ELSEVIER

Journal of Molecular Structure (Theochem) 574 (2001) 39–45

THEO CHEM

www.elsevier.com/locate/theochem

Arginine–phosphate salt bridges in protein–DNA complexes: a Car–Parrinello study

D. Frigyes[a,1], F. Alber[b], S. Pongor[a], P. Carloni[a,*]

[a]*International Center for Genetic Engineering and Biotechnology, AREA Science Park, Padriciano 99, 34012 Trieste, Italy*
[b]*International School for Advanced Studies (SISSA) and Istituto Nazionale di Fisica della Materia (INFM), Via Beirut 2-4, 34014 Trieste, Italy*

Received 2 October 2000; revised 8 January 2001; accepted 9 January 2001

Abstract

We present a gradient-corrected density functional (DFT)-based molecular dynamics (MD) study on hydration and dynamics of arginine–phosphate adducts in proteins. Calculations are carried out using the BLYP recipe for the exchange-correlation functional. We focus on two representative H-bond patterns found in protein–DNA complexes in the presence and in the absence of water molecules. Our structural models include methylguanidinium (representing the arginine side chain), dimethylphosphate (representing the phosphate moiety in DNA) and water molecules H-bonding the complex. Our DFT-MD simulations, carried out at room temperature, point to hydration as a possible key factor for molecular recognition of the bidentate complex. Furthermore they suggest that hydration is accompanied by significant polarization effects.

Keywords: Arginine–phosphate salt bridges; Protein–DNA complexes; Car–Parrinello study

1. Introduction

Protein–DNA interactions take part in a variety of fundamental biological phenomena, from packaging to replication of DNA, from gene-expression to recombination [1–3]. In these complexes, salt bridges between DNA phosphate and positively charged protein residues (arginine and, to a lesser extent [4], lysine) are always present and play an important role in the complex stability [3].

In spite of the relevance of arginine–phosphate recognition in protein–DNA complexes, few quantum theoretical studies have appeared so far in the literature. These investigations provide structural properties and energetical aspects of *bidentate* patterns (Fig. 1a) in quantum-chemical model complexes at Hartree–Fock (HF) [5,6] and semi-empirical [6–8] levels of theory.

Here we present an ab initio molecular dynamics study on two typical examples of mono- and bidentate complexes (Fig. 1a and b). In contrast to previous work, our investigation includes explicit treatment of solvation and finite temperature effects, which are crucial for molecular recognition processes in protein–DNA interactions [9].

The simulations are based on the Car–Parrinello (CP) method [10]. Along with the BLYP exchange-correlation functional [11,12], the CP approach is a powerful tool in describing structure and dynamics of H-bonded systems, such as liquid water [13–15],

* Corresponding author. Tel.: +39-040-3787-463; fax: +39-040-3787-528.
E-mail address: carloni@sissa.it (P. Carloni).
[1] Present address: Department of General and Inorganic Chemistry, Eötvös Loránd University, P.O. Box 32, H-1518 Budapest 112, Hungary.

0166-1280/01/$ - see front matter
PII: S0166-1280(01)00368-2

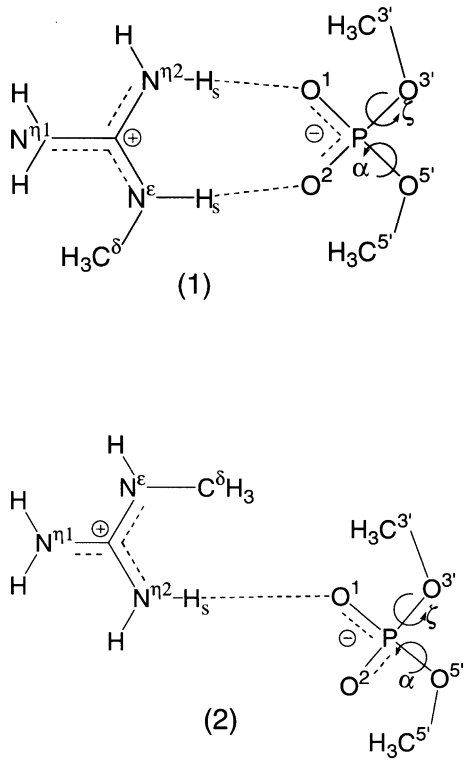

Fig. 1. Arginine–phosphate H-bond patterns in protein–DNA interactions [25]: (a) bidentate and (b) monodentate complexes, modeled as MGD and DMP. Hydrogen atoms taking part in salt bridges are denoted H_s.

