) = (-\beta)^{n-1} C_n(\lambda), \quad n \geq 1. \tag{10}$$

When assuming a linear dependence of $H(\lambda)$ on λ as in (3), one has to third order [10]

$$C_1(\lambda) = \langle \Delta H_{BA} \rangle_\lambda \tag{11a}$$

$$C_2(\lambda) = \langle (\Delta H_{BA} - C_1)^2 \rangle_\lambda \tag{11b}$$

$$C_3(\lambda) = \langle (\Delta H_{BA} - C_1)^3 \rangle_\lambda. \tag{11c}$$

Thus, the free energy difference ΔF_{BA} can be calculated as a Taylor expansion (TE) of order N:

$$\Delta F_{BA} = \sum_{n=1}^{N} \frac{1}{n!} F^{(n)}(\lambda = 0). \tag{12}$$

As the OSP method, the TE method requires only one MD simulation to calculate ΔF_{BA}. The TE, however, suffers a severe difficulty: Not only the ensemble averages in the coefficients C_n must be converged, but also the Taylor series expansion itself. In practice the higher-order coefficients converge very slowly, which implies that a long simulation trajectory is needed to calculate ΔF_{BA} [10].

2.6. Linear Response (LR)

The linear response (LR) method [9, 11] is a first-order Taylor expansion of the AOSP formula:

$$\Delta F_{BA} = \frac{1}{2}[\langle \Delta H_{BA} \rangle_A - \langle \Delta H_{AB} \rangle_B] \tag{13}$$

If the distribution of ΔH_{BA} in the ensembles A and B is assumed to be Gaussian, the variance terms (11b) cancel out on average:

$$\langle \Delta H_{BA} - \langle \Delta H_{BA} \rangle_A \rangle_A - \langle \Delta H_{AB} - \langle \Delta H_{AB} \rangle_B \rangle_B = 0, \tag{14}$$

which makes Eq. (13) exact up to second order inclusive.

Another way to interpret the LR formula (13) is to consider it an average of the first-order TE expansions. As for the AOSP, two simulations are needed when applying the LR method.

3. MODEL SYSTEM AND COMPUTATIONAL PROCEDURE

The model system chosen is an aspartic acid (Asp) solute molecule in aqueous solution. The topology and nomenclature of the aspartic acid is shown in Figure 1. The aspartic acid was modeled using the GROMOS96 biomolecular force field, parameter set 43A1 [24]. The force field parameters for the different protonation states of the Asp solute are given in Table I. Water was modeled using the simple point charge (SPC) model [25].

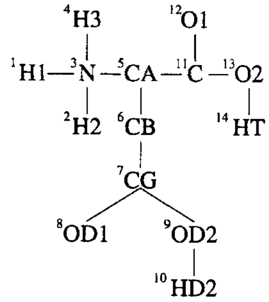

FIGURE 1 Topology and nomenclature of the aspartic acid solute.

The MD simulations were carried out using the simulation package GROMOS96 [24, 26]. The aspartic acid was solvated in a periodic truncated octahedron with 755 water molecules. The density of the system was $1001.13\,\mathrm{kg\,m^{-3}}$. All simulations were carried out at constant volume and with twin-range cut-off radii of 0.8 and 1.4 nm for the non-bonded forces. The non-bonded forces in the intermediate range were calculated every 5 time steps of 2 fs. For the electrostatic interactions beyond 1.4 nm, a reaction-field approximation [27] was used with the dielectric constant

TABLE I Force field parameters for the different protonation states of the aspartic acid. The atom names and numbering are specified in Figure 1. IAC denotes the integer atom code, which determines the van der Waals interaction, ICB the bond type code, ICT the bond-angle type code, ICQ the improper dihedral type code, and ICP the torsional dihedral type code as defined in [24]. The subscripts d and p stand for the deprotonated and the protonated state of a titration site. Atoms with IAC = 18 do not have van der Waals interactions [24], thus the hydrogen atoms of the three titration sites, H3, HD2 and HT, are dummy atoms regarding the non-bonded interaction in the deprotonated state

Atom		*Deprotonated*		*Protonated*	
Number	*Name*	*IAC_d*	*$Charge_d$ [e]*	*IAC_p*	*$Charge_p$ [e]*
1	H1	18	0.415	18	0.248
2	H2	18	0.415	18	0.248
3	N	6	−0.830	7	0.129
4	H3	18	0	18	0.248
5	CA	12	0	12	0.127
6	CB	13	0	13	0
7	CG	11	0.270	11	0.530
8	OD1	2	−0.635	1	−0.380
9	OD2	2	−0.635	3	−0.548
10	HD2	18	0	18	0.398
11	C	11	0.270	11	0.530
12	O1	2	−0.635	1	−0.380
13	O2	2	−0.635	3	−0.548
14	HT	18	0	18	0.398

Bond	ICB_d	ICB_p
1–3	2	2
2–3	2	2
3–4	2	2
3–5	20	20
5–6	26	26
5–11	26	26
6–7	26	26
7–8	5	4
7–9	5	12
9–10	1	1
11–12	5	4
11–13	5	12
13–14	1	1

Bond angle	ICT_d	ICT_p
1–3–5	10	10
1–3–2	9	9
1–3–4	9	9
2–3–4	9	9
2–3–5	10	10
4–3–5	10	10
3–5–6	12	12
3–5–11	12	12
11–5–6	12	12
5–6–7	14	14
6–7–8	21	29
6–7–9	21	18
8–7–9	37	32
7–9–10	11	11
5–11–12	21	29
5–11–13	21	18
12–11–13	37	32
11–13–14	11	11

Improper dihedral	ICQ_d	ICQ_p
11–5–13–12	1	1
5–3–11–6	2	2
7–8–9–6	1	1

Torsional dihedral	ICP_d	ICP_p
1–3–5–11	14	14
3–5–6–7	17	17
3–5–11–13	20	20
5–6–7–9	20	20
6–7–9–10	3	3
5–11–13–14	3	3

$\varepsilon_{RF} = 54$ [28]. The temperature was weakly coupled to a temperature bath of 298 K with a coupling time of 0.1 ps [29]. The configurations of the trajectories were saved every 0.5 ps.

