Theor Chem Acc (2012) 131:1274
DOI 10.1007/s00214-012-1274-z

REGULAR ARTICLE

First-principle-based MD description of azobenzene molecular rods

Silvio Pipolo · Enrico Benassi · Giorgia Brancolini · Michal Valášek · Marcel Mayor · Stefano Corni

Received: 26 July 2012 / Accepted: 30 August 2012 / Published online: 4 October 2012

Abstract Extensive density functional theory (DFT) calculations have been performed to develop a force field for the classical molecular dynamics (MD) simulations of various azobenzene derivatives. Besides azobenzene, we focused on a thiolated azobenzene's molecular rod (4′-{[(1,1′-biphenyl)-4-yl]diazenyl}-(1,1′-biphenyl)-4-thiol) that has been previously demonstrated to photoisomerize from trans to cis with high yields on surfaces. The developed force field is an extension of OPLS All Atoms, and key bonding parameters are parameterized to reproduce the potential energy profiles calculated by DFT. For each of the parameterized molecule, we propose three sets of parameters: one best suited for the trans configuration, one for the cis configuration, and finally, a set able to describe both at a satisfactory degree. The quality of the derived parameters is evaluated by comparing with structural and vibrational experimental data. The developed force field opens the way to the classical MD simulations of self-assembled monolayers (SAMs) of azobenzene's molecular rods, as well as to the quantum mechanics/molecular mechanics study of photoisomerization in SAMs.

Keywords Force field parameterization · DFT calculations · Azobenzenes · IR spectra

Electronic supplementary material The online version of this article (doi:10.1007/s00214-012-1274-z) contains supplementary material, which is available to authorized users.

S. Pipolo · E. Benassi · G. Brancolini · S. Corni (✉)
Centro S3, CNR Istituto di Nanoscienze, via G. Campi 213/a, 41125 Modena, Italy
e-mail: stefano.corni@nano.cnr.it

S. Pipolo
Dipartimento di Fisica, Università di Modena e Reggio Emilia, via G. Campi 213/a, 41125 Modena, Italy

M. Valášek · M. Mayor
Institute of Nanotechnology, Karlsruhe Institute of Technology, Hermann-von-Helmholtz Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

M. Mayor
Department of Chemistry, University of Basel, St. Johanns-Ring 19, 4056 Basel, Switzerland

1 Introduction

Molecules undergoing conformational changes upon interaction with light have the potential to be used as molecular machines [1–3]. In particular, when these molecules are orderly supported on a surface, they can express a considerable mechanical work that allows the execution of light-propelled actions at the nanoscale [4–6]. Molecular rods containing the azobenzene moiety (4′-{[(1,1′-biphenyl)-4-yl]diazenyl}-(1,1′-biphenyl)-4-thiol, thio-2-DA in the following) and supported on gold surfaces have been found to be optically and reversibly switchable with high yield from trans to cis and back [4, 5, 7]. Moreover, they were found to be able to perform surprising amounts of mechanical work upon isomerization from cis to trans [5]. The underlying microscopic mechanism behind these experimental findings is not completely understood. Due to the dense packing of the molecular rods, cooperative mechanisms for the photoisomerization from trans to cis have been suggested, but such cooperative mechanisms have not been microscopically characterized. On the contrary, it has been suggested that the delocalization of excitations may hamper the photoisomerization in compact self-assembled monolayers of azobenzene derivatives (azo-SAMs) [8]. Even the rate of the thermal back reaction

has been experimentally found to be affected by the SAMs environment.

Molecular mechanics simulations offer the best computational compromise to study thermal phenomena (such as thermal back isomerization from cis to trans) in (atomistically) large systems as a SAM. Molecular modeling at the classical atomistic level has already proved to be useful in the interpretation of the behavior of a SAM [9, 10]; in particular, it has been successfully applied to the study of elastic properties of azobenzene SAMs at the nanoscale [11].