H-bonded aqueous solutions [16,17], and a variety of biological systems [18–22] at room temperature. Thus the CP method appears ideally suited to treat these complexes.

In this work, we have identified two patterns, which are most likely fully hydrated in the real macromolecular complexes. This is a common feature of the protein–DNA complexes whose structure have been solved so far [23,24]. Then, we carry out simulations for the complexes in which hydration is fully taken into account.

2. Computational details

2.1. Model systems

Two structural models were considered. They involve the energetically most stable bidentate (**1**) (Fig. 1a) and the most frequently occurring monodentate (**2**) (Fig. 1b) arginine–phosphate interaction patterns [25]. Arginine and DNA–phosphate were modeled as methylguanidinium (MGD) [5,6,26–28] and dimethylphosphate (DMP) [5,8], respectively. Inspection of the Protein Data Bank (PDB) [23] pointed to the bidentate complex in PUT3 transcription factor–DNA complex [29] (PDB code: 1zme) (Fig. 2a) and the monodentate complex in the StRE–SREBP-1a complex [30] (PDB code: 1am9) (Fig. 2c). In these X-ray structures, arginine–phosphate terminal groups are putatively fully hydrated, although most waters could not be detected by the X-ray diffraction experiment [29,30]. The model complexes are displayed in Fig. 2. **1** is constructed from Arg C38 and Cyt B14 (Fig. 2b) of the PUT3 transcription factor–DNA complex (PDB code: 1zme [29]); **2** is constructed from Arg B334 and Gua G44 of the human StRE–SREBP-1a complex (Fig. 2d) (PDB code: 1am9 [30]). Hydrogen atoms were added assuming standard bond lengths and angles. Initial water positions were determined by performing classical molecular dynamics simulations using the Cerius2 program package [31]. MGD and DMP were constrained to their crystallographic positions and immersed in a $9 \times 15.6 \times 14.7$ Å^3 simulation box containing 57 water molecules. We used the DREIDING force field [32]. For each of the two models, one force-field-based MD simulation was performed for 150 ps at 300 K. Constant temperature simulation was achieved by coupling the system to a Nosé–Hoover thermostat [33]. Periodic boundary conditions were applied and the electrostatic interactions being calculated with the Ewald-sum method [34]. Only the water molecules H-bonding to the MGD–DMP complex at the last snapshot from the classical MD (i.e. after 150 ps) were included in our quantum chemical models (**1** and **2**, Fig. 1). The applied H-bonding criteria concerning the donor (D) and acceptor (A) atoms were $d(\mathrm{A}\cdots\mathrm{H}) < 2.5$ Å and $\mathrm{D–H}\cdots\mathrm{A} > \angle 90°$. We did not include all the water molecules as full solvation would not correctly model the rather hydrophobic environment around the terminal CH_3 groups.

This procedure led to the inclusion of seven water molecules in model **1** and eight water molecules in model **2** as a starting point in our ab initio MD simulations.