In the TI calculations, 16 λ points were simulated for the different reactions. For all λ points, $\langle \partial H/\partial\lambda \rangle_\lambda$ turned out to be converged after 100 ps. The free energy estimates based on the approximative one-step and two-step methods were obtained from MD trajectories of 4 ns, from which the first 200 ps were taken as equilibration period and thus omitted. For the non-bonded interactions of the perturbed atoms, a soft-core potential was used [24, 26, 30] with soft-core parameter values $\alpha_{LJ} = \alpha_C = 0.5$.

4. RESULTS AND DISCUSSION

Aspartic acid possesses three titration sites. This yields a total of $2^3 = 8$ different protonation states of this solute. To simplify the notation, we will denote the protonation state of a titration site with 0 if it is deprotonated and with 1 if it is protonated. The protonation state of the Asp solute containing three titration sites will be characterised by a three digit code of zeros and ones in the order of the sites amine (N)–carboxylate (side chain OD2)–carboxylate (O2). Combining each of the 8 protonation states with each other gives a total of $7 \cdot 8/2 = 28$ different reaction pathways to be considered forwards and backwards.

The values of the free energy differences for all 28 reaction paths calculated forwards and backwards (if different) using seven of the approximative methods discussed are presented below and compared to the exact TI values. As a measure of the accuracy of the different approximative methods, we will use the *coefficient of determination* R^2, which is the square of the multiple correlation coefficient when fitting the data obtained by the approximative method under consideration linearly to the TI data. The closer R^2 is to 1, the more accurate is the approximative method. Furthermore, the convergence behaviour of each method is discussed and the different methods are analysed in terms of additivity of the free energy differences.

4.1. Thermodynamic Integration (TI)

The results obtained from TI will be used as the reference values for the free energy differences obtained by the other methods in order to assess their accuracy. 16 λ points were calculated for the different paths. As for the convergence behaviour of the TI, all λ points were well converged after 100 ps. Therefore, 1600 ps of simulation are needed in total per pathway.

Between the configurations saved every 0.5 ps, $\partial H/\partial\lambda$ was hardly correlated, and therefore the standard deviation was a suitable measure for the error in $\langle\partial H/\partial\lambda\rangle$. For each path, a Monte Carlo integration was carried out: The value of $\langle\partial H/\partial\lambda\rangle$ was chosen randomly from the normal distribution at each λ point, and the free energy difference and the error were taken as the mean value and standard deviation of the mean value of 100 such integrals. The results are presented in Table II: The initial states are listed in the rows and the final states are listed in the columns.

Not unexpectedly, the triple (de)protonation, 000 to 111, or *vice versa*, shows the largest error. To further test the accuracy of TI, seven backwards paths were calculated. There is almost no hysteresis, the biggest hysteresis (sum of forwards and backwards results) is about 2% for the reaction 100 to 101. All other free energy differences of the backwards paths are well within the error estimates.

The natural state of the aspartic acid at pH = 7 is the state 100. This can be seen in Table II by noting that all free energy differences with 100 as the initial state are positive and all free energy differences with 100 as the final state are negative. The most unfavourable state is 011. All free energy differences with 011 as the initial state are negative and all differences with 011 as the final state are positive.

4.2. One-Step Perturbation (OSP)

The free energy differences calculated with the one-step perturbation (OSP) method as well as the relative errors to thermodynamic integration (TI) are presented in Table III. The relative errors to TI range up to 2700%. If only changes in one protonation site are considered, the errors are all below 100%. However, 11 of the signs (out of a total of 56) are not correct. For the OSP, the coefficient of determination is $R^2 = 0.709$.

The convergence behaviour of the OSP method is shown in panels A, B, D, and E of Figure 2.

Two examples to illustrate convergence were chosen: the reaction 010 to 110, which yields good results for all methods considered, and 101 to 110, which yields rather poor results for all methods. For both reactions we observe the typical saw tooth behaviour of the free energy differences: Every time a structure is sampled in ensemble A that is favourable for the Hamiltonian of ensemble B, there is a drop in the corresponding curve. Therefore, it is almost impossible to determine when the OSP has converged – there is no way to predict whether another drop will occur.

TABLE II Free energy differences and errors computed with thermodynamic integration (TI), Eq. (4), in [kJ/mol]. The protonation states of the Asp solute are indicated using a zero or a one for each of the three sites (N, OD2, O2), see Figure 1. Values not calculated are indicated as n.a. Initial states are indicated by rows, final states by columns

	000	001	010	011	100	101	110	111
000	n.a.	252.2 ± 1.7	175.0 ± 1.5	456.7 ± 3.0	− 285.5 ± 1.6	− 96.3 ± 2.6	− 84.1 ± 2.6	129.5 ± 3.5
001	n.a.	n.a.	− 76.5 ± 2.1	200.4 ± 1.6	− 530.9 ± 1.6	− 346.3 ± 1.4	− 334.4 ± 2.2	− 119.4 ± 2.6
010	− 176.3 ± 2.0	82.1 ± 2.5	n.a.	275.0 ± 1.5	− 462.3 ± 1.7	− 265.6 ± 2.5	− 261.9 ± 1.5	− 42.6 ± 2.6
011	n.a.	n.a.	n.a.	n.a.	− 732.6 ± 2.3	− 545.3 ± 1.8	− 534.7 ± 2.0	− 318.0 ± 1.3
100	285.3 ± 1.5	n.a.	n.a.	n.a.	n.a.	184.1 ± 1.5	196.3 ± 1.6	414.4 ± 2.5
101	n.a.	n.a.	n.a.	n.a.	− 183.9 ± 1.4	n.a.	10.5 ± 2.0	230.6 ± 1.5
110	n.a.	n.a.	n.a.	n.a.	− 199.2 ± 1.6	n.a.	n.a.	217.4 ± 1.5
111	− 128.4 ± 3.6	n.a.	n.a.	n.a.	− 412.3 ± 2.7	n.a.	n.a.	n.a.