While optical phenomena cannot be directly simulated with these models (unless mapping to ab initio results is used [12]), the proper parameterization and set up of a molecular mechanics (MM) model of azo-SAMs are a mandatory step toward QM/MM (quantum mechanics/molecular mechanics) calculations [13, 14]. This strategy has been in fact used for azobenzene by Bockmann et al. [15] who used density functional theory (DFT)-based ab initio molecular dynamics simulations to set up a GROMOS-like MM description of azobenzene (diphenyldiazene, 1-DA in the following). This description was later exploited to study azobenzene's containing systems in a variety of applications [16–18]. Moreover, a classical force field for a peptide derivative of azobenzene has been set up to simulate AFM experiments [19].

In this paper, we present the derivation of intramolecular classical force field parameters for (1) 1-DA, (2) its para-sulfanyl derivative (thio-1-DA), (3) bis[(1,1′-biphenyl)-4-yl] diazene (2-DA), and (4) its already introduced para-sulfanyl derivative thio-2-DA (see Figs. 1, 2), all in trans and cis forms, based on accurate DFT scans of the potential energy surface (PES) of such systems. In particular, we extend OPLS/AA (OPLS All Atoms) [20] by properly parameterize degrees of freedom that are specific to the studied molecules. While we have mentioned previous studies on force field parameterization of 1-DA [15, 19], we are not aware of any parameterization specifically developed for thio-1-DA, 2-DA, and thio-2-DA. Moreover, for each of the four parameterized molecule, we are proposing three sets of parameters, one specific for the trans form, one for the cis form, and one able to describe both.

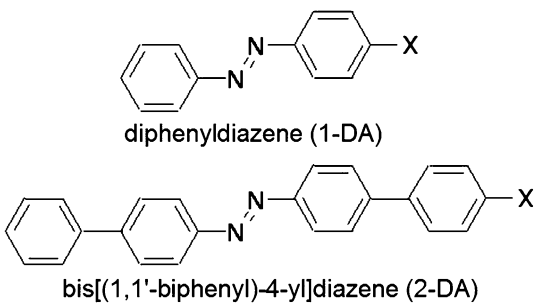

Fig. 1 X = H, SH. Names refer to X = H

To verify the quality of the resulting parameters, we compare calculated structures and vibrational spectra with experimental data, either from the literature or presented here. These comparisons show the reliability of the derived parameters: the equilibrium bond lengths, angles, and dihedrals for thio-2-DA are reproduced with RMS errors around 0.03 Å, 2.5°, and 6.5° (without considering the inter-ring dihedrals), respectively; the errors on vibrational frequencies ($\sim$90 cm^{-1}), although far from spectroscopic accuracy, are those expected for a class I force field [21]. The force fields derived here will be used in the future both to simulate ground state processes in azo-SAMs and as a basis to develop a QM/MM description of photoisomerization in azo-SAMs. The derived parameters are quite general and can be used to simulate these molecules in situations other than SAMs.

2 Methods

2.1 Overview of the force field form and parameterization strategy

The MM energy is calculated within a OPLS/AA-based force field. In detail, the contributions to the potential function of the system are the following: harmonic bonds, harmonic angles, harmonic improper dihedrals, and periodic dihedrals (that we express in the Ryckaert–Bellemans form) for the so-called bonding interactions, whereas 12-6 Lennard-Jones interactions and fixed point charges located at the atom positions were used to describe dispersion, repulsion, and electrostatics (i.e., "non-bonding") interactions. The functional form is thus the following:

$$\begin{aligned} E^{MM} = & \sum_{\text{bonds}} \mathbf{k}_b (b - \mathbf{b_0})^2 + \sum_{\text{angles}} \mathbf{k}_{\boldsymbol{\alpha}} (\alpha - \boldsymbol{\alpha_0})^2 \\ & + \sum_{\text{dih\_imp}} \mathbf{k}_{\boldsymbol{\delta}} (\delta - \boldsymbol{\delta_0})^2 + \\ & + \sum_{\text{dih}} \sum_{n=1}^{4} \frac{\mathbf{C_n}}{2} \left(1 + (-1)^{n-1} \cos(n\delta)\right) \\ & + \sum_{i<j} \frac{\mathbf{q_i q_j}}{r_{ij}^2} + \sum_{i<j} 4\varepsilon_{ij} \left[\left(\frac{\boldsymbol{\sigma}_{\mathbf{ij}}}{r_{ij}}\right)^{12} - \left(\frac{\boldsymbol{\sigma}_{\mathbf{ij}}}{r_{ij}}\right)^{6} \right] \end{aligned} \quad (1)$$

where "i" and "j" refer to atoms, the first four terms refer to "bonding" interactions, the last two to "non-bonding" interactions, and force field parameters are highlighted in bold.

Our starting point was the OPLS/AA-based force field proposed in Ref. [19]. Lennard-Jones (LJ) parameters for the benzene backbone are taken from the standard OPLS/AA set of parameters. They have been originally derived to simulate liquid benzene and have been successfully tested

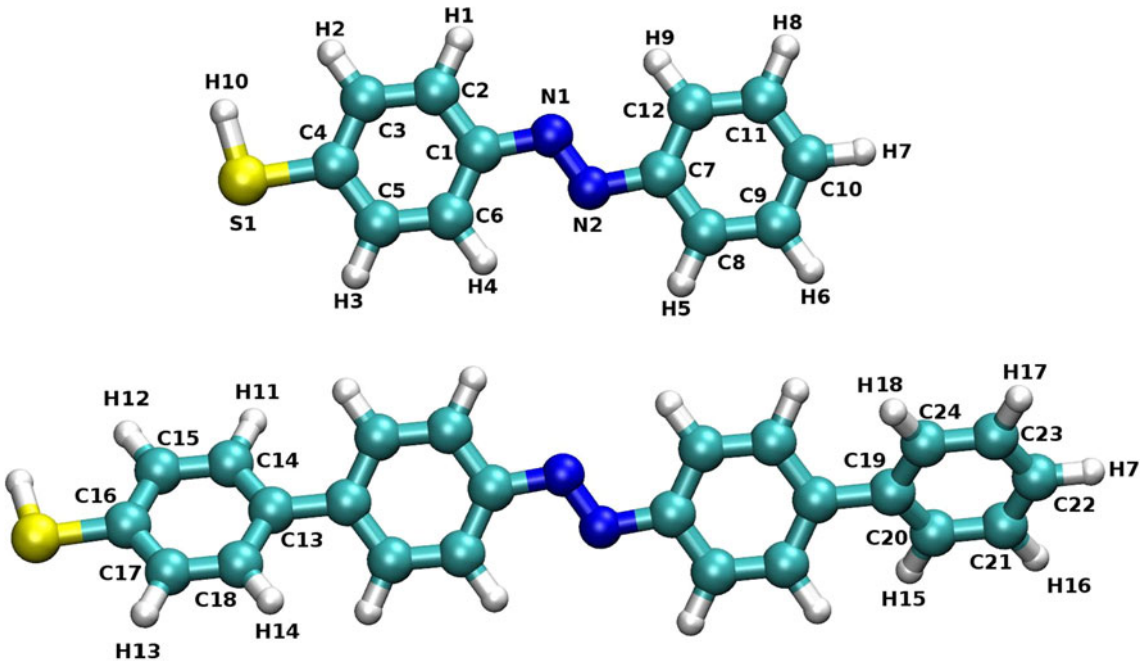

Fig. 2 Representation of 4-(2-phenyldiazenyl)benzenethiol (thio-1-DA, top) and 4′-{[(1,1′-biphenyl)-4-yl]diazenyl}-(1,1′-biphenyl)-4-thiol (thio-2-DA, bottom) in trans configuration. Atom labels are reported. Legend of colors: *white* hydrogen; *light blue* carbon; *dark blue* nitrogen; *yellow* sulfur

