Journal of Computer-Aided Molecular Design **18:** 41–53, 2004.

Force-field parametrization and molecular dynamics simulations of Congo red

Marcin Król[a,*], Tomasz Borowski[b], Irena Roterman[a], Barbara Piekarska[c], Barbara Stopa[c], Joanna Rybarska[c] & Leszek Konieczny[c]

[a]*Department of Bioinformatics and Telemedicine, Collegium Medicum, Jagiellonian University, Kopernika 17, 31-501 Kraków, Poland;* [b]*Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Kraków, Poland;* [c]*Institute of Medical Biochemistry, Collegium Medicum, Jagiellonian University, Kopernika 7, 31-034 Kraków, Poland*

Received 21 July 2003; accepted in revised form 19 December 2003

Key words: diazo dyes, force-field parametrization, molecular dynamics simulations, supramolecular protein ligands

Abstract

Congo red, a diazo dye widely used in medical diagnosis, is known to form supramolecular systems in solution. Such a supramolecular system may interact with various proteins. In order to examine the nature of such complexes empirical force field parameters for the Congo red molecule were developed. The parametrization of bonding terms closely followed the methodology used in the development of the charmm22 force field, except for the calculation of charges. Point charges were calculated from a fit to a quantum mechanically derived electrostatic potential using the CHELP-BOW method. Obtained parameters were tested in a series of molecular dynamics simulations of both a single molecule and a micelle composed of Congo red molecules. It is shown that newly developed parameters define a stable minimum on the hypersurface of the potential energy and crystal and *ab initio* geometries and rotational barriers are well reproduced. Furthermore, rotations around C-N bonds are similar to torsional vibrations observed in crystals of diphenyl-diazene, which confirms that the flexibility of the molecule is correct. Comparison of results obtained from micelles molecular dynamics simulations with experimental data shows that the thermal dependence of micelle creation is well reproduced.

Introduction

For many years, Congo red (CR) (Figure 1) and other related diazo dyes have been commonly used to stain amyloid proteins [1–3], including Alzheimer plaque proteins and prions [4] and cellulose. Moreover, thermodynamical destabilization, incurred by unfolding conditions or physiological activity, may make non-amyloid proteins (especially those with high β-structure content) likely targets for diazo dyes binding. Clear examples of such binding are immunoglobulins induced to dye attachment by interactions with antigens [5, 6] and serpines [7].

*To whom correspondence should be addressed. Phone/Fax: +4812 4214057; E-mail: mykrol@cyf-kr.edu.pl

Congo red and related dyes have long been known to be able to self-assemble in water solution to form rod-like or ribbon-like micellar structures. It was shown recently [8–12] that the ability to self-assemble may, in fact, influence ligation capabilities. Actually, Skowronek et al. [11] have shown that only those dyes which are able to form stable micellar structures are efficient protein ligands. Therefore, it is assumed that Congo red may bind to proteins not only as a single molecule [13], but as a micelle as well. In fact, we believe that micelles composed of diazo dyes may constitute a new class of protein ligands characterized by a nonspecific mode of binding to proteins. The detailed molecular mechanism of such binding is still unclear; however, growing evidence indicates that a ribbon-like

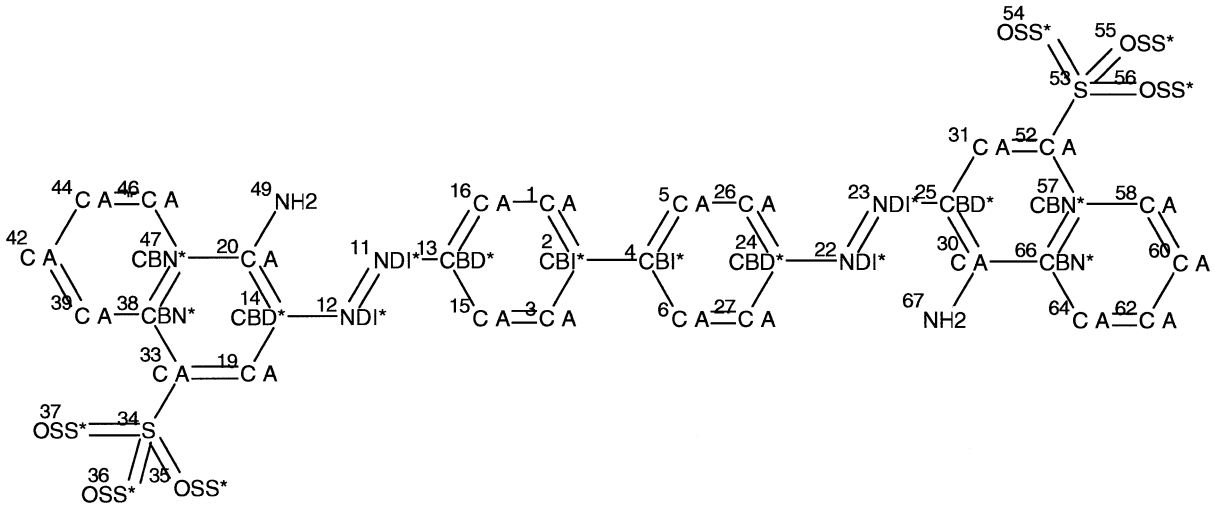

Figure 1. Congo red molecule. charmm [22] atom types are shown, new atom types (developed in the paper) are indicated by an asterisk.

micellar fragment interacts by adhesion with peptide chains in β-conformation, resulting in inclusion of the former into a β-sheet.

Difficulties in crystallization of Congo red–protein complexes make the computational approach (molecular mechanics) the method of choice to test and compare certain hypotheses on micelle formation and interaction with a β-sheet. This, however, requires that a proper parameter set for diazo dyes, which would be consistent with a protein force field, be developed.

Since we are interested in studying interactions of Congo red with proteins, we have decided to choose a well-tested protein force field and develop missing parameters within this force field, rather than choose a force field for small molecules, where parameters for Congo red would be easily accessible, but the energy function for proteins would not be as well described. Therefore, we have chosen to use the CHARMM [14] molecular modeling package, and, as a consequence, we had to develop missing parameters for Congo red in the charmm22 [15] force field. A thorough search through literature has revealed that molecular mechanics parameters for Congo red have already been obtained for the AMBER force field [16, 17]; however, no CHARMM parameters were available. The methodology used in charmm22 force-field development [18] was applied to define the missing parameters, except for the point charge generation procedure.

In the present article we describe the Congo red parametrization process and validation of the obtained parameters. Furthermore, results of the molecular dynamics simulations of various micelles in explicit solvent are presented.