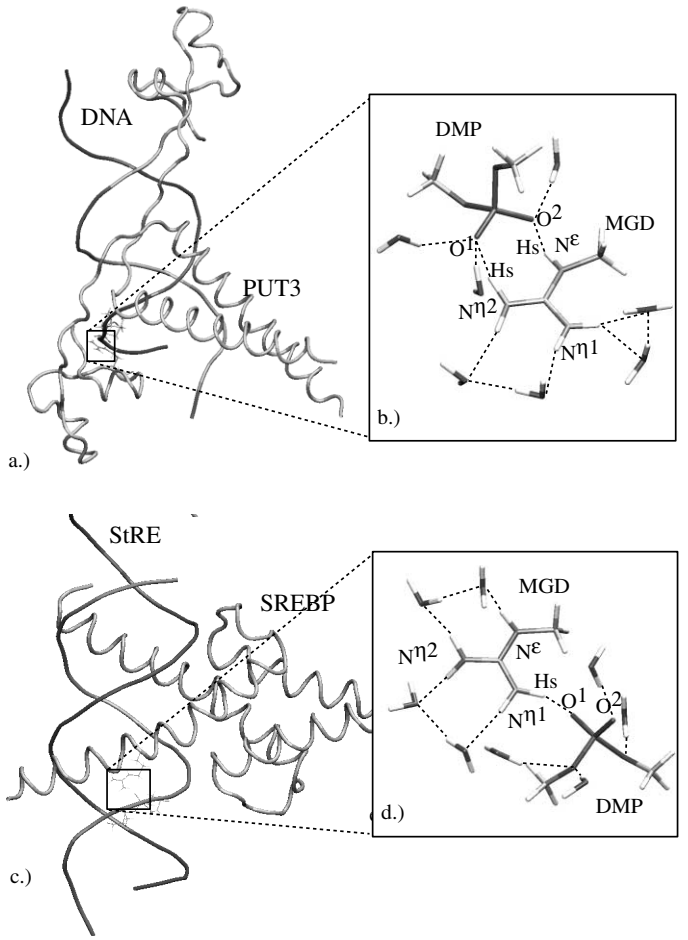

Fig. 2. Protein–DNA complexes and the corresponding model systems used in the molecular dynamics simulations: (a) PUT3 transcription factor bound to DNA [29]; (b) initial configuration of model **1** used in the quantum chemical calculations; (c) SREBP–StRE bound to DNA [30]; (d) initial model **2** used in our quantum chemical calculation. Protein cartoon is made with the VMD program package [48].

Electronic properties were calculated on representative sets of snapshots (10 structures per model) taken from the trajectory of our ab initio molecular dynamics of **1** and **2**. The corresponding complexes *without* water molecules were also considered in our quantum-chemical calculations. These complexes were used to evaluate changes in structural and electronic properties upon hydration.

2.2. Computational methodology

Car–Parrinello-MD [10] simulations for complexes **1** and **2** has been carried out using the CPMD program by Hutter et al. [35]. The electronic structure was described in the density functional approach using the BLYP [11,12] gradient-corrected exchange-correlational functional. Complexes (**1**) and (**2**) with and without water were inserted in a box of $17 \times 13 \times 14$ and $16.7 \times 12.5 \times 12.5$ Å^3, respectively. All the complexes were treated as isolated systems [36]. A plane-wave basis set was used up to an energy cutoff of 70 Ry. The interactions between the ionic cores and valence electrons were described by Troullier–Martins type pseudopotentials [37]. The timestep was 0.121 fs and the fictitious electronic mass was set to 1000 au. The simulations were performed for 1.5 ps at 300 K coupling the system to a Nosé–Hoover thermostat [33,38,39]. In both complexes,

position constraints were applied on the C_δ and $C^{3'}$ and $C^{5'}$ methyl carbon atoms of MGD and DMP, respectively, in order to mimic the rigidity of protein and DNA backbones. All calculations were performed on a CRAY T3e and SP3 parallel computers.

2.3. Calculated properties

Hydration was analyzed in terms of coordination numbers derived from a radial distribution function [40] for the O_{DMP}–H_{water} and H_{MGD}–O_{water} contacts with a cutoff of 2.5 Å. The effect of solvation on the electronic structure was estimated by (i) the electron density differences ($\Delta\rho$) defined as:

$$\Delta\rho(r) = \rho^{cpw}(r) - \rho^{cp}(r) - \rho^{w}(r),$$

where $\rho^{cpw}(r)$ denotes the charge density of the entire complex; $\rho^{cp}(r)$ of the complex without water and $\rho^{w}(r)$ the charge density of the corresponding water structure; (ii) by the geometrical analysis of the centers of maximally localized Wannier orbitals (WOC) [41,42]. The displacement of WOCs caused by solvent effects will be used as a measure of bond polarity changes [42]. Our analysis follows the procedure in Ref. [42] by calculating the bond ionicity indices (BI), which correlate to changes in Pauling electronegativity in a series of XH_n compounds (X = Mg,H,P,C,N,Cl,O). The higher the absolute value of BI, the more polarized is the corresponding bond. Here, we have calculated the BI_{NH} values for the N–H bonds at four snapshots along the trajectory (t = 0.36, 0.72, 1.08, 1.42 ps) of **1** and **2**, with and without the inclusion of water molecules.