also within SAMs [11]. For the remaining LJ parameters, we refer to the work of Schäfer [19]. All LJ parameters are summarized in Table S.1 (see Online Resource 1). Atomic charges were obtained by a RESP procedure as described below. Here, we focus on the bonding part of the force field. The parameterization of such interactions is done by minimizing the differences between a scan of the quantum mechanics potential energy surface (QM-PES) and the corresponding molecular mechanics potential energy surface (MM-PES). The function $F^{\min}$ to be minimized is a sum on each point I of the scan

$$F^{\mathrm{Min}} = \sum_I \left(E_I^{\mathrm{MMred}} - E_I^{QM} + \sum_{\mathrm{bond-sel}} k_b (b_I - b_0)^2 + \sum_{\mathrm{angle-sel}} k_\alpha (\alpha_I - \alpha_0)^2 + \sum_{\mathrm{dih-sel}} \sum_{n=1}^{4} \frac{C_n}{2} \left(1 + (-1)^{n-1} \cos(n\delta_I)\right) \right)^2 \quad (2)$$

where:

$$E_I^{\mathrm{MMred}} = E_I^{MM} - \left(\sum_{\mathrm{bond-sel}} k_b (b_I - b_0)^2 + \sum_{\mathrm{angle-sel}} k_\alpha (\alpha_I - \alpha_0)^2 + \sum_{\mathrm{dih-sel}} \sum_{n=1}^{4} \frac{C_n}{2} \left(1 + (-1)^{n-1} \cos(n\delta_I)\right) \right) \quad (3)$$

E^{MMred} includes the classical potential energy terms that are not optimized during the minimization procedure. The MM-PES is a function of the force field parameters; here, the optimized subset (bond-sel, angle-sel, dihedral-sel) of the total set of used parameters are reported explicitly, together with their bonding functions, in the last three terms of the target function $F^{\min}$. Those parameters are optimized by a specifically implemented software in order to reach the best fitting of the QM-PES. To this aim, an extensive sampling of the QM-PES of the two systems in gas phase is done by means of scanning procedures of internal coordinates. Another equally viable technique would have been to avoid large PES scans (entirely or avoiding just the stiff, high force constant modes) and include instead hessian information from minimum energy conformations and/or gradient and hessian information along selected mode scans [21–25]. The QM information included in the fitting is comparable, and the overall computational cost is smaller for these approaches, but the scanning of QM-PES is inherently parallelizable which may be convenient depending on the kind of available computational resources.

The minimization procedure exploits the Levenberg–Marquardt algorithm implemented in the MINPACK [26, 27] library. The MM-PES needs to be evaluated at every minimization step; hence, a FORTRAN-based interface between GROMACS and MINPACK has been built and used. An equal weight is used for each point of

the MM-PES, whereas two criteria of convergence are established through the setting of a tolerance threshold, according to the "lmdif1.f" routine of minpack [26, 27]: the fitting procedure is stopped if the algorithm estimates that either the relative error on F^{min} between two consecutive iterations is at most equal to the value of the threshold or the relative error between the set of parameters at a certain iteration and the estimated optimal set is at most equal to the threshold. The tolerance parameter has been set here to 10^{-10}.

Internal coordinates are grouped into types in case they refer to chemically analogous atoms (see Table 1). As an example, we use the same harmonic function for the α2(C1N1N2) and the α2(C12N2N1) contributions to the MM-PES (i.e., the CNN harmonic function) and they belong to the same type, called CNN (see Fig. 2; Table 1). QM-PES scans of different coordinates belonging to the same type are considered together in the parameterization since they are used for the optimization of the same parameters (e.g., force constant and equilibrium angle of the CNN harmonic function).

Finally, since it has been documented that QM calculations at the B3LYP/cc-pVTZ level of theory overestimate vibrational frequencies [28], we scaled bond and angle force constants by a numerical factor (0.9392), the square of the value previously suggested for scaling vibrational frequencies at the level of theory used in the present calculation of QM-PESs [28]. The details of the parameterization procedure will be given in Sect 3.2.