Methods

Parametrization protocol

The CHARMM energy function [15] is a sum of various elements for internal and nonbonded terms and is given as follows:

$$V = \sum_{bounds} K_b(b-b_0)^2 + \sum_{angles} K_\theta(\theta-\theta_0)^2 + \\ + \sum_{Dihedrals} K_\phi(1+\cos(n\phi-\delta)) + \sum_{1,3pairs} K_{UB}(S-S_0)^2 + \\ + \sum_{Improper} K_w(w-w_0)^2 + \sum_{Nonbonded} \varepsilon_{ij} \left(\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{6} + \frac{q_i q_j}{4\pi D r_{ij}} \right)$$

where K denotes force constants for bonds, angles, torsions, Urey-Bradley terms and improper torsions, b_0, θ_0, S_0 and w_0 are equilibrium values, q are point charges and $R_{min,ij}$ is the equilibrium distance between two atom types i,j (for details see Ref. 14). In the current work we have not optimized van der Waals parameters (ε_{ij}, $R_{min,ij}$), as they are fairly insensitive to the surrounding environment and direct transfer of these parameters based on analogy with atoms already within the force field yields satisfactory results, especially in case of ligands that interact with biological molecules [18].

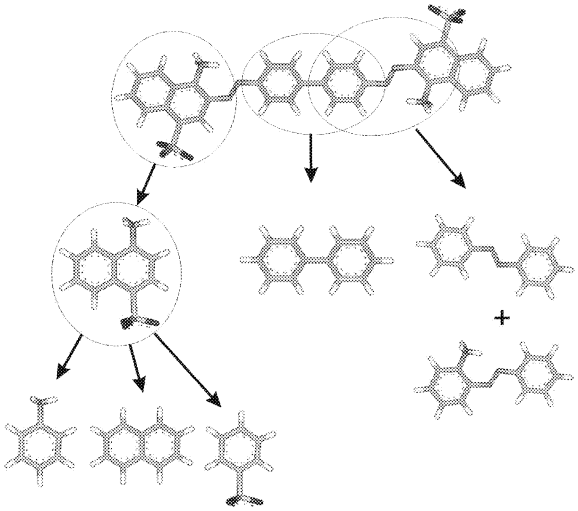

Figure 2. Choice of model compounds.

In the first step of parametrization we have divided the molecule into a set of model compounds, as it would be impossible to optimize force constants for the whole molecule. The model compounds are shown in Figure 2.

It should be noted that it is difficult to partition a Congo red molecule into model compounds which could be easily merged. In order to parametrize the N = N double bond we have decided to use diphenyl-diazene, although naphthalene-phenyl-diazene would be more suitable. It should be stressed that we attempted to keep the number of new atom types as small as possible; however, this attempt should have not impaired the quality of the parameters derived. All new atom types defined are shown in Figure 1. All model compounds were parametrized separately according to the same protocol listed below:

1. Equilibrium geometries of the model compounds were chosen based on quantum mechanical calculations (QM) and, where available, crystal data for bond, angle and dihedral values. QM geometry optimizations of neutral model compounds were performed using the hybrid density functional method B3LYP [19] with the 6-31G(d) basis set. In case of a negatively charged sulfonaphthalene molecule the 6-31+G(d) basis set was employed to accommodate an additional charge.
2. As we were unable to define model compounds that could be easily merged into a complete Congo red molecule, we decided to use another method of deriving point charges than the one proposed in the original charmm22 force-field paper [15]. The original method implied that point charges on atoms are adjusted to reproduce interaction energies and minimum energy geometries calculated at the HF/6-31G(d) level for supermolecules consisting of a model compound and a TIP3P water molecule. In the next step model compounds were linked by deleting two hydrogen atoms, moving their charges on the atoms to which they were attached and creating a bond between these atoms. However, we were not able to define model compounds that could be linked in such a way as to create a complete Congo red molecule. Therefore, we wanted to derive point charges with a method that would allow for recalculation of charges after the model compounds had been linked. For that reason, we have obtained partial charges for the model molecules from a fit to a quantum mechanical electrostatic potential calculated for the optimized model compounds using the CHELP-BOW method [20, 21]. In this method potential points used to derive charges are sampled with Boltzmann weights after their occurrence in actual simulations. The points are selected randomly within 8 Å from any atom in a molecule with a limit of 2500 points per atom. The original van der Waals parameters of the method were changed to fit those taken from the CHARMM force field. The method is shown to be superior to the CHELP [22], CHELPG [23], and Merz-Kollman [24] methods mainly because it uses a larger number of points chosen over a broader space. If the parametrized molecule was symmetrical, symmetry related atoms were constrained to the same charge.
3. Force constants associated with the bond lengths, bond angles and improper torsions were adjusted by fitting scaled *ab initio* vibrational data for the model systems. A scaling factor was set to 0.9216 – a proper value for B3LYP calculations [25]. Analysis of the vibrational spectra and potential energy distributions was performed with the Molvib [26] module of CHARMM. Availability of contributions from internal coordinate motions and frequencies made it possible to take both into account during force constants optimization.
4. Force constants associated with the dihedral angle for rotation around the bond linking phenyls in biphenyl, the C-N bond in diphenyl-diazene, the N = N double bond and rotation of amino and sulfonic groups were adjusted by fitting torsional barriers of rotation derived by quantum mechanical calculations. In the case of the N = N double bond

a potential energy landscape was sampled every 30 degrees using the CASSCF/6-31G(d) method. The active space in the CASSCF method was the Π-system of the diazodibenzene molecule (12 Π orbitals of phenyl rings and Π and Π* orbitals of an N = N bond). In all remaining torsions data were sampled every 15 degrees and B3LYP/6-31G(d) (6-31+G(d) in case of ionized sulfonic groups) was used. Torsional force constants were adjusted to reproduce both the height of the barrier and the location of the minima and maxima on the energy landscape plot.

5. The parametrization process was accomplished in an iterative way – after torsional force constants had been optimized, a check was done to see how equilibrium geometries and a vibrational data fit were affected and necessary modifications were made. This was repeated until convergence of all parameters was achieved. Point charges were calculated only once and were not included in the iterative process.

After parameters for all model compounds had been developed, they were used to generate topology and parameter files for the whole Congo red molecule. Force constants and equilibrium geometries were taken directly from the model compounds and point charges were recalculated for the whole Congo red molecule using the method described above. Recalculation of charges was necessary, as there were ambiguities during the 'merging' of the model compounds. Point charges on symmetry related atoms were constrained to have the same charge.

Subsequently molecular mechanics calculations were performed to verify the quality of the parametrization process.

Computational details

All molecular mechanics (MM) calculations were performed using the CHARMM (c27b2) program [14]. Energy minimizations of the model compounds were carried out to a final gradient of 10^{-6} kcal/mol*Å. MM frequencies were determined using the VIBRAN facility of CHARMM, and the analysis of contributions from various internal coordinates to frequencies was performed with the MOLVIB module. The choice of internal coordinates for ring systems was based upon recommendations of Pulay [27].