3. Results

MD-averaged structural properties (such as N–O distances and phosphate α and ζ torsion angles) of complex **1** is in fairly good agreement with the X-ray structure (Table 1). In contrast, complex **2** exhibits small discrepancies with the experimental data.

Hydration is analyzed in terms of water coordination numbers. Let us focus first on complex **1**. Phosphate O^1 and arginine $N^{\eta 2}$–H are fully solvated, as indicated by the water coordination numbers reported in Fig. 3a. In particular, a stable structural motif involving the $N^{\eta 2}$–H···O^1=P contact and three water molecules (Fig. 3b) is formed after 0.5 ps and it is kept during the rest of the simulation. The corresponding N_ε–H···O^2=P hydration pattern cannot be formed because of the presence of a methyl group bound to N_ε. Indeed, water H-bonds are relatively weak and phosphate O^2 is only partially hydrated (see again Fig. 3a).

A plot of the electronic density difference $\Delta\rho$, which describes charge redistribution upon hydration of the complex, shows that the solvent significantly polarizes arginine and the phosphate oxygens interacting with it (Fig. 4). In particular, the electron density shifts along the NH_s bonds, which interact with the phosphate, from nitrogen towards hydrogen, suggesting that the polarity of these bonds decreases in the presence of water (similar results have been obtained for other snapshots of the dynamics). Consistently, the calculated Pauling electronegativities [42] ($\Delta\chi$) of the bond, which are estimated from bond ionicity indexes (see Section 2), *decrease* on passing from the gas phase to hydrated complex (Table 1). This suggests that the NH_s bonds are more polarized towards the hydrogen atom and therefore the NH_s bonds are less polar in the presence of water. In contrast, NH bonds interacting with water ($N^{\eta 1}$–H, $N^{\eta 2}$–H in Fig. 3) exhibit *reverse* polarization (Table 1).

We now turn our attention to complex **2**. In this complex, Arg binds phosphate in monodentate fashion. As expected, the MD-averaged Arg–phosphate H-bond length is rather similar to those in the bidentate complex (Table 1). The very short H-bond length in the starting X-ray structure is most likely to be an artifact of the X-ray refinement, as such strong H-bonds are very rarely encountered in biological systems [9].

Phosphate interacts in **2** more extensively with the solvent than in **1** (Fig. 3c); in particular, phosphodiester oxygen $O^{5'}$ forms weak H-bond interactions with a water molecule, which also interacts with O^2 (Fig. 3c). However, the difference density plot does not exhibit significant differences with that of **1**.

4. Discussion and conclusions

We have investigated the effects of aqueous solvation on stability and electronic structure of arginine–phosphate interactions occurring in protein–DNA complexes.

Table 1
Selected averaged properties (d — bond length, ζ,α-torsion angles, BI_{NH} — bond ionicity indices, $\Delta\chi_{NH}$ — differences in Pauling electronegativity) of model **1** and **2**. Standard deviations are given in parentheses, experimental data are in square brackets. Values in curly brackets correspond to data derived by vacuo calculations without water solvent. Notations correspond to those in Fig. 1

	1	**2**
$d(OH_s)$ (Å)	1.78(0.19)	1.67(0.10)
$d(NH_s)$ (Å)	1.05(0.03)	1.06(0.03)
$BI_{NH_s}(N)$	−0.146(0.014)	−0.156(0.030)
	{−0.157(0.014)}	{−0.168(0.029)}
$BI_{NH}(N)$	−0.091(0.013)	−0.088(0.011)
	{−0.079(0.008)}	{−0.079(0.006)}
$\Delta\chi_{NH_s}$	1.364(0.131) {1.467 (0.131)}	1.458(0.280) {1.570(0.271)}
$\Delta\chi_{NH}$	0.850(0.121) {0.738(0.075)}	0.822(0.103) {0.738(0.056)}
$d(N_\varepsilon–O)$ (Å)	2.70(0.08) [2.71]	
$d(N_\eta–O)$ (Å)	2.89(0.15) [2.77]	2.71(0.09) [2.41]
ζ	−164(13)° [−152°]	−73(19)° [−85°]
α	−60(8)° [−43°]	−74(18)° [−79°]