2.2 QM computational details

A full geometry optimization of the electronic ground state of all the isolated species of Fig. 1 (with X = H, SH), both in trans and in cis conformation, was obtained in vacuo at Hartree–Fock (HF), density functional theory (DFT), and second-order Møller–Plesset perturbation theory (MP2) level using the following basis sets: 6-31G(d), 6-31+G(d), 6-31G(d,p), 6-31+G(d,p), 6-311G(d), 6-311+G(d), 6-311 G(d,p), 6-311+G(d,p), 6-311++G(d,p), cc-pVDZ, cc-pVTZ, and cc-pVQZ. For DFT calculations, the Becke three-parameter Lee–Yang–Parr (B3-LYP) exchange–correlation functional [29] was employed. Gaussian 09 computational package [30] was used for these calculations.

The optimized geometries were compared with published data (see Sect. 3.1), in order to identify the level(s) of theory providing the results in best agreement with the experimental ones. The approach which best matches with the X-ray data with a reasonable computational cost was DFT by employing cc-pVTZ basis set. Therefore, this basis set was selected for the subsequent calculations.

From the optimized geometries, some internal coordinates were selected and then scanned (with/without constrained optimization) at DFT B3-LYP/cc-pVTZ level of theory, in order to obtain the potential energy hypersurfaces profiles. The list of the selected internal coordinates is presented in Table 1. In particular, the isomerization reaction coordinates (i.e., the inversion angle α2 and the

Table 1 Internal coordinates subjected to QM scan, grouped in types

Type	Label	Range(s)/step(s)	Constrained optimization?
NN	b1(N1,N2)	Eq ± 0.50 Å/0.05 Å; eq ± 0.10 Å/0.01 Å	No
NC	b2(N1,C1)	Eq ± 0.50 Å/0.05 Å; eq ± 0.10Å/0.01	No
	b2(N2,C7)	Eq ± 0.50 Å/0.05 Å; eq ± 0.10 Å/0.01 Å	No
CS	b3(C4,S1)	Eq ± 0.50 Å/0.05 Å; eq ± 0.10 Å/0.01 Å	No
CCN	α1(C2,C1,N1)	Eq ± 10.0°/2.5°	No
	α1(C8,C7,N2)	Eq ± 10.0°/2.5°	No
CNN	α2(C1,N1,N2)*	Eq ± 180.0°/variable step	Yes
	α2(C7,N2,N1)*	Eq ± 180.0°/variable step	Yes
CCS	α3(C3,C4,S1)	Eq ± 20.0°/2.5°	No
	α3(C5,C4,S1)	Eq ± 20.0°/2.5°	No
CCC	α4(C4,C13,C14)	Eq ± 5.0°/2.5°	No
	α4(C9,C10,C19)	Eq ± 5.0°/2.5°	No
CCNN	δ1(C2,C1,N1,N2)	Eq ± 360.0°/7.5°	Yes
	δ1(C8,C7,N2,N1)	Eq ± 360.0°/7.5°	Yes
CNNC	δ2(C1,N1,N2,C7)*	Eq ± 360.0°/variable step	Yes
CCSH	δ3(C3,C4,S1,H10)**	Eq ± 120.0°/5.0°; eq ± 4.0°/2.0°	No
CCCC	δ4(C3,C4,C13,C14)	Eq ± 360.0°/7.5°	No
	δ4(C11,C10,C19,C24)	Eq ± 360.0°/7.5°	No

Legend of labels: b = bond length; α = bond angle; δ = dihedral angle. "range/step" indicates the range of the scan around the equilibrium value and the step used in the scan. "eq" stands for "equilibrium value." * Both trans and cis minima were considered as starting point for the scanning. Around such minima, the step was smaller than in the remaining portion of PES. ** For trans only