X-ray crystallographic data for model compounds were obtained from the Cambridge Structural Database [28, 29]. Programs for the CHELP-BOW charge fitting procedure were kindly delivered by Prof. Ulf Ryde.

All quantum mechanical calculations were performed with the Gaussian98 [30] program. Geometry optimizations of model compounds were carried out with the Opt=tight keyword, which resulted in a predicted energy change of approximately 10^{-10} Hartree (approx. 10^{-7} kcal/mol).

Molecular dynamics simulation of a single Congo red molecule in vacuum was performed for 640.5 ps. First, the molecule was minimized and heated for 50 ps to a temperature of 1000 K. This was followed by 150 ps of constant temperature and 440.5 ps slow cooling MD simulations with the Nose thermostat [31]. At the end of the cooling phase the temperature was 247 K and the molecule was minimized.

Starting structures of Congo red micelles were created using the authors' own software. All micelles were composed of 4 trans molecules and the distance between individual molecules was set to 4.5 Å. The micelle generation protocol has been presented elsewhere [5, 11] and is here described only briefly. In short, the source Congo red molecule was placed with the central benzidine bond in the center of the coordinate system, the bond defined the x-axis and a benzene ring plane defined the xy-plane. Positions of other molecules were determined using three transformations: T_y – translation along the y-axis, R_z – rotation around the z-axis, and T_z – translation along the z-axis (this parameter was fixed at 4.5 Å – the optimal van der Waals distance between consecutive Congo red molecules). Every second molecule was rotated by 180° around the long internal molecule axis to produce antiparallel mutual orientation of dipole moments.

Subsequently, micelles were placed in rectangular boxes containing TIP3P water molecules. Na^+ and Cl^- ions were added to the simulation cell in the amount necessary to mimic a 2.8% salt concentration.

After setting up a simulation cell the system was minimized and heated from 200 K to the desired temperature by 5 K every 100 steps. After heating, 5 ns NPT [32] production runs were performed at 300, 340 and 450 K.

In all solvent simulations PME electrostatics [33] with grid density ~ 1 point/Å, $\kappa = 0.32$, sixth interpolation order was used.

In all MD simulations all bonds between hydrogen atoms and the H-H distance in water were fixed using the SHAKE [34] algorithm, which allowed a time step of 2 fs.

Table 1. Difference in bond lengths and angles between target, QM and CHARMM optimized geometries. CR stands for Congo red.

	BOND		ANGLE	
	Average difference (Å)	Standard deviation (Å)	Average difference (deg)	Standard deviation (deg)
Biphenyl	0.03	0.0038	0.32	0.13
Diphenyl-diazene	0.025	0.0036	0.86	0.21
2-Phenylazo-phenylamine	0.023	0.0038	1.27	0.24
Sulfobenzene	0.012	0.0035	0.33	0.07
CR (compared to target)	0.017	0.0022	2.32	0.41
CR (compared to QM)	0.022	0.0030	1.82	0.3
CR after SA (compared to QM)	0.021	0.0029	2.85	0.46

Results and discussion

Model compounds parametrization

Biphenyl

Equilibrium geometries for a biphenyl molecule were taken from Ref. 35 and a QM optimized structure. Average differences in target and CHARMM optimized bond lengths and angles for the biphenyl molecule are given in Table 1. In general, the bond lengths and angles fall within the range of experimental errors. It should be noted that introduction of a new atom type, CBI, was necessary to obtain a reasonable value of the phenyl–phenyl bond length. Van der Waals parameters for the CBI atom type were taken directly from the CA atom (aromatic carbon in the charmm22 force field). However, in the case of a biphenyl molecule a key problem is a proper parametrization of the dihedral between two phenyl planes. This dihedral angle is positioned in the center of the Congo red molecule and its value is crucial for the conformation of the whole molecule. Moreover, it was suggested [11] that the twist of two symmetrical parts of Congo red influences its ability to form a micelle. Biphenyl QM calculations determined the value of the dihedral to be 38.4°. In the crystal structure phenyl rings were planar. It is clearly seen that rotation around this bond does not require much energy. Indeed, the QM barrier of rotation (shown in Figure 3A, solid line) indicates that the difference between planar and twisted conformations is about 2 kcal/mol. Moreover, Baudour [36] has experimentally estimated the height of the barrier to be 2.2 kcal/mol and 2.6 kcal/mol for gaseous biphenyl planar and perpendicular configurations, respectively. We have set the equilibrium value of the dihedral to its QM value and the rotational energy profile obtained in CHARMM is shown in Figure 3A, dashed line. Overall perfect agreement of the QM and MM rotational energy profiles is observed for the model compound. What is more, the calculated energy barrier is in very good agreement with the experimental data. In the case of the Congo red molecule the agreement is worse (dot-dashed line). However, attempts to correct the value of the barrier height proved to be unsuccessful, as lowering the energy of the barrier worsens the overall agreement of the potential energy scan (PES) and negatively influences other parameters. Furthermore, it should be noted that attachment of naphthalene rings to a biphenyl moiety substantially alters the rotational energy profile of the molecule and no experimental or theoretical data on this profile for the whole Congo red molecule are readily available.

Frequencies and internal coordinate contributions to the normal modes of vibrations are in good agreement with the target B3LYP/6-31G(d) frequency calculations. The RMS difference between scaled *ab initio* and empirical frequencies is given in Table 2. It should be noted that the lowest energy mode corresponding to rotation around the bond linking two phenyl rings has been parametrized to fit the rotational barrier and not the IR spectrum; therefore, it was excluded from RMS calculation. A higher RMS for modes associated with bond and angle parameters was observed in all model compounds and may be caused by two effects. First, modes with a frequency larger than ~ 800 cm^{-1} contain contributions from a large number of internal coordinates and a proper parametrization of these modes is, therefore, very difficult. Second, we have not modified force

Table 2. RMS difference between CHARMM and QM derived frequencies for all, torsional and bonding modes.

	Total RMS difference (cm^{-1})	Torsional modes RMS difference (cm^{-1})	Bond and angles RMS difference (cm^{-1})
Biphenyl	36.1	26.6	40.5
Diphenyl-diazene	36.7	23.6	46.8
2-Phenylazo-phenylamine	29.3	22.2	34.3
Sulfobenzene	37.6	39.7	35.3
1-Amino-4-sulfonaphthalene	28.2	22.6	31.7

constants present in the original charmm22 force field. Consequently, our ability to change force constants to reproduce frequencies was limited only to bonds, angles and dihedrals involving new atom types. Overall, the agreement between QM and MM frequencies is very good and indicates that bonding parameters are correct.