TABLE III Free energy differences as computed with the one-step perturbation (OSP) method, Eq. (7), and the relative errors with respect to the thermodynamic integration (TI) results. See also caption of Table II

	000	001	010	011	100	101	110	111
Free energy differences [kJ/mol]								
000	n.a.	410.2	338.7	927.7	−160.5	413.6	434.4	1216.0
001	−84.0	n.a.	186.8	366.3	−372.2	−224.7	33.6	308.9
010	−7.9	350.7	n.a.	433.0	−293.0	192.7	−130.1	496.2
011	50.6	−8.0	−115.2	n.a.	−370.8	−343.3	−337.9	−216.3
100	405.9	677.1	626.5	1018.0	n.a.	348.9	358.5	878.8
101	585.1	443.2	743.0	682.8	−31.9	n.a.	215.9	380.9
110	550.6	685.6	378.1	682.2	−26.8	257.4	n.a.	361.6
111	881.4	560.6	563.9	434.5	20.0	−78.9	−74.2	n.a.
Relative errors to TI [%]								
000	n.a.	62.3	93.1	103.1	43.5	530.2	616.7	841.9
001	66.7	n.a.	346.1	82.5	29.8	35.0	110.2	359.7
010	95.5	326.8	n.a.	57.5	36.6	172.8	50.2	1281.0
011	111.2	96.0	57.8	n.a.	49.3	36.9	36.7	31.9
100	42.1	27.6	35.5	38.9	n.a.	89.1	82.7	112.1
101	508.3	27.8	180.0	25.1	82.5	n.a.	2050.0	65.2
110	554.8	105.1	44.4	27.5	86.4	2670.0	n.a.	66.4
111	788.3	370.6	1240.5	36.6	104.8	65.7	65.9	n.a.

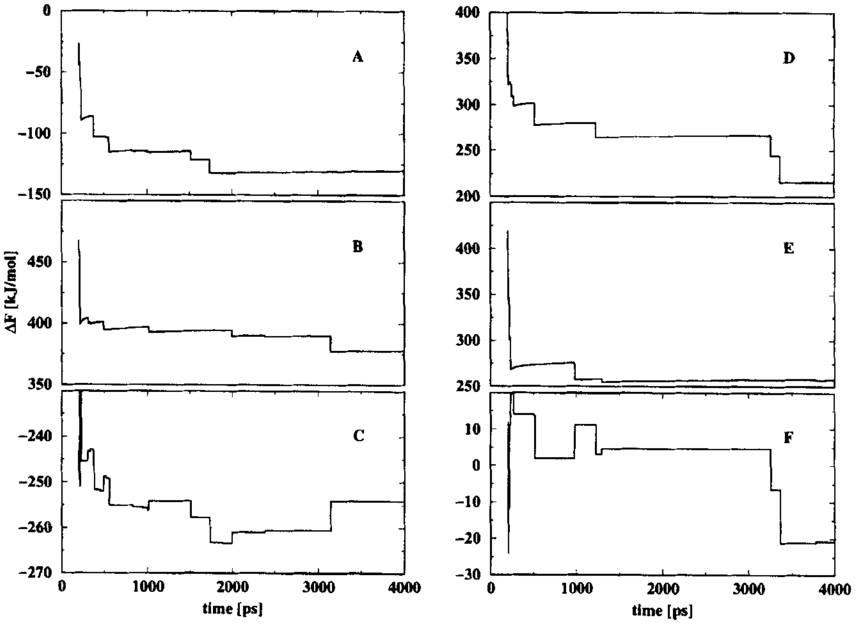

FIGURE 2 Convergence behaviour of one-step perturbation (OSP) and the averaged one-step perturbation (ASOP) methods: Panel A shows the convergence of OSP for the reaction 010 to 110, panel B the convergence of OSP for 110 to 010 (the inverse reaction), and panel C the convergence of AOSP for 010 to 110. Panel D shows the convergence of OSP for 101 to 110, panel E the convergence of OSP for 110 to 101 (the inverse reaction), and panel F the convergence of AOSP for 101 to 110.

4.3. Averaged One-Step Perturbation (AOSP)

The free energy differences and relative errors to TI as calculated with the averaged one-step perturbation (AOSP) method are presented in Table IV. Except for the reaction 101 to 110, all of the signs are correct, and except for four reactions, all of the errors are below 10%. Especially the reactions involving only one change in protonation are very accurate. Their relative errors are between 0% for 000 to 100 and 6% for 001 to 011. The coefficient of determination is $R^2 = 0.998$, which confirms the accuracy of the method.

The convergence behaviour of the AOSP is shown in panels C and F of Figure 2. The saw tooth behaviour of the OSP method is still visible in the convergence of the AOSP method. Both panels C and F show that the convergence of the AOSP is very slow, and it is hardly possible to tell when the free energy differences are converged.

4.4. Half-Way Averaged One-Step Perturbation (HAOSP)

Table V shows the free energy differences and their relative errors to TI as calculated with the half-way averaged one-step perturbation (HAOSP) method. Even though one might expect the HAOSP to be more accurate than the AOSP, since the perturbation steps from ensembles *A* and *B* are

TABLE IV Free energy differences as computed with the averaged one-step perturbation (AOSP) method, Eq. (8), and the relative errors to the thermodynamic integration (TI) results. See also caption of Table II

	000	001	010	011	100	101	110	111
Free energy differences [kJ/mol]								
000	n.a.	247.1	173.3	438.6	−283.2	−85.7	−58.1	167.3
001	−247.1	n.a.	−82.0	187.1	−524.7	−334.0	−326.0	−125.8
010	−173.3	82.0	n.a.	274.1	−459.8	−275.1	−254.1	−33.9
011	−438.6	−187.1	−274.1	n.a.	−694.4	−513.0	−510.0	−325.4
100	283.2	524.7	459.8	694.4	n.a.	190.4	192.7	429.4
101	85.7	334.0	275.1	513.0	−190.4	n.a.	−20.8	229.9
110	58.1	326.0	254.1	510.0	−192.7	20.8	n.a.	217.9
111	−167.3	125.8	33.9	325.4	−429.4	−229.9	−217.9	n.a.
Relative errors to TI [%]								
000	n.a.	2.0	0.6	4.0	0.7	10.4	29.8	28.7
001	2.0	n.a.	6.6	6.5	1.1	3.5	2.4	5.0
010	1.1	0.0	n.a.	0.0	0.4	3.4	2.7	19.1
011	4.0	6.5	0.0	n.a.	5.2	5.9	4.5	2.2
100	0.7	1.1	0.4	5.2	n.a.	3.3	1.5	3.4
101	10.4	3.5	3.4	5.9	3.3	n.a.	310.0	0.0
110	29.8	2.4	2.7	4.5	3.0	310.0	n.a.	0.0
111	29.7	5.0	19.1	2.2	4.1	0.0	0.0	n.a.