Table 2 Selected geometrical parameters of trans-1-DA

Label	*a* Theo.	*b* Theo.	*c* Theo.	*d* Theo.	*e* Theo.	*f* Theo.	*g* Theo.	*h* Exp.	*i* Exp.	*l* This work	*m* This work	*n* This work
b1(N1,N2)	1.253	1.24	1.259	1.243	1.243	1.268	1.267	1.247	1.268	1.258	1.248	1.247
b2(N1,C1)	1.428	1.42	1.438	1.423	1.422	1.417	1.420	1.428	1.427	1.419	1.417	1.417
α1(C2,C1,N1)	n.a.	n.a.	n.a.	115.8	115.5	115.1	115.3	115.3	n.a.	115.3	115.5	115.5
α1(C6,C1,N1)	n.a.	n.a.	n.a.	124.2	124.5	124.6	124.8	123.7	123.0	124.7	124.7	124.7
α2(C1,N1,N2)	116.8	115.0	114.2	115.0	115.1	113.7	114.8	114.1	114.5	115.2	115.5	115.6
δ1(C2,C1,N1,N2)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
δ2(C1,N1,N2,C7)	180.0	180.0	180.0	180.0	180.0	180.0	180.0	180.0	180.0	180.0	180.0	180.0

Bond lengths are in Å. Bond and dihedral are angles in degrees. *a:* CASSCF(6,6)/4-31G [42]; b: CASSCF(6,5)/double zeta with contraction (421/31) for N, minimal zeta with contraction (43/4) for C, uijneveldt's (5 s) for H [43]; *c:* CASSCF(6,6)/6-31G(d) for N and nearest C atoms, 6-31G otherwise [44]; *d:* CASSCF(10,8)/6-31G(d) [35]; *e:* CASSCF(14,12)/6-31G(d) [38]; *f:* MP2/cc-pVTZ [45]; *g:* DFT/BP86/TZVP [45]; *h:* Exp. (X-ray diffraction) [36]; *i:* Exp. (electron diffraction) [37]; *l:* DFT/B3-LYP/6-31 + G(d,p) [present work]; *m:* DFT/B3-LYP/cc-pVTZ [present work]; *n:* DFT/B3-LYP/cc-pVQZ [present work]

Table 3 Selected geometrical parameters of cis-1-DA

Label	*a* Theo.	*b* Theo.	*c* Theo.	*d* Theo.	*e* Theo.	*f* Theo.	*g* Exp.	*h* This work	*i* This work	*j* This work
b1(N1,N2)	1.24	1.240	1.242	1.242	1.261	1.255	1.253	1.249	1.239	1.239
b2(N1,C1)	1.43	1.472	1.357 (1.437)	1.435	1.432	1.437	1.449	1.437	1.433	1.433
α1(C2,C1,N1)	n.a.	n.a.	117.6 (117.7)	117.1	117.0	116.4	117.3	116.5	117.8	116.8
α1(C6,C1,N1)	n.a.	n.a.	121.9 (121.8)	122.4	122.2	122.9	122.5	122.9	122.7	122.7
α2(C1,N1,N2)	123.0	121.9	122.7	122.9	120.8	124.1	121.9	124.0	124.2	124.2
δ1(C2,C1,N1,N2)	56.0	60.8	62.8		53.6	48.4	53.3	51.0	52.2	52.2
δ2(C1,N1,N2,C7)	0.0	0.0	3.2	4.2	7.3	11.4	8.0	9.5	9.3	9.4

Bond lengths are in Å. Bond and dihedral angles are in degrees. *a:* CASSCF(6,5)/double zeta with contraction (421/31) for N, minimal zeta with contraction (43/4) for C, Duijneveldt's (5 s) for H [43]; *b:* CASSCF(6,6)/6-31G(d) for N and nearest C atoms, 6-31G otherwise [44]; *c:* CASSCF(10,8)/6-31G(d) [35]; *d:* CASSCF(14,12)/6-31G(d); CASPT2(14,12)/6-31G(d) [38]; *e:* MP2/cc-pVTZ [37]; *f:* DFT/BP86/TZVP [37]; *g:* Exp. (x-ray diffraction) [46]; *h:* DFT/B3-LYP/6-31 + G(d,p) [present work]; *i:* DFT/B3-LYP/cc-pVTZ [present work]; *j:* DFT/B3-LYP/cc-pVQZ [present work]