Diphenyl-diazene

Equilibrium geometries for the diphenyl-diazene molecule were taken from Refs. 37–39 and the QM optimization. Average differences in target and CHARMM optimized bond lengths and angles are given in Table 1. In general, the bond lengths and angles fall within the range of experimental errors. The exceptions are angles between azo bonds and phenyl rings. The differences are approximately 2.5°. We have introduced two new atom types, i.e., CBD – carbon atom that binds azo nitrogen, and NDI – azo nitrogen atom. The original CHARMM force field does not have a double bonded nitrogen atom bound to another nitrogen atom and, as a consequence, the NDI atom type must have been introduced. Van der Waals parameters for that atom type were taken from an N atom. The CBD atom type was introduced to sustain the necessary C-N bond length and C-C-N bond angle, although the latter was achieved with minor success. Its van der Waals parameters were taken from a CA atom.

During the optimization of N = N dihedral parameters we could not use single determinant QM methods (such as B3LYP) to generate a potential energy scan (PES). Consequently, we have chosen the CAS-SCF model to probe the energy profile of the torsion. The QM barrier of rotation is shown in Figure 3B, solid line. A MM profile (dashed line) was calculated for a diphenyl-diazene molecule and this is why both profiles are different to a certain extent. However, the main objective was to reproduce the height of the barrier and this goal was accomplished. The main difference – disparity in the minimum depth for diphenyl-diazene at 0° and 180° by approximately 5 kcal/mol – is caused by the fact that the *trans*-diphenyl-diazene has lower energy compared to the *cis* form because of possible steric clashes in the latter form, whereas in case of CH_3-N=N-CH_3 such an effect does not exist and, consequently, both minima have the same energy.

Dihedral parameters for rotation around the CBD-NDI bond were also parametrized to reproduce the QM PES calculated for the diphenyl-diazene molecule. The energy of the QM derived barrier was approximately 7 kcal/mol and is well reproduced by the MM energy profile (Figure 3C). Frequencies and internal coordinate contributions to the normal modes of vibrations are also in good agreement with the target B3LYP/6-31G(d) frequency calculations. The RMS difference between scaled *ab initio* and empirical frequencies is given in Table 2. It should be noted that in case of a diphenyl-diazene molecule two low energy modes have been parametrized to fit a rotational barrier and not an IR spectrum; therefore, they were excluded from RMS calculation.

Naphthalene

The naphthalene moiety was parametrized in several consecutive steps. At first we tried to parametrize the whole 1-amino-4-sulfonaphthalene; however, this proved to be very difficult, as the number of normal modes and numbers of contributions in each mode were too large to efficiently set force constants. As a result, we decided to split 1-amino-4-sulfonaphthalene into a set of smaller model compounds, parametrize them, and use these parameters as starting values in parametrization of the whole 1-amino-4-sulfonaphthalene. The smaller compounds were: unsubstituted naphthalene, phenylamine and sulfobenzene. Moreover, during the parametrization

Table 3A. Equilibrium bond distances (b_0, in Å) and angles (θ_0, in degrees) and Urey-Bradley equilibrium distances (S_0, in Å), and stretching (K_b, in kcal/mol/Å^2) and bending (K_θ, in kcal/mol/rad^2) and Urey-Bradley (K_{UB}, in kcal/mol/Å^2) force constants for bonds and angles involving new atom types.

Bond	b_0	K_b	Angle	θ_0	K_θ	S_0	K_{UB}
CA-CBI	1.4000	295.0	CBI-CA-CA	121.31	32.0	2.4362	27.0
CA-CBN	1.4260	280.0	CA-CBI-CA	117.71	32.0	2.4102	27.0
CBI-CBI	1.4895	267.0	CA-CBI-CBI	121.20	27.0	2.5160	22.0
CBN-CBN	1.4355	198.0	CA-CBN-CA	122.58	40.0	2.4102	35.0
CA-CBD	1.4075	310.0	CA-CBN-CBN	118.71	40.0	2.5160	35.0
CBD-NDI	1.4078	300.0	CBD-CA-CBN(CA)	120.55	35.0	2.4063	20.0
NH2-CA	1.4836	330.0	CBD-NDI-NDI	113.00	85.0		
NDI-NDI	1.2680	560.0	CA-CBD-NDI	120.53	85.0		
S-CA	1.7996	350.0	CA-CBD-CA	118.94	36.0	2.4173	31.0
S-OSS	1.4725	500.0	CBN-CA-CA	120.74	40.0	2.4362	35.0
			NH2-CA-CBD	110.00	47.0		
			S-CA-CA	119.56	20.0		
			S-CA-CBN	122.44	21.0		
			OSS-S-CA	105.38	75.0		
			OSS-CA-OSS	113.18	85.0		
			HP-CA-CBI	119.59	23.0	2.1572	15.0
			HP-CA-CBD	118.82	22.0	2.1572	14.0
			HP-CA-CBN	119.57	30.0	2.1572	22.0
			H-NH2-CA	109.63	45.0		

process it turned out that an amino group should also be attached to the diazo compound, to obtain the CBD-CA-N angle force constant. Therefore calculations for 2-phenylazo-phenylamine were also performed.

We introduced two new atom types in the molecule: CBN – a central carbon atom of the naphthalene moiety, and OSS — a sulfonic oxygen atom. Van der Waals parameters of the former were taken from CA and the latter from O atoms. These atom types were necessary to reasonably represent desired equilibrium geometries and QM frequencies. As was mentioned above, it was impossible to adjust all force constants for the whole 1-amino-4-sulfonaphthalene molecule. However, force constants taken from smaller model compounds were a reasonable starting point for further optimization. RMS differences between MM and QM frequencies are given in Table 2. It is interesting to note that even though 1-amino-4-sulfonaphthalene was a much more difficult molecule to adjust force constants compared to biphenyl and diphenyl-diazene, RMS differences are not larger. This indicates that bonding parameters are of good quality.

Table 3B. Torsional parameters K_ϕ (kcal/mol), n, δ (degrees) for selected dihedrals involving new atom types.

Dihedral	K_ϕ	n	δ
CA-CBI-CBI-CA	0.13	4	0.0
CA-CBI-CBI-CA	0.78	2	180.0
CA-CBD-NDI-NDI	3.00	2	180.0
CBD-NDI-NDI-CBD	27.40	2	180.0
CBD-NDI-NDI-CBD	10.00	1	180.0
-S-CA-	0.1	3	180.0
-NH2-CA-	0.6	3	180.0

Congo red

After all the model compounds had been parametrized, we merged them to create a complete Congo red molecule. All bonding parameters were sustained; they are given in Table 3A (force constants and equilibrium geometries) and Table 3B (selected dihedral parameters). We also decided to recalculate partial atomic charges. This was necessary as linking model compounds raised ambiguities on how to sum up and divide partial charges on particular atoms. As the

charges were obtained from an electrostatic potential fit, we decided that it would be better to calculate them once more for the whole molecule, rather than get unreasonable charges by wrong linking. The comparison of charges obtained for Congo red and model compounds is given in Table 4. The average difference in charges is 0.072 ± 0.057e. Although a change in charges influences force constants and equilibrium geometry parameters, this influence should not be great, as differences between charges are not large. As differences in CHARMM optimized and target bond lengths and angles for the Congo red molecule indicate (given in Table 1), modifying charges after the other parameters were developed has not impaired the results obtained.