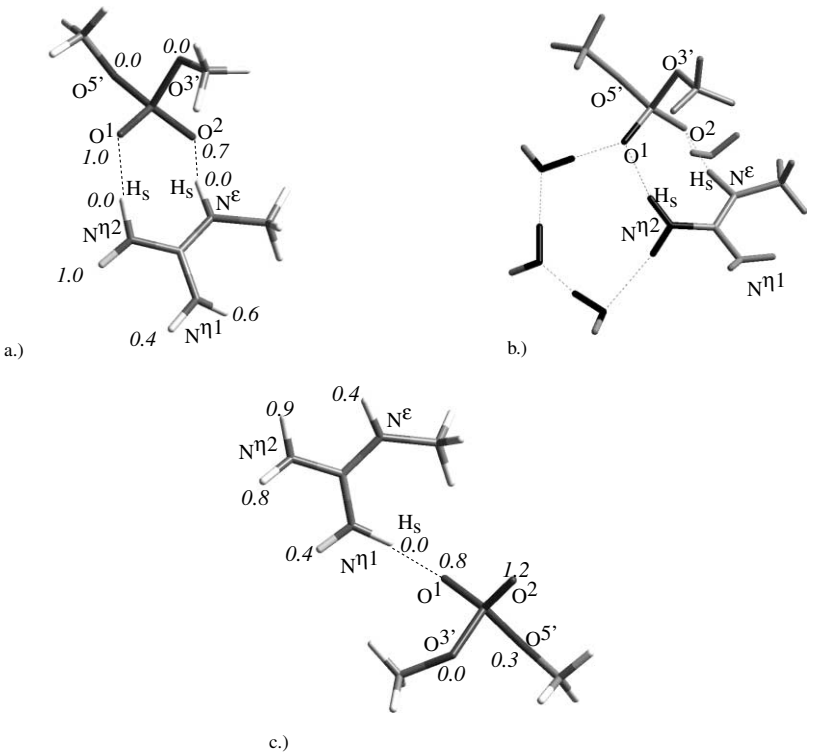

Fig. 3. Solvation of the arginine–phosphate complexes: (a) water coordination numbers of selected MGD and DMP atoms in **1** (water molecules were not drawn for the sake of clarity) typical structure of a cyclic water–MGD–DMP cluster (black) found in the simulation of **1** (only those water molecules, H-bonding to the complex are shown), (c) water coordination numbers of selected MGD and DMP atoms in **2**.

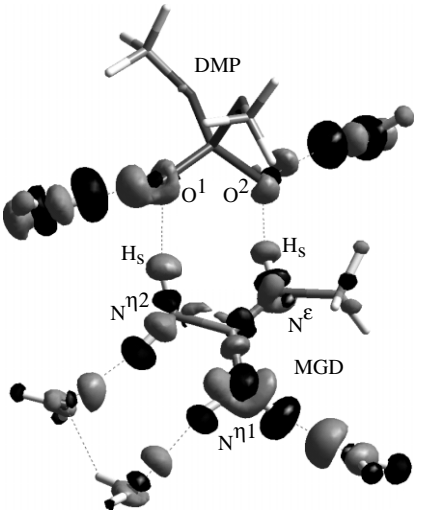

Fig. 4. Isodensity contour plot of the electron density difference $\Delta\rho$ in **1**. Black, — 0.002 e/a.u.^{3}; gray, 0.002 e/a.u.^{3}. Only those water molecules H-bonding to the complex are shown.

Solvation appears to be an essential ingredient to describe these complexes, fully confirming and providing further evidence of previous calculations [6,46,47].

Our calculations reveal that a cluster of water molecules connecting the phosphate and the arginine moieties is formed and is stable during the dynamics in complex **1** (Fig. 3b), suggesting that continuum solvation models, included in a quantum-mechanical calculation of arginine–phosphate adducts, may not be appropriate for describing hydration phenomena. Inspection of the PDB database [23] shows that this peculiar H-bonded complex is present in a variety of protein–DNA complexes, from the retinoic acid repressor [43] — to Cre recombinase [44] — and to *E. coli* Trp repressor [45]. This indicates that this pattern may be important for the molecular recognition of the two moieties.