Table 4 Selected geometrical parameters of trans- and cis-thio-1-DA

Label	trans			cis		
	a	*b*	*c*	*a*	*b*	*c*
b1(N1,N2)	1.259	1.250	1.249	1.250	1.240	1.240
b2(N1,C1)	1.413	1.411	1.410	1.433	1.429	1.428
b2(N2,C7)	1.417	1.416	1.415	1.436	1.432	1.431
b3(C4,S1)	1.776	1.774	1.770	1.783	1.776	1.772
α1(C2,C1,N1)	115.8	115.9	115.9	116.0	116.3	173.7
α1(C8,C7,N2)	115.4	115.5	115.5	116.9	117.1	117.2
α2(C1,N1,N2)	115.5	115.5	115.6	124.4	124.5	124.5
α2(C7,N2,N1)	115.6	115.5	115.6	124.1	124.4	124.4
α3(C3,C4,S1)	122.9	123.1	123.1	123.0	123.1	123.1
α3(C5,C4,S1)	117.4	117.5	117.5	117.8	117.7	117.7
δ1(C2,C1,N1,N2)	180.0	180.0	180.0	144.4	142.9	142.9
δ1(C8,C7,N2,N1)	180.0	180.0	180.0	52.6	53.4	53.4
δ2(C1,N1,N2,C7)	180.0	180.0	180.0	10.2	10.1	10.1

Bond lengths are in Å. Bond and dihedral angles are in degrees. *a:* DFT/B3-LYP/6-31 + G(d,p); *b:* DFT/B3-LYP/cc-pVTZ; *c:* DFT/B3-LYP/cc-pVQZ

Table 5 Selected geometrical parameters of trans- and cis-thio-2-DA

Label	2-DA				thio-2-DA			
	trans		cis		trans		cis	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*
b1(N1,N2)	1.261	1.251	1.251	1.240	1.261	1.251	1.249	1.241
b2(N1,C1)	1.415	1.413	1.433	1.430	1.415	1.412	1.431	1.429
b2(N2,C7)	1.415	1.413	1.433	1.430	1.415	1.412	1.431	1.429
b3(C4,S1)	–	–	–	–	1.785	1.778	1.785	1.778
α1(C2,C1,N1)	115.7	115.9	116.7	117.1	115.7	115.9	116.7	116.8
α1(C8,C7,N2)	115.7	115.9	116.7	117.1	115.7	115.9	116.6	117.0
α2(C1,N1,N2)	115.2	115.5	124.1	124.3	115.2	115.6	124.1	124.4
α2(C7,N2,N1)	115.2	115.5	124.1	124.3	115.3	115.5	124.1	124.4
α3(C3,C4,S1)	–	–	–	–	123.1	123.2	123.1	123.2
α3(C5,C4,S1)	–	–	–	–	118.0	118.0	118.0	118.0
α4(C4,C13,C14)	120.9	120.9	120.9	120.9	121.1	121.2	121.2	121.2
α4(C9,C10,C19)	120.9	120.9	120.9	120.9	121.1	121.1	121.2	121.2
δ1(C2,C1,N1,N2)	179.0	178.2	48.9	50.5	178.9	178.1	140.0	140.1
δ1(C8,C7,N2,N1)	179.0	178.2	48.9	50.5	178.1	178.0	138.1	138.2
δ2(C1,N1,N2,C7)	179.9	179.9	10.9	10.6	179.8	179.8	10.4	10.4
δ3(C3,C4,S1,H10)	179.9	179.9	10.9	10.6	0.1	0.9	2.1	2.5
δ4(C3,C4,C13,C14)	38.9	37.6	37.4	37.8	37.5	36.1	37.2	36.9
δ4(C11,C10,C19,C24)	38.9	37.6	37.4	37.8	38.9	37.7	38.4	38.7