Dashed–dotted lines in Figure 3A–C show CHARMM potential energy scans for the most important dihedrals of the Congo red molecule parametrized against QM rotational barriers. Rotation around the central bond was already discussed in the biphenyl section and is not repeated here.

The rotational energy barrier for the NDI = NDI bond is well preserved in both forms of the Congo red molecule compared to QM data for diphenyl-diazene. Although absolute values differ by about 10 kcal/mol, the difference between the energy of the minimum and the maximum is sustained. Rotational parameters for the CBD-NDI dihedral angle constitute a more complex problem, as they must describe a rotational energy barrier in two distinct environments – rotation of the 1-amino-4-sulfonaphthalene moiety (e.g. rotation around C19-C14-N12-N11) and rotation of the 1-amino-2-azo-4-sulfonaphthalene moiety (e.g. rotation around N12-N11-C13-C16). Analysis of the *ab initio* derived CBD-NDI barriers of rotation in both environments for a 2-phenylazo-phenylamine molecule revealed that we are not able to simultaneously reproduce rotational energies in both cases. To solve this problem we would have to introduce different atom types for C14 and C13, which would further complicate the parametrization problem. The dashed–dotted line in Figure 3C shows the rotational energy barrier for the rotation around C19-C14-N12-N11. Inspection of the picture reveals that the difference between the minimum energies at 180° and 0° is approximately 6 kcal/mol for the Congo red molecule. However, there is no such difference in case of an unsubstituted diphenyl-diazene molecule (solid and dashed lines in Figure 3C). In order to verify if this destabilization is due to steric clashes between amino and azo nitrogen atoms in the Congo red molecule, we have calculated QM energies of two conformers of 2-phenylazo-phenylamine with the dihedral angle defined by C19-C14-N12-N11 equal to 0° and 180°, respectively. The former conformer has a lower energy by 2.6 kcal/mol. This may partially account for the destabilization of the 180° conformation in Figure 3C for both forms of Congo red.

Test simulations

Single molecule

The simulated annealing (SA) MD protocol is a useful tool in the search for minima on the potential energy surface of a molecule [41]. Contrary to a minimization or a low-temperature MD simulation, the high temperature of a simulated annealing allows for an effective scanning of the potential energy surface, as even high-energy barriers between local minima can be easily crossed. We have used such a protocol to test if the developed parameters define a stable, global minimum on the potential energy surface of the molecule. After 640.5 ps of the annealing simulation we have minimized the molecule and measured the most important geometrical parameters. Average differences in QM and post-annealed bond lengths and angles are given in Table 1. Figure 4 shows QM optimized and post-annealed structures of the molecule superimposed on each other.

It is clearly seen that the mutual orientation of different fragments is very similar, with the exception of diazo bonds. Closer trajectory analysis reveals that interconversion between two conformations occurs through simultaneous torsional rotation around NDI-CBD bonds. This effect is well known in diphenyl-diazene and related compounds and is the origin of the orientational disorder and shrinkage of the N-N bond in crystals [39]. Therefore, the fact that such interconversions occur during MD simulation indicates not only that the developed parameters define a stable minimum on the potential energy surface of the molecule, but also that torsional vibrations are well reproduced. The inability of a MD simulation at 1000 K to produce an alternative conformation denotes that the equilibrium geometry of the molecule will be sustained over the range of temperatures used in standard MD simulations.

In order to further check if overall conformational stability is not the result of damping of important intramolecular motions or freezing out molecular degrees of freedom, we have calculated root mean square fluctuations (RMSFs) about the average struc-

Table 4. Comparison of charges calculated for the Congo red (CR) molecule and model (MODEL) compounds. For atom numbering see Figure 1.

Atom	CR	MODEL	Diff.	Atom	CR	MODEL	Diff.
C1	−0.036	−0.104	0.068	S34	0.924	1.152	−0.228
C2	0.002	0.035	−0.033	O35	−0.541	−0.672	0.131
C3	−0.139	−0.104	−0.035	O36	−0.541	−0.672	0.131
C4	0.002	0.035	−0.033	O37	−0.541	−0.672	0.131
C5	−0.139	−0.104	−0.035	C38	0.198	0.038	0.160
C6	−0.036	−0.104	0.068	C39	−0.120	−0.175	0.055
N11	−0.183	−0.165	−0.018	C41	−0.050	−0.109	0.059
N12	−0.099	−0.142	0.043	C43	−0.158	−0.086	−0.072
C13	0.209	0.208	0.001	C45	−0.109	−0.227	0.118
C14	0.125	0.139	−0.014	C47	−0.151	−0.273	0.122
C15	−0.064	−0.035	−0.029	N49	−0.768	−0.841	0.073
C16	−0.240	−0.214	−0.026	C52	−0.106	−0.089	−0.017
C19	−0.191	−0.167	−0.023	S53	0.924	1.152	−0.228
C20	0.269	0.345	−0.076	O54	−0.541	−0.672	0.131
N22	−0.183	−0.165	−0.018	O55	−0.541	−0.672	0.131
N23	−0.099	−0.142	0.043	O56	−0.541	−0.672	0.131
C24	0.209	0.208	0.001	C57	0.198	0.038	0.160
C25	0.125	0.139	−0.014	C58	−0.120	−0.175	0.055
C26	−0.064	−0.035	−0.029	C60	−0.050	−0.109	0.059
C27	−0.240	−0.214	−0.026	C62	−0.158	−0.086	−0.072
C30	0.269	0.345	−0.076	C64	−0.109	−0.227	0.118
C31	−0.191	−0.167	−0.023	C66	−0.151	−0.273	0.122
C33	−0.106	−0.089	−0.017	N67	−0.768	−0.841	0.073

Table 5. Root mean square fluctuations (RMSF) (Å) for all Congo red atoms determined from a 640.5 ps long simulated annealing of the neutral form.