In the bidentate complex **1**, hydration significantly increases the polarity of the arginine NH bonds interacting with water. This is probably the result of the electrostatic interactions between water oxygens and arginine hydrogens, which become more positively charged in the hydrated complex (Fig. 5). In contrast, the effect of the NH_s bonds (which interact with phosphate) is the opposite. This may be due to the fact that the charge on the phosphate oxygens O1 or O2 is screened by the presence of the solvent (Fig. 5). However, these polarization effects might be

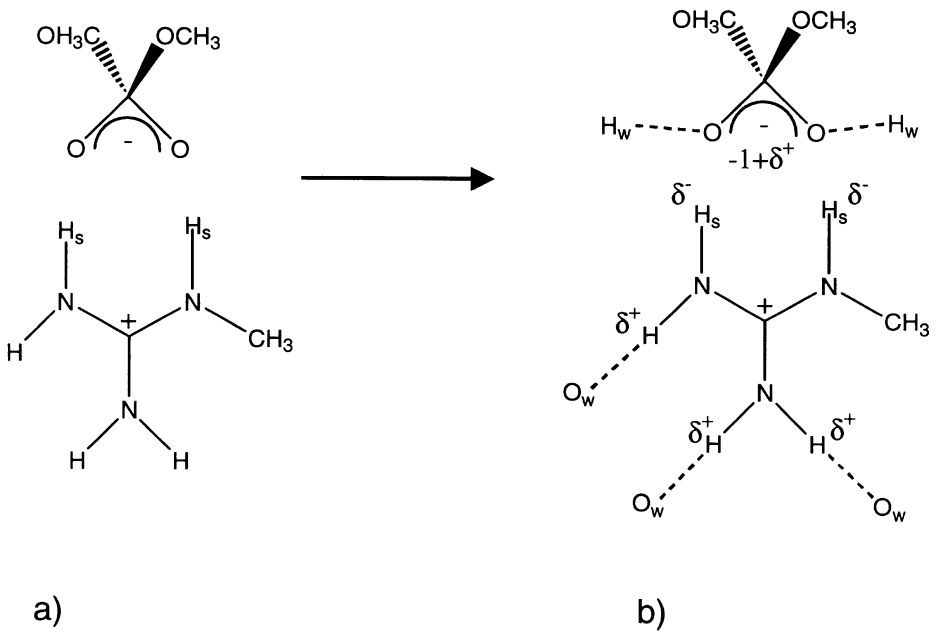

Fig. 5. Schematic view on the polarization effects upon solvation observed in the density difference plot of **1**.

overestimated by the use of a relatively small, isolated complex instead of the real macromolecular complex.

Standard force-fields for biological simulations, which do not include polarization effects, may encounter difficulties in correctly describing hydration phenomena in these systems.

Acknowledgements

This project was supported by an INFM parallel computational grant (project number: P0010/1999). The help of Katalin Nadassy in selection and handling of protein–DNA 3D structures is gratefully acknowledged. D. Frigyes is a recipient of an ICGEB postdoctoral fellowship.