TABLE V Free energy differences as computed with the half-way averaged one-step perturbation (HAOSP) method, Eq. (9), and the relative errors to the thermodynamic integration (TI) results. See also caption of Table II

	000	001	010	011	100	101	110	111
Free energy differences [kJ/mol]								
000	n.a.	336.3	273.1	686.5	−340.8	−23.6	−5.9	331.3
001	−336.3	n.a.	−74.8	280.5	−636.5	−398.1	−318.2	−81.8
010	−273.1	74.8	n.a.	362.5	−567.3	−275.4	−300.8	−5.4
011	−686.5	−280.5	−362.5	n.a.	−873.2	−643.4	−614.2	−368.4
100	340.8	636.5	567.3	873.2	n.a.	275.2	282.6	670.8
101	23.6	398.1	275.4	643.4	−275.2	n.a.	7.8	316.7
110	5.9	318.2	300.8	614.2	−282.6	−7.8	n.a.	309.7
111	−331.3	81.8	5.4	368.4	−670.8	−316.7	−309.7	n.a.
Relative errors to TI [%]								
000	n.a.	33.3	56.0	50.2	19.3	75.0	92.9	155.8
001	33.3	n.a.	1.3	40.0	19.8	14.7	4.8	31.1
010	54.5	8.5	n.a.	31.6	22.7	3.4	14.6	88.1
011	50.2	40.0	31.6	n.a.	19.1	18.0	14.8	15.8
100	19.3	19.8	22.7	19.1	n.a.	49.5	43.9	61.8
101	75.0	14.7	3.4	18.0	49.7	n.a.	20.0	37.4
110	92.9	4.8	14.6	14.8	41.7	20.0	n.a.	42.4
111	157.8	31.1	88.1	15.8	62.6	37.4	42.4	n.a.

smaller, this is not the case. Although the errors are almost all below 100%, there are only few errors below 10%. This shows also in the coefficient of determination: $R^2 = 0.973$. However, all of the signs were correctly calculated.

The reason for the inaccuracy of this method is that the artificially introduced intermediate state at $\lambda = 0.5$ is not well sampled. Unlike in the AOSP, there are no simulations involved using the Hamiltonian of the ensemble at $\lambda = 0.5$. As for the convergence behaviour of the HAOSP method, it suffers from the same problems as the OSP and the AOSP methods; it is hard to determine when the free energy differences have converged. This can be seen in Figure 3.

4.5. Taylor Expansion (TE)

The free energy differences and their relative errors to TI were calculated with the Taylor expansion (TE) up to third order. The results are presented in Tables VI, VII and VIII. For first-order TE, the relative errors to TI are quite large. They range up to over 4000% for the reaction 110 to 101. None of the errors is below 50%. Furthermore, 20 of the signs (out of 56) of the free energy differences are wrong. The results improve slightly when taking into account the second order term. Although most of the errors are still

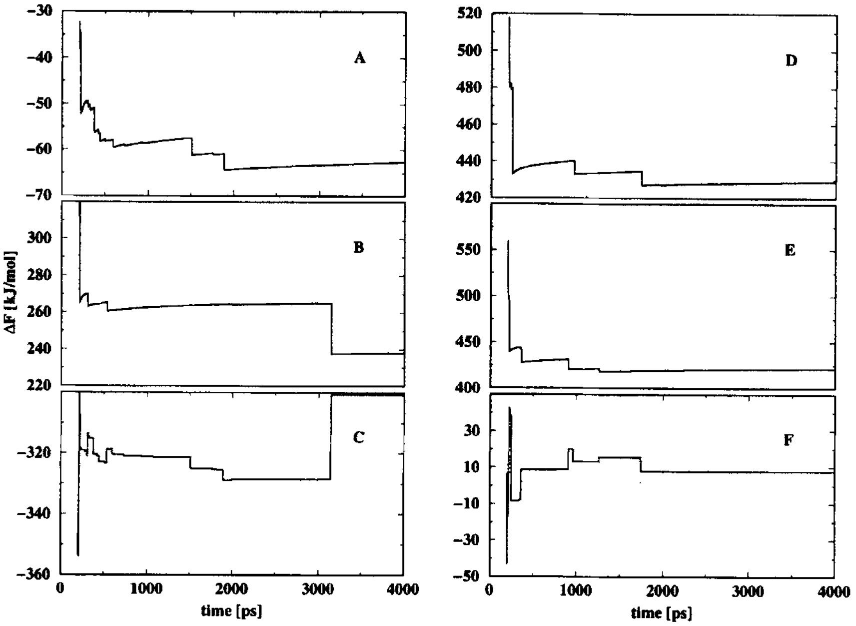

FIGURE 3 Convergence behaviour of the half-way averaged one-step perturbation (HAOSP) method: Panel A shows the convergence of the free energy difference from $\lambda = 0$ to $\lambda = 0.5$ for the reaction 010 to 110, panel B shows the same free energy difference from $\lambda = 1$ to $\lambda = 0.5$ and panel C the convergence of HAOSP for 010 to 110. Panel D shows the convergence of the free energy difference from $\lambda = 0$ to $\lambda = 0.5$ for the reaction 101 to 110, panel E shows the same free energy difference from $\lambda = 1$ to $\lambda = 0.5$ and panel F the convergence of HAOSP for 101 to 110.

TABLE VI Free energy differences as computed with the first-order Taylor expansion (TE) method, Eqs. (10)–(12), and the relative errors to the thermodynamic integration (TI) results. See also caption of Table II

	000	001	010	011	100	101	110	111
Free energy differences [kJ/mol]								
000	n.a.	543.3	472.1	1142.9	−4.8	702.6	675.7	1510.8
001	36.3	n.a.	359.3	486.1	−213.4	−77.8	302.4	600.8
010	109.5	518.7	n.a.	563.9	−101.8	458.4	10.3	725.1
011	313.8	118.5	42.3	n.a.	−123.5	−157.2	−182.2	−63.0
100	506.5	807.6	778.9	1205.9	n.a.	482.4	491.7	1099.8
101	822.5	561.5	967.6	876.3	87.5	n.a.	427.0	508.9
110	747.8	888.1	489.3	808.6	87.2	409.2	n.a.	500.8
111	1184.0	773.6	776.0	540.3	291.0	56.5	60.1	n.a.
Relative errors to TI [%]								
000	n.a.	115.5	169.7	150.4	98.2	831.2	903.6	1070.5
001	114.3	n.a.	572.4	142.5	59.8	77.5	190.4	605.0
010	161.9	531.7	n.a.	104.7	77.9	273.2	104.2	1826.2
011	168.9	159.0	115.3	n.a.	83.2	71.2	65.9	80.4
100	77.5	52.1	68.4	64.6	n.a.	162.0	150.5	165.5
101	756.2	62.1	264.5	60.7	148.1	n.a.	4160.0	120.9
110	789.3	165.6	87.0	51.1	143.7	4190.0	n.a.	130.4
111	1025.0	549.6	1745.2	70.0	170.6	124.8	127.7	n.a.