Atom	RMSF	Atom	RMSF	Atom	RMSF	Atom	RMSF	Atom	RMSF
C1	1.34	C15	1.40	H29	2.00	H43	2.15	H57	3.50
C2	0.96	C16	1.32	C30	1.73	C44	2.71	C58	0.87
C3	1.41	H17	2.03	C31	0.75	H45	3.45	C59	1.17
C4	0.95	H18	1.95	H32	1.42	C46	2.66	H60	1.26
C5	1.48	C19	0.68	C33	0.76	H47	3.39	C61	1.96
C6	1.36	C20	1.68	S34	1.93	C48	1.70	H62	2.15
H7	1.98	H21	1.28	O35	2.69	N49	2.59	C63	2.80
H8	2.06	N22	0.79	O36	2.63	H50	3.17	H64	3.59
H9	2.18	N23	0.96	O37	2.58	H51	2.72	C65	2.77
H10	2.06	C24	0.89	H38	3.16	C52	0.85	H66	3.56
N11	0.78	C25	0.83	C39	0.90	S53	2.17	C67	1.74
N12	0.96	C26	1.48	C40	1.19	O54	2.93	N68	2.71
C13	0.89	C27	1.33	H41	1.23	O55	2.91	H69	2.84
C14	0.85	H28	2.18	C42	1.95	O56	2.79	H70	3.36

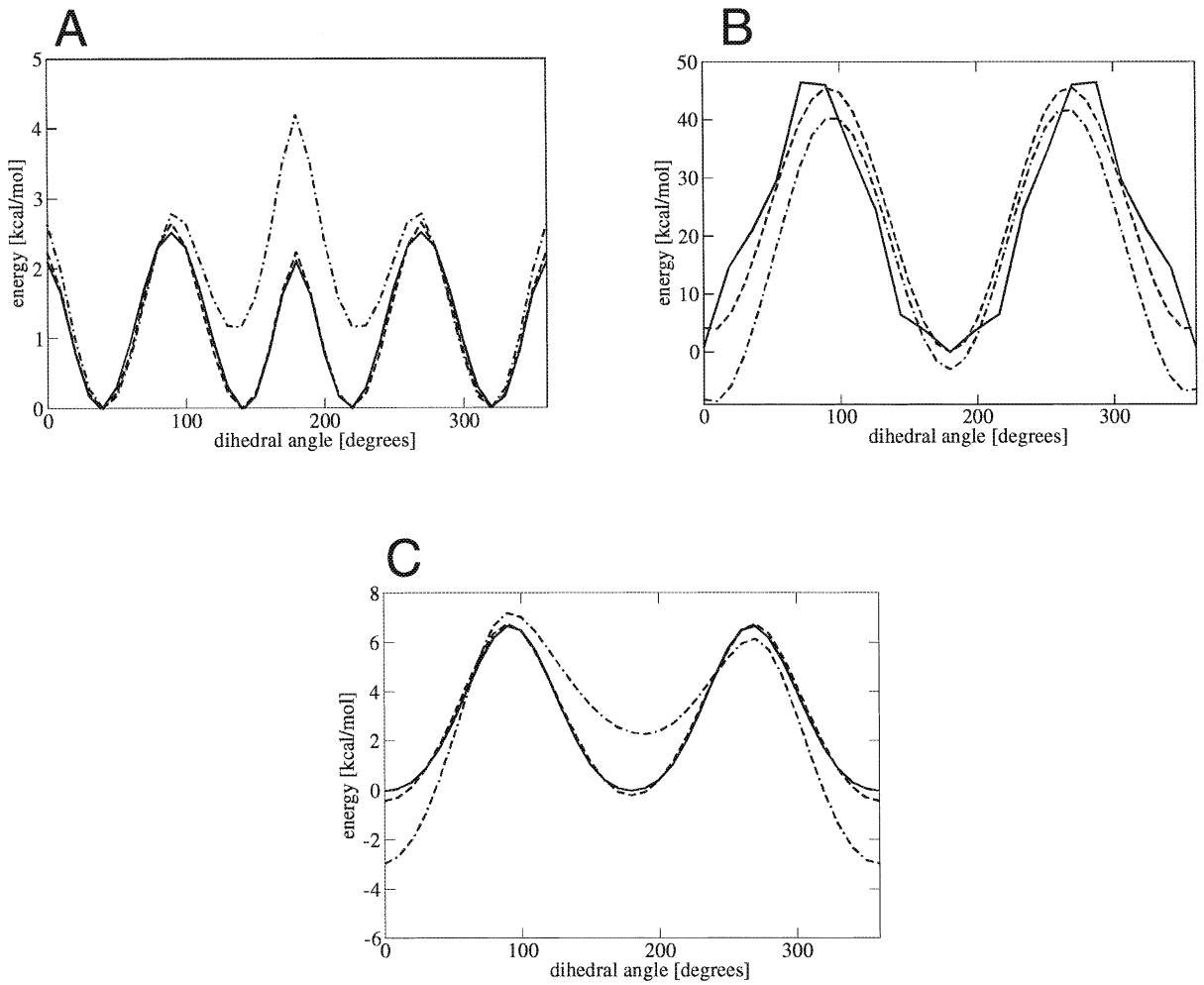

Figure 3. Barriers of rotation obtained from *ab initio* calculations for model compounds (solid line), CHARMM calculations for model compounds (dashed line), and CHARMM calculations for Congo red (dash–dotted line). (A) Rotation around the CBI-CBI bond, (B) rotation around the NDI-NDI bond, (C) rotation around the NDI-CBD bond.

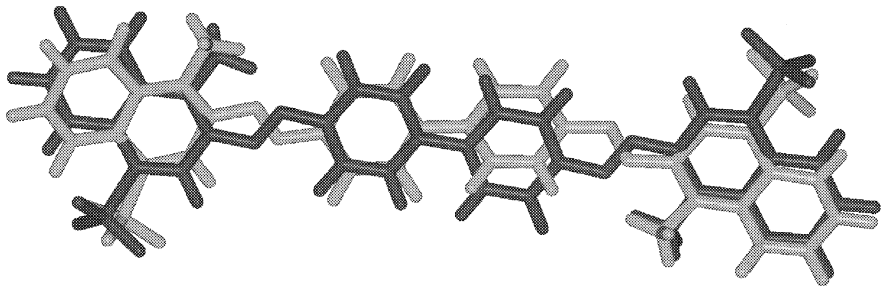

Figure 4. QM derived (dark grey) and post-annealed (light grey) structures of the Congo red molecule superimposed on each other.

Table 6. Results of micelles MD simulations. The first column gives a micelle structure; the first number represents T_y (Å) and the second number represents R_z (degrees). E means that the micellar structure was obtained during the first 300 ps of the simulation (equilibration period). Numbers show time (in ps) necessary to obtain a stable micellar structure. Brackets indicate that a high-energy structure was obtained (Figure 5D).