References

[1] A.A. Travers, DNA–Protein Interactions, Chapman and Hall, London, 1993.
[2] C. Branden, J. Tooze, Introduction to Protein Structure, Garland Publishing/Taylor and Francis, New York, 1999.
[3] G.M. Blackburn, M.J. Gait, Nucleic Acids in Chemistry and Biology, IRL Press at Oxford University Press, Oxford, 1990.
[4] Y. Mandel-Gutfreund, O. Schueler, H. Margalit, J. Mol. Biol. 253 (1995) 370.
[5] D.W. Deerfield, H.B. Nicholas Jr., R.G. Hiskey, L.G. Pedersen, Proteins 6 (1989) 168.
[6] J. Mavri, H.J. Vogel, Proteins 24 (1996) 495.
[7] N.V. Kumar, G. Govil, Biopolymers 23 (1984) 1995.
[8] N. Gresh, B. Pullman, Theor. Chim. Acta 52 (1979) 67.
[9] G.A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structures, Springer, Berlin, 1991.
[10] R. Car, M. Parrinello, Phys. Rev. Lett. 55 (1985) 2471.
[11] A. Becke, Phys. Rev. A 38 (1988) 3098.
[12] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.
[13] M. Sprik, J. Hutter, M. Parrinello, J. Chem. Phys. 105 (1996) 1142.
[14] P.L. Silvestrelli, M. Bernasconi, M. Parrinello, Chem. Phys. Lett. 277 (1997) 478.
[15] D. Marx, M.E. Tuckerman, J. Hutter, M. Parrinello, Nature 397 (1999) 601.
[16] M. Sprik, J. Phys.: Condens. Matter 8 (1996) 9405.
[17] F. Alber, G. Folkers, P. Carloni, J. Mol. Struct. (Theochem) 489 (1999) 237.
[18] L. De Santis, P. Carloni, Proteins 37 (1999) 611.
[19] C. Rovira, M. Parrinello, Biophys. J. 78 (2000) 93.
[20] C. Rovira, M. Parrinello, Chem. Eur. J. 5 (1999) 250.
[21] U. Roethlisberger, P. Carloni, Int. J. Quantum Chem. 73 (1999) 209.
[22] S. Piana, P. Carloni, Proteins 39 (2000) 26.
[23] F.C. Bernstein, T.F. Koetzle, J.B. Williams Jr., E.F. Meyer, M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, J. Mol. Biol. 112 (1977) 535.
[24] K. Nadassy, S.J. Wodak, J. Janin, Biochemistry 38 (1999) 1999.
[25] B. Raman, C. Guarnaccia, K. Nadassy, S. Zahariev, C. Acatrinei, G. Pongor, S. Pongor, submitted for publication.
[26] P.G. Mezey, J.J. Ladik, S. Suhai, Theor. Chim. Acta 51 (1979) 323.
[27] J.B.O. Mitchell, J.M. Thornton, J. Singh, S.L. Price, J. Mol. Biol. 226 (1992) 251.
[28] X. Barril, C. Alemán, M. Orozco, F.J. Luque, Proteins 32 (1998) 67.
[29] K. Swaminathan, P. Flynn, R.J. Reece, R. Marmorstein, Nat. Struct. Biol. 4 (1997) 751.
[30] A. Parraga, L. Bellsolell, A.R. Ferre-D'Amare, S.K. Burley, Structure 6 (1998) 661.
[31] Molecular Simulations Inc. Cerius2, 1997.
[32] S.L. Mayo, B.D. Olafson, W.A. Goddard III, J. Phys. Chem. 94 (1990) 8897.
[33] S. Nosé, Mol. Phys. 52 (1984) 255.
[34] D. Frenkel, B. Smit, Understanding Molecular Simulation: From Algorithms to Applications, Academic Press, San Diego, 1996.
[35] J. Hutter, M. Bernasconi, P. Focher, E. Fois, S. Godecker, M. Parrinello, M. Tuckermann, CPMD, MPI für Festkörperforschung and IBM Zurich Research Laboratory, 1995.
[36] R.N. Barnett, U. Landman, Phys. Rev. B 48 (1993) 2081.
[37] N. Troullier, J.L. Martins, Phys. Rev. B 43 (1991) 1993.
[38] S. Nosé, J. Chem. Phys. 81 (1984) 511.
[39] W.G. Hoover, Phys. Rev. A 31 (1985) 1695.
[40] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Oxford University Press, Oxford, 1989.
[41] P.L. Silvestrelli, Phys. Rev. B 59 (1999) 9703.
[42] F. Alber, G. Folkers, P. Carloni, J. Phys. Chem. B 103 (1999) 6121.
[43] F. Rastinejad, T. Perlmann, R.M. Evans, P.B. Sigler, Nature 375 (1995) 203.
[44] F. Guo, D.N. Gopaul, G.D. van Duyne, Nature 389 (1997) 40.
[45] Z. Otwinowski, R.W. Schevitz, R.G. Zhang, C.L. Lawson, A. Joachimiak, R.Q. Marmorstein, B.F. Luisi, P.B. Sigler, Nature 335 (1988) 321.
[46] C.L. Brooks III, M. Karplus, B. Montgomery-Pettit, Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics, Wiley, New York, NY, 1988.
[47] P. Carloni, M. Sprik, W. Andreoni, J. Phys. Chem. B 104 (2000) 823.
[48] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graphics 14 (1996) 33.