TABLE VII Free energy differences as computed with the second-order Taylor expansion (TE) method, Eqs. (10)–(12), and the relative errors to the thermodynamic integration (TI) results. See also caption of Table II

	000	001	010	011	100	101	110	111
Free energy differences [kJ/mol]								
000	n.a.	154.1	76.4	202.8	−494.3	−424.9	−467.1	−358.1
001	−277.1	n.a.	−389.0	100.9	−793.1	−580.8	−1163.7	−599.8
010	−205.9	−258.0	n.a.	180.3	−963.8	−778.3	−485.0	−395.0
011	−670.1	−193.5	−274.7	n.a.	−1463.9	−1008.4	−788.2	−578.0
100	253.2	390.5	189.7	396.2	n.a.	97.4	96.6	185.8
101	−73.1	319.1	−181.7	269.7	−219.8	n.a.	−291.5	123.3
110	−18.4	−55.0	239.0	400.3	−225.3	−176.6	n.a.	118.3
111	−387.3	−28.5	−117.8	297.7	−610.7	−244.6	−242.1	n.a.
Relative errors to TI [%]								
000	n.a.	38.9	56.0	55.5	73.0	341.7	455.9	377.5
001	9.5	n.a.	410.5	49.5	49.4	67.6	248.2	403.4
010	16.5	414.6	n.a.	34.2	108.4	193.2	85.4	838.1
011	46.7	3.0	0.0	n.a.	99.9	85.0	47.4	82.0
100	11.2	26.4	58.9	45.9	n.a.	46.7	50.5	55.1
101	176.0	7.8	168.7	50.5	19.1	n.a.	3010.0	46.5
110	121.4	116.5	8.4	25.1	13.1	1660.0	n.a.	45.6
111	201.6	123.5	380.9	6.3	48.1	6.1	11.1	n.a.

TABLE VIII Free energy differences as computed with the third-order Taylor expansion (TE) method, Eqs. (10)–(12), and the relative errors to the thermodynamic integration (TI) results. See also caption of Table II

	000	001	010	011	100	101	110	111
Free energy differences [kJ/mol]								
000	n.a.	528.2	543.3	1342.3	−523.9	−103.2	−208.4	359.8
001	−283.9	n.a.	497.1	630.3	−716.1	−828.8	−460.8	−519.2
010	−224.6	369.6	n.a.	520.9	−656.3	−211.3	−460.3	−66.5
011	−455.4	−149.9	−322.7	n.a.	483.4	−762.9	−499.2	−422.4
100	253.2	928.1	566.9	2228.3	n.a.	575.9	460.3	1174.6
101	168.6	301.0	−64.0	160.6	−193.0	n.a.	988.1	461.0
110	−34.2	2771.3	281.8	946.9	−239.4	417.3	n.a.	517.6
111	−124.6	341.2	327.1	296.0	−845.2	−309.7	−168.8	n.a.
Relative errors to TI [%]								
000	n.a.	109.5	210.3	194.1	83.5	6.2	147.6	178.3
001	12.3	n.a.	754.0	214.5	34.9	139.3	37.7	335.3
010	27.3	350.0	n.a.	89.1	42.0	20.4	75.9	54.8
011	0.2	25.0	17.1	n.a.	166.0	39.8	6.5	32.8
100	11.2	74.9	22.5	204.2	n.a.	212.5	134.7	183.6
101	75.0	13.0	124.2	70.5	4.9	n.a.	9770.0	100.0
110	140.5	729.3	7.3	77.2	20.1	4270.0	n.a.	138.2
111	2.3	185.7	676.2	6.6	104.8	34.4	22.1	n.a.

quite high, there are some reactions for which an almost exact free energy difference is obtained. Most of the errors for reactions with a change in only one protonation site are below 100%. There are only 9 signs wrongly calculated. When going to third-order TE, the results get slightly worse again. The reason for this is, as we will see, the slow convergence for the third-order TE. However, there remain only 5 signs wrongly calculated. The coefficients of determination are $R^2 = 0.595$, $R^2 = 0.825$, and $R^2 = 0.561$ for first, second and third-order TE respectively.

Panels A, B, D, and E of Figure 4 show the convergence behaviour of the TE. We can see that the first and second-order TE converge quite quickly, whereas the convergence behaviour of the third-order TE is much slower.

4.6. Linear Response (LR)

Table IX shows the free energy differences and their relative errors to TI as computed with the linear response (LR) method. All the relative errors to TI are below 60%. Considering the reactions with changes in only one protonation site, the errors are even within about 10%. All the signs are correct. The coefficient of determination is $R^2 = 0.997$.