Micelle structure	300 K	340 K	450 K
0_0	E	E	E
0_5	E	E	E
0_10	3032	E	1305
0_15	440	E	400
0_20	1064	1094	350
0_25	(1612)	E	E
0_30	814	620	E
0_35	(615)	(662)	(3100)
5_5	1469	E	E
5_10	E	E	E
5_15	E	E	E
5_20	2784	E (trans → cis)	E
5_25	765	484	E (trans → cis)
5_35	1275	495	535
10_5	E	E	E
10_10	E	E	550
10_15	672	E	E
10_20	1774	362	E (trans → cis)
10_35	1212	E	NA

ture. They are given in Table 5. Average RMSF for all heavy atoms is 1.63 Å. Relatively high values of RMSF confirm that the molecule has a large amount of conformational flexibility and is not artificially stabilized by too strong force constants. It is interesting to point out that visual trajectory analysis has revealed that there is a frequent rotation around the central bond in the molecule, and that that *cis* conformation is also quite often sampled. In fact, the molecule spent 19% of the simulation time in the cis and 81% in the trans conformation. This is probably caused by a larger barrier height for the Congo red molecule compared to the model compound (see Figure 3A). However, our attempts to lower the value of the barrier height for the dye molecule proved to be unsuccessful.

Micelles simulations

We have performed a series of molecular dynamics simulation studies for Congo red micelles in a water solution. Results of all simulations are summarized in Table 6.

Firstly, data shown in Table 6 indicate that heating results in faster micelle structure generation. During the equilibration period (E), 13 starting structures simulated at T = 340 K converged to the low energy micelle structure (Figure 5A–C), while simulations at T = 300 K produced only 6 low energy micelle structures in the same period. It should be mentioned, however, that all starting structures simulated at T = 300 K yielded stable micelles during simulations of 5 ns. The systems that required longer simulations also converged faster to the low energy conformation if simulated at elevated temperatures (0_10, 0_25, 5_5, 5_20, 5_25, 5_35, 10_20, 10_35). The exceptions include 0_20, where heating to 450 K was necessary in order to accelerate micelle formation.

Further increase of the temperature generally does not expedite micelle formation (with the exception of 0_20). In the case of large R_z values consecutive dye molecules may interlace – the order of naphthalene rings belonging to the same molecule is different on both sides of a micelle (such a situation occurs for the 0_35 micelle). As a result, biphenyl fragments of two consecutive molecules intersect and the energy of such a system is higher (Figure 5D). Visual analysis of the micelle MD trajectories indicates that created micelles are stable and are not disrupted, regardless of the temperature of simulations. However, micelles simulated at 450 K show large fluctuations and translations of individual molecules with respect to each other.

These observations indicate that micelles are likely to be formed at elevated temperatures, most probably because some additional kinetic energy is needed to reorganize the position of Congo red molecules between water and ions. If the kinetic energy falls below a certain threshold, as in 300 K simulations, molecules need much more time to reorganize. However, if the temperature is increased, molecules gain additional energy, which enables them to reorient and form stable stacked micelles quicker. These outcomes are in excellent agreement with experimental data [12], which show that micelle formation is inhibited, or proceeds slowly, at low temperatures.

Moreover, Skowronek et al. [12] have shown that one of the indicators of a micelle creation process is a NMR measured downfield chemical shift alteration of a naphthalene proton, which comes close to a charged sulfonic group of another Congo red molecule in the supramolecular organization. The distance between the sulfonic group and the proton was estimated to be around 4 Å. In post-dynamical structures of micelles calculated with the newly developed parameters the

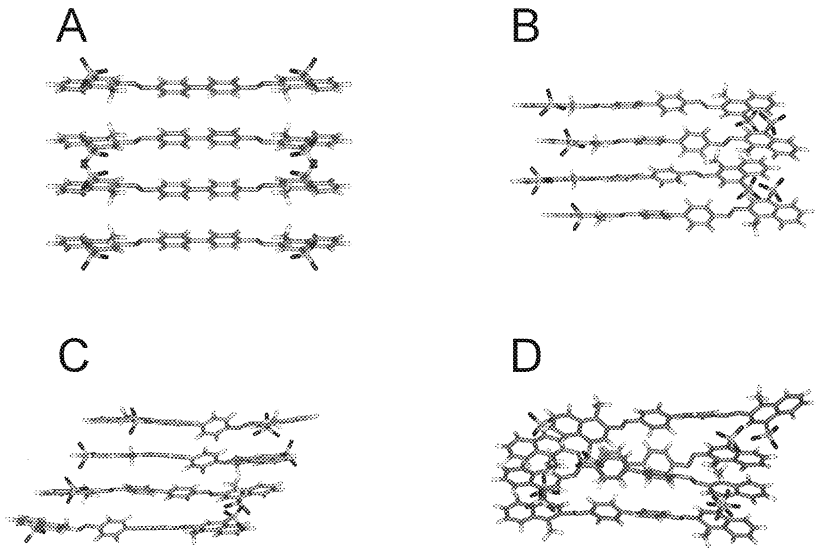

Figure 5. Different types of Congo red micelles post-dynamical structures. (A–C) Low-energy micelles, (D) high-energy micelle.

distance between the sulfonic group and the proton varied between 3.8–4.4 Å depending on the starting micelle structure. These distance values are, in fact, similar to the experimentally measured value. This further indicates that the newly developed parameters yield a micelle structure which is in accord with experimental results.

Conclusions

The charmm22 force field has been successfully extended to incorporate parameters for diazo dyes. The parametrization process of all bonding parameters followed the methodology used in the development of the original force field. This, together with unchanged van der Waals parameters, guarantees that the newly developed parameters are consistent with the protein part of the force field and can be used in simulations of protein-diazo dyes complexes. Various high level *ab inito* data, such as equilibrium geometries, IR frequencies and rotational barriers are adequately reproduced for model systems and for a single Congo red molecule. Rotational vibrations of azo bonds confirm that developed parameters not only define a stable structure, but also allow for conformational flexibility to be in accord with experimental data. Moreover, experimental data on micelle formation are well reproduced. It is especially important that thermal criteria for micelle formation agree with experiments, as this indicates that the new parameters can be used to study intermolecular interactions of Congo red molecules in mesophase structure constitution and docking calculations. Furthermore, other diazo dyes, such as trypan blue and Evans blue, can be easily parametrized with the use of the developed parameters. The newly developed parameters will also be used to design and test new self-associating molecules, which could become potent, non-specific liquid crystalline protein ligands.

Acknowledgements

This study was supported by the Polish State Committee for Scientific Research (Grant 4 T11F 015 22). QM calculations were performed on a SGI2800 high performance computer at Academic Computer Center of Stanislaw Staszic University of Mining and Metallurgy in Kraków, Poland (grant KBN/SGI_ORIGIN_2000/UJ/106/1998).