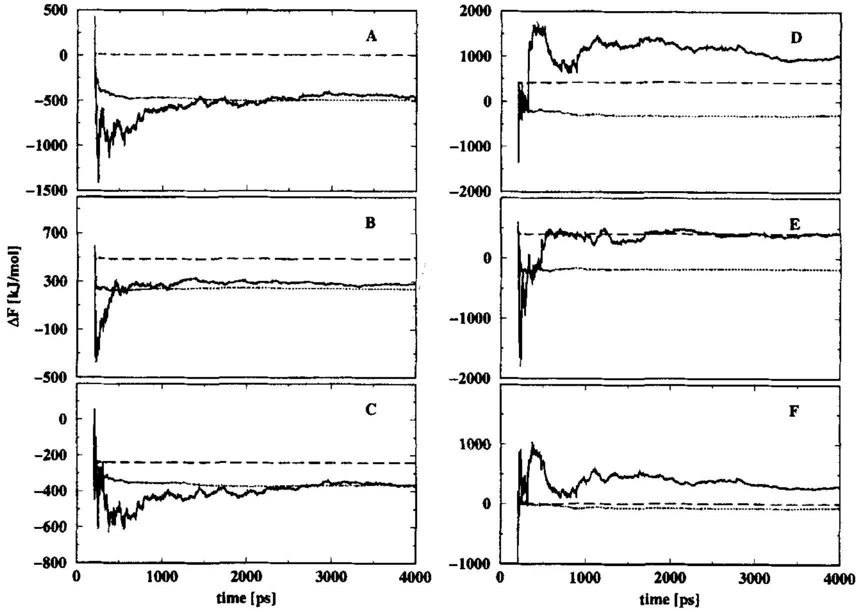

FIGURE 4 Convergence behaviour of free energy differences as calculated with the Taylor expansion (TE) method. The long dashed line represents the first-order, the dotted line the second-order and the solid line the third-order TE. Panel A shows the free energy difference for the reaction 010 to 110, panel B the backwards reaction, and panel C shows the average of the forwards and backwards reaction for 010 to 110. Panel D shows the free energy difference for the reaction 101 to 110, panel E the backwards reaction and panel F shows the average of the forwards and backwards reaction for 101 to 110.

TABLE IX Free energy differences as computed with the linear response (LR) method, Eq. (13), and the relative errors to the thermodynamic integration (TI) results. See also caption of Table II

	000	001	010	011	100	101	110	111
Free energy differences [kJ/mol]								
000	n.a.	253.5	181.3	414.6	−255.7	−59.9	−36.0	163.4
001	−253.5	n.a.	−79.7	183.8	−510.5	−319.7	−292.9	−86.4
010	−181.3	79.7	n.a.	260.8	−440.4	−254.6	−239.5	−25.4
011	−414.6	−183.8	−260.8	n.a.	−664.7	−516.8	−495.4	−301.6
100	255.7	510.5	440.4	664.7	n.a.	197.4	202.3	404.4
101	59.9	319.7	254.6	516.8	−197.4	n.a.	8.9	226.2
110	36.0	292.9	239.5	495.4	−202.3	−8.9	n.a.	220.3
111	−163.4	86.4	25.4	301.6	−404.4	−226.2	−220.3	n.a.
Relative errors to TI [%]								
000	n.a.	0.4	3.4	9.2	10.2	37.5	57.1	25.6
001	0.4	n.a.	4.0	8.0	3.8	7.5	12.3	26.9
010	2.8	2.4	n.a.	5.1	4.5	4.2	8.4	40.5
011	9.2	8.0	5.1	n.a.	9.2	5.1	7.3	5.0
100	10.2	3.8	4.5	9.2	n.a.	7.1	2.5	2.4
101	37.5	7.5	4.2	5.1	7.1	n.a.	10.0	1.7
110	57.1	12.3	8.4	7.3	1.5	10.0	n.a.	0.9
111	27.3	26.9	40.5	5.0	1.7	1.7	0.9	n.a.

The convergence behaviour of the LR method can be studied in panels C and F of Figure 4. As was mentioned in the methods section, the LR is an average of two first-order TE results. We can see that the LR converges very fast: For both reactions shown in Figure 4, the free energy difference hardly changes after about 50 ps. The averages of the second- and third-order TE converge more slowly.

When comparing the average of the first and second-order TE, we can see that the fluctuations of the electrostatic intra-solute and solute–solvent interactions cannot be neglected in the case of protonation and deprotonation reactions: The difference between the averages of the first- and second-order TE is quite large. Since the second-order TE is more accurate than the first-order TE, one would expect that the average of the second-order TE is more accurate than the LR result. This, however, is not the case. The coefficient of determination for the averaged second-order TE is $R^2 = 0.955$. Apparently, not the second-order term, but all higher order terms of the TE almost cancel out on average.

4.7. Thermodynamic Cycles

To test the additivity of the free energy differences, the free energy changes along all 56 thermodynamic cycles between three protonation states were calculated. As a measure of the accuracy of the closure of the cycles, we consider the quotient

$$Q = -\frac{\Delta F(1 \text{ to } 2) + \Delta F(2 \text{ to } 3)}{\Delta F(3 \text{ to } 1)}. \tag{15}$$

The values of Q as well as the mean value $\bar{Q}$ and the standard deviation from the mean value ΔQ are shown in Table X.

TABLE X $Q = -(\Delta F(1 \text{ to } 2) + \Delta F(2 \text{ to } 3))/\Delta F(3 \text{ to } 1)$ for all thermodynamic cycles consisting of three protonation states of the Asp solute. $\bar{Q}$ and ΔQ denote the mean value and the standard deviation from the mean value respectively