References

1. Glenner, G.G., Eanes, E.D. and Page, D.L., Histochem. Cytochem., 20 (1972) 821.
2. Sunde, M. and Blake, C.C.F., Quart. Rev. Biophys., 31 (1998) 1.
3. Westermark, G.T., Johnson, K.H. and Westermark, P., Methods Enzymol., 309 (1999) 3.
4. Sipe, J.D., Annu. Rev. Biochem., 61 (1992) 947.
5. Roterman, I., No, K.T., Piekarska, B., Kaszuba, J., Pawlicki, R., Rybarska, J. and Konieczny, L., J. Physiol. Pharmacol., 44 (1993) 213.

6. Piekarska, B., Skowronek, M., Rybarska, J., Stopa, B., Roterman, I. and Konieczny, L., Biochemie, 78 (1996) 183.
7. Roterman, I., Rybarska, J., Konieczny, L., Skowronek, M., Stopa, B., Piekarska, B. and Bakalarski, G., Comput. Chem., 22 (1998) 61.
8. Stopa, B., Konieczny, L., Piekarska, B., Roterman, I., Rybarska, J. and Skowronek, M., Biochemie, 79 (1997) 23.
9. Stopa, B., Górny, M., Konieczny, L., Piekarska, B., Rybarska, J., Skowronek, M. and Roterman, I., Biochemie, 80 (1998) 963.
10. Konieczny, L., Piekarska, B., Rybarska, J., Skowronek, M., Stopa, B., Tabor, B., Dąbrós, W., Pawlicki, R. and Roterman, I., Folia Histochem. Cytochem., 35 (1997) 203.
11. Skowronek, M., Roterman, I., Konieczny, L., Stopa, B., Rybarska, J. and Piekarska, B., J. Comput. Chem., 21 (2000) 656.
12. Skowronek, M., Stopa, B., Konieczny, L., Rybarska, J., Piekarska, B., Szneler, E., Bakalarski, G. and Roterman, I., Biopolymers, 46 (1998) 267.
13. Turnell, W.G. and Finch, J.T., J. Mol. Biol., 227 (1992) 1205.
14. Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S. and Karplus, M., J. Comput. Chem., 4 (1983) 187.
15. MacKerell, A.D. Jr., Bashford, D., Bellot, M., Dunbrack, R.L. Jr., Evenseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E. III, Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. and Karplus M., J. Phys. Chem. B, 102 (1998) 3586.
16. Skowronek, M., Roterman, I., Konieczny, L., Stopa, B., Rybarska, J., Piekarska, B., Górecki A. and Król, M., Comput. Chem., 24 (2000) 429.
17. Woodcock, S. and Henrissat, B., Biopolymers, 36 (1995) 201.
18. MacKerell, A.D., Jr. In Becker, O.M., MacKerell, A.D., Jr., Roux, B. and Watanabe, M. (Eds.) Computational Biochemistry and Biophysics. Marcel Dekker, Inc., New York, 2001, p. 7.
19. Barone, V., Chem. Phys. Lett., 226 (1994) 392.
20. Sigfridsson, E. and Ryde, U., J. Comput. Chem., 19 (1998) 377.
21. Sigfridsson, E., Ryde, U. and Bush, B.L., J. Comput. Chem., 23 (2002) 351.
22. Chirlain, L.E. and Francl, M.M., J. Comput. Chem., 6 (1987) 894.
23. Bessler, B.H., Merz, K.M. and Kollman, P.A., J. Comput. Chem., 11 (1990) 431.
24. Breneman, C.M. and Wiberg, K.B., J. Comput. Chem., 11 (1990) 361.
25. Foresman, J.B. and Frisch, Æ. (Eds.) Exploring Chemistry with Electronic Structure Methods. Gaussian, Inc., Pittsburgh, PA, 1996.
26. Wiorkiewicz-Kuczera, J., Kuczera K. and Karplus, M., MOLVIB program, Harvard University, MA, 1989.
27. Pulay, P., Fogarasi, G., Pang, F. and Boggs, J.E., J. Am. Chem. Soc., 101 (1979) 2550.
28. Allen, F.H., Acta Crystallogr. B, 58 (2002) 380.
29. Bruno, I.J., Cole, J.C., Edgington, P.R., Kessler, M., Macrae, C.F., McCabe, P., Pearson, J. and Taylor, R., Acta Crystallogr. B, 58 (2002) 389.
30. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Zakrzewski, V.G. Montgomery, J.A. Jr., Stratmann, R.E. Burant, J.C. Dapprich, S. Millam, J.M. Daniels, A.D. Kudin, K.N. Strain, M.C. Farkas, O. Tomasi, J. Barone, V. Cossi, M. Cammi, R. Mennucci, B. Pomelli, C. Adamo, C. Clifford, S. Ochterski, J. Petersson, G.A. Ayala, P.Y. Cui, Q. Morokuma, K. Malick, D.K. Rabuck, A.D. Raghavachari, K. Foresman, J.B. Cioslowski, J. Ortiz, J.V. Baboul, A.G. Stefanov, B.B. Liu, G. Liashenko, A. Piskorz, P. Komaromi, I. Gomperts, R. Martin, R.L. Fox, D.J. Keith T., Al-Laham, M.A. Peng, C.Y. Nanayakkara, A. Gonzalez, C. Challacombe, M. Gill P.M.W., Johnson, B.G. Chen, W. Wong, M.W. Andres, J.L. Head-Gordon, M. Replogle E.S. and Pople, J.A., Gaussian 98, revision A.9, Gaussian, Inc., Pittsburgh, PA, 1998.
31. Nose, S., J. Chem. Phys., 81 (1984) 511.
32. Feller, S.E., Zhang, Y., Pastor, R.W. and Brooks, B.R., J. Chem. Phys., 103 (1995) 4613.
33. Essman, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. and Pedersen, L.G., J. Chem. Phys., 103 (1995) 8577.
34. Ryckaert, J., Ciccotti, G. and Berendsen, H., J. Comput. Phys., 23 (1977) 327.
35. Charbonneau G.-P. and Delugeard, Y., Acta Crystallogr. B, 33 (1977) 1586.
36. Baudour, J.L., Acta Crystallogr. B, 47 (1991) 935.
37. Brown, C.J., Acta Crystallogr., 21 (1966) 146.
38. Bouwstra, J.A., Schouten, A. and Kroon, J., Acta Crystallogr. C, 39 (1983) 1121.
39. Harada, J., Ogawa, K. and Tomoda, S., Acta Crystallogr. B, 53 (1997) 662.
40. Brown, C.J. and Yadav, H.R., Acta Crystallogr. C, 43 (1987) 1087.
41. Becker, O.M., In Becker, O.M., MacKerell, A.D., Jr., Roux, B. and Watanabe, M., (Eds.), Computational Biochemistry and Biophysics. Marcel Dekker Inc., New York, 2001, pp. 82–83.