Cycle	*TI*	*OSP*	*AOSP*	*HAOSP*	*TE*1	*TE*2	*TE*3	*LR*
000,001,010	1.00	75.28	0.95	0.96	−8.24	−1.14	4.56	0.96
000,001,011	0.99	−15.35	0.99	0.90	−3.28	0.38	2.54	1.05
000,001,100	0.98	−0.09	0.98	0.88	−0.65	2.52	0.74	1.00
000,001,101	0.98	−0.32	1.01	2.62	−0.56	−5.83	1.78	1.10
000,001,110	0.98	−0.81	1.36	−3.07	−1.13	−54.91	1.97	1.09
000,001,111	1.03	−0.81	0.72	0.77	−0.97	−1.15	0.07	1.02
000,010,011	0.98	−15.25	1.02	0.93	−3.30	0.38	2.34	1.07
000,010,100	1.01	−0.11	1.01	0.86	−0.73	3.50	0.45	1.01
000,010,101	0.94	−0.91	1.19	0.10	−1.13	−9.60	−1.97	1.22
000,010,110	1.03	−0.38	1.39	4.69	−0.65	−22.22	2.43	1.61
000,010,111	1.03	−0.95	0.83	0.81	−1.01	−0.82	3.83	0.95
000,011,100	0.97	−1.37	0.90	0.55	−2.01	4.98	−7.21	0.98
000,011,101	0.92	−1.00	0.87	−1.83	−1.20	−11.01	−3.44	1.71
000,011,110	0.93	−1.07	1.23	−12.25	−1.28	−31.84	24.65	2.24
000,011,111	1.08	−0.81	0.68	0.96	−0.91	−0.97	7.38	0.69
000,100,101	1.05	−0.32	1.08	2.78	−0.58	−5.43	−0.31	0.97
000,100,110	1.06	−0.36	1.56	9.86	−0.65	−21.63	−1.86	1.48
000,100,111	1.00	−0.81	0.87	1.00	−0.92	−0.80	5.22	0.91
000,101,110	1.02	−1.14	1.83	2.68	−1.51	−38.96	25.88	1.42
000,101,111	1.05	−0.90	0.86	0.88	−1.02	−0.78	2.87	1.02
000,110,111	1.04	−0.90	0.95	0.92	−0.99	−0.90	2.48	1.13
001,010,011	0.99	77.53	1.03	1.02	−7.79	−1.08	6.79	0.98
001,010,100	1.01	0.16	1.03	1.01	−0.32	3.46	0.17	1.02
001,010,101	0.99	−0.86	1.07	0.88	−1.46	3.66	−0.95	1.04
001,010,110	1.01	−0.08	1.03	1.18	−0.42	−15.90	−0.01	1.09
001,010,111	1.00	−1.22	0.92	0.98	−1.40	−27.48	−1.26	1.22
001,011,100	1.00	0.01	0.97	0.93	−0.45	3.49	−1.20	0.94
001,011,101	0.99	−0.05	0.97	0.91	−0.58	2.84	0.44	1.04
001,011,110	1.00	−0.04	0.99	1.05	−0.34	−12.51	−0.05	1.06
001,011,111	0.98	−0.27	1.10	1.07	−0.55	−16.73	−0.61	1.36
001,100,101	1.00	0.05	1.00	0.91	−0.48	2.18	0.47	0.98
001,100,110	1.00	0.02	1.02	1.11	−0.31	−12.67	0.09	1.05
001,100,111	0.97	−0.90	0.76	−0.42	−1.15	−21.29	−1.34	1.23
001,101,110	1.00	0.01	1.09	1.23	−0.39	−15.87	−0.06	1.06
001,101,111	0.97	−0.28	0.83	0.99	−0.56	−16.03	1.08	1.08
001,110,111	0.98	−0.70	0.86	0.10	−1.04	−36.64	−0.17	0.84
010,011,100	0.99	−0.10	0.91	0.90	−0.56	6.76	−1.77	0.92
010,011,101	1.02	−0.12	0.87	1.02	−0.42	−4.56	−3.78	1.00
010,011,110	0.99	−0.25	0.93	0.84	−0.78	2.54	−0.08	0.98

TABLE X (Continued)

Cycle	*TI*	*OSP*	*AOSP*	*HAOSP*	*TE*1	*TE*2	*TE*3	*LR*
010,011,111	1.01	−0.38	1.51	1.09	−0.65	−3.38	−0.30	1.60
010,100,101	1.05	−0.07	0.98	1.06	−0.39	−4.77	−1.26	0.95
010,100,110	1.01	−0.17	1.05	0.95	−0.80	3.63	0.69	0.99
010,100,111	1.12	−1.04	0.90	−19.17	−1.29	−6.61	−1.58	1.41
010,101,110	0.97	−1.08	1.16	0.89	−1.81	4.48	−2.76	1.02
010,101,111	0.82	−1.02	1.34	−7.65	−1.25	−5.56	−0.76	1.12
010,110,111	1.04	−0.41	1.07	−1.66	−0.66	−3.11	−0.17	0.75
011,100,101	1.00	0.03	0.98	0.93	−0.41	5.07	−6.60	0.90
011,100,110	1.00	0.02	0.98	0.96	−0.46	3.42	−1.00	0.93
011,100,111	1.00	−1.17	0.81	0.55	−1.81	4.29	−5.60	0.86
011,101,110	1.00	0.19	1.05	1.03	−0.33	3.25	−0.24	1.02
011,101,111	0.99	−0.09	0.87	0.89	−0.65	2.97	1.02	0.96
011,110,111	1.00	−0.05	0.90	0.83	−0.59	2.25	−0.06	0.91
100,101,110	0.98	21.05	0.88	1.00	−10.43	−0.86	6.53	1.02
100,101,111	1.00	−36.49	0.98	0.88	−3.41	0.36	1.23	1.05
100,110,111	1.00	−36.01	0.96	0.88	−3.41	0.35	1.16	1.04
101,110,111	0.99	7.32	0.86	1.00	−16.41	−0.71	4.86	1.01
$\bar{Q}$	1.00	0.98	1.02	0.29	−1.72	−6.20	1.20	1.09
ΔQ	0.01	2.22	0.03	0.50	0.37	1.70	0.73	0.03

It is not surprising that the methods that yield accurate results for the free energy differences are also accurate in terms of additivity. The best results in this respect are obtained from TI, AOSP, and LR. For the OSP, $\bar{Q}$ is almost correct, but ΔQ is very large. However, this could be an indication that there is a systematic sampling problem involved in the OSP method that makes the AOSP method accurate. As for the TE method, the additivity of the free energy differences of the first-order TE is much better than the one of the second-order TE. This might explain why LR is more accurate than the averaged second-order TE.

5. CONCLUSIONS

We have compared seven fast but approximate methods to compute free energy differences of chemical reactions by means of molecular dynamics (MD) for the case of protonation and deprotonation reactions of aspartic acid in aqueous solution. As reference values for the free energy differences, the values calculated with thermodynamic integration (TI) were taken.

It can be concluded that none of the approximate methods involving only one MD trajectory (one-step perturbation (OSP) and Taylor expansion (TE)) yield accurate results. From the methods involving two MD trajectories (averaged one-step perturbation (AOSP), half-way averaged one-step perturbation (HAOSP) and linear response (LR)), the AOSP and

LR both proved to be accurate, whereas the HAOSP was less accurate than AOSP and LR.

The fastest convergence was achieved with the first-order TE and LR methods: The free energy differences have converged after about 100 ps. The second-order TE converges after about 100 to 800 ps. The AOSP and third-order TE converge much more slowly: For single cases, the sampling time of 3.8 ns was not long enough for the free energy difference to converge. The OSP, AOPS and HAOSP are unpredictable in terms of convergence, as there are many sudden changes in their convergence curves that do not become smaller with a long simulation time.

Considering both aspects, accuracy and convergence, the linear response (LR) method is the best method tested to calculate free energy differences of protonation and deprotonation reactions, at least for small molecules.

References

[1] van Gunsteren, W. F., Beutler, T. C., Fraternali, F., King, P. M., Mark, A. E. and Smith, P. E. (1993). "Computation of free energy in practice: choice of approximations and accuracy limiting factors", In: van Gunsteren, W. F., Weiner, P. K. and Wilkinson, A. J. (Eds.), *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*, **2**, 315–348. ESCOM Science Publishers, Leiden, The Netherlands.

[2] Allen, M. P. and Tildesley, D. J. (1987). *Computer Simulation of Liquids*, Clarendon Press, Oxford.

[3] Mezei, M. and Beveridge, D. L. (1986). "Free energy simulations", *Annals New York Acad. Sci.*, **482**, 1.

[4] Beveridge, D. L. and DiCapua, F. M. (1989). "Free-energy *via* molecular simulation – applications to chemical and biomolecular systems", *Annu. Rev. Biophys. Biophys. Chem.*, **18**, 431.

[5] Straatsma, T. P. and McCammon, J. A. (1992). "Computational alchemy", *Annu. Rev. Phys. Chem.*, **43**, 407.

[6] Mark, A. E. (1998). "Free energy perturbation calculations", In: von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer III, H. F. and Schreiner, P. R. (Eds.), *Encyclopaedia of Computational Chemistry*, **2**, 1070–1083. Wiley, Chichester.

[7] King, P. M. (1993). "Free energy *via* molecular simulation: A primer", In: van Gunsteren, W. F., Weiner, P. K. and Wilkinson, A. J. (Eds.), *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*, **2**, 267–314. Escom Science Publisher, Leiden.

[8] King, G. and Barford, R. A. (1993). "Calculation of electrostatic free-energy differences with a time-saving approximate method", *J. Phys. Chem.*, **97**, 8798.

[9] Levy, R. M., Belhadj, M. and Kitchen, D. B. (1991). "Gaussian fluctuation formula for electrostatic free-energy changes in solution", *J. Chem. Phys.*, **95**, 3627.

[10] Smith, P. E. and van Gunsteren, W. F. (1994). "Predictions of free energy differences from a single simulation of the initial state", *J. Chem. Phys.*, **100**, 577.

[11] Åqvist, J., Medina, C. and Samuelsson, J.-E. (1994). "A new method for predicting binding affinity in computer-aided drug design", *Protein Eng.*, **7**, 385.

[12] Hummer, G. and Szabo, A. (1996). "Calculation of free-energy differences from computer simulations of initial and final states", *J. Chem. Phys.*, **105**, 2004.

[13] Kong, X. J. and Brooks III, C. L. (1996). "λ-dynamics: A new approach to free energy calculations", *J. Chem. Phys.*, **105**, 2414.

[14] Liu, H., Mark, A. E. and van Gunsteren, W. F. (1996). "Estimating the relative free energy of different molecular states with respect to a single reference state", *J. Phys. Chem.*, **100**, 9485.

[15] Schäfer, H., van Gunsteren, W. F. and Mark, A. E. (1999). "Estimating relative free energies from a single ensemble: Hydration free energies", *J. Comput. Chem.*, **20**, 1604.

[16] Carlson, H. A. and Jörgensen, W. L. (1999). "An extended linear-response method for determining free-energies of hydration", *J. Phys. Chem.*, **99**, 10667.

[17] Daura, X., Hünenberger, P. H., Mark, A. E., Querol, E., Avilés, F. X. and van Gunsteren, W. F. (1996). "Free energies of transfer of Trp analogs from chloroform to water: Comparison of theory and experiment and the importance of adequate treatment of electrostatic and internal interactions", *J. Am. Chem. Soc.*, **118**, 6285.

[18] McDonald, N. A., Carlson, H. A. and Jörgensen, W. L. (1997). "Free energies of solvation in chloroform and water from a linear response approach", *J. Phys. Org. Chem.*, **10**, 563.

[19] Guo, Z. Y. and Brooks III, C. L. (1998). "Rapid screening of binding affinities: Application of the λ-dynamics method to a trypsin-inhibitor system", *J. Am. Chem. Soc.*, **120**, 1920.

[20] Chen, X. and Tropsha, A. (1999). "Generalized linear response method: Application to hydration free energy calculations", *J. Comput. Chem.*, **20**, 749.

[21] Kirkwood, J. G. (1935). "Statistical mechanics of fluid mixtures", *J. Chem. Phys.*, **3**, 300.

[22] Zwanzig, R. W. (1954). "High-temperature equation of state by a perturbation method. I. Nonpolar gases", *J. Chem. Phys.*, **22**, 1420.

[23] Postma, J. P. M., Berendsen, H. J. C. and Haak, J. R. (1982). "Thermodynamics of cavity formation in water: A molecular dynamics study", *Faraday Symp. Chem. Soc.*, **17**, 55.

[24] van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P., Mark, A. E., Scott, W. R. P. and Tironi, I. G. (1996). *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Zürich.

[25] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F. and Hermans, J. (1981). "Interaction models for water in relation to protein hydration", In: Pullman, B. (Ed.), *Intermolecular Forces*, pp. 331–342. Reidel, Dordrecht.

[26] Scott, W. R. P., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Krüger, P. and van Gunsteren, W. F. (1999). "The GROMOS biomolecular simulation program package", *J. Phys. Chem. A*, **103**, 3596.

[27] Tironi, I. G., Sperb, R., Smith, P. E. and van Gunsteren, W. F. (1995). "A generalized reaction field method for molecular dynamics simulations", *J. Chem. Phys.*, **102**, 5451.

[28] Smith, P. E. and van Gunsteren, W. F. (1995). "Reaction field effects on the simulated properties of liquid water", *Mol. Simulation*, **15**, 233.

[29] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. and Haak, J. R. (1984). "Molecular dynamics with coupling to an external bath", *J. Chem. Phys.*, **81**, 3684.

[30] Beutler, T. C., Mark, A. E., van Schaik, R. C., Gerber, P. R. and van Gunsteren, W. F. (1994). "Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations", *Chem. Phys. Letters*, **22**, 529.