Bond lengths are in Å. Bond and dihedral angles are in degrees. *a:* DFT/B3-LYP/6-31 + G(d,p); *b:* DFT/B3-LYP/cc-pVTZ

rotation angle δ2) were scanned with higher accuracy respect to the others, with the following principles: (1) scanning included relaxation of the remaining molecular coordinates, (2) shorter steps, especially around trans and cis minima were applied, (3) inclusion of both isomerization directions (i.e., from trans to cis and from cis to trans) was considered. In case of bond lengths, the coordinates were first scanned including a wider range (with a wider step) and then a finer scanning around the minima, in the quasi-harmonic region, was applied. All these results were used for setting up the force field. For the central moiety of thio-2-DA, we have assumed the same potential energy hypersurfaces profiles of thio-1-DA. Gaussian 09 computational package [30] was used for these calculations.

The optimized geometries were submitted to the calculation of the RESP charges. RESP and ESP charges to be included in classical force fields can be derived with different strategies. In this work, RESP charge derivation involved three steps (1) geometry optimization, (2) molecular electrostatic potential (MEP) computation using the optimized geometry from the first step, and (3) fitting of the charges centered on the atoms to the MEP resulted in the second step. Molecular symmetry and charge equivalency have been imposed during the charge derivation procedures. The MEP calculation is done through Connolly's surface [31], using the GAFF-like RESP derivation [32]. Data are reported in Table 8. Four additional sets of charges are obtained by combining different methodologies for the calculation of the MEP and the fitting procedure; a detailed description and results are reported in the "Online Resource 1."

2.3 Molecular dynamics computational details

The MM-PES scan is obtained by evaluating the molecular mechanic energy of all the QM structures defined by the scanning procedure described above. A geometrical combination rule is used for the Lennard-Jones parameters (i.e., the parameters for each couple of interacting atoms are calculated as the square root of the product of the single-atom tabulated parameters) and the non-bonding 1–4 interactions (the non-bonding interaction between the two external atoms involved in the definition of a proper dihedral) are halved. Shifted Lennard-Jones and PME-Coulomb were used for the van der Waals and electrostatic interactions. The shifted cutoff for the Lennard-Jones potential is applied between 1.1 and 1.2 nm, and a direct-space cutoff of 1.2 nm is used for the PME-Coulomb contribution. Atomic charges are derived as mentioned above. The molecule is embedded in a 6.090 nm × 6.153 nm × 7.074 nm rectangular box with periodic boundary conditions applied in all of the three directions,

Table 6 Partial comparison between the experimental (Exp.) [4] and the computed (QM and MM) geometries for substituted trans-2-DA-X

Label	Exp. X = $SOCH_3$	QM X = SH	MM X = SH
b1(S1,C16)	1.816	1.778	1.780
b2(N1,N2)	1.258	1.251	1.265
b3(N1,C1)	1.447	1.412	1.417
b3(N2,C7)	1.449	1.412	1.412
α1(S1,C16,C15)	121.1	123.2	118.8
α1(S1,C16,C17)	119.2	117.9	118.6
α2(N1,N2,C7)	114.5	115.5	115.5
α3(N1,C1,C2)	115.6	115.9	116.7
α3(N1,C1,C6)	125.2	125.0	123.9
α2(N2,N1,C1)	113.5	115.4	115.2
α3(N2,C7,C8)	115.2	115.9	117.2
α3(N2,C7,C12)	125.1	125.0	124.0
δ1(S1,C16,C15,C14)	175.5	180.0	179.5
δ1(S1,C16,C17,C18)	176.3	179.8	179.7
δ2(N1,N2,C7,C8)	169.5	178.0	179.8
δ2(N1,N2,C7,C12)	9.7	2.1	0.5
δ3(N1,C1,C2,C3)	179.5	179.9	179.8
δ3(N2,N1,C1,C2)	169.5	178.1	179.8
δ4(N2,N1,C1,C6)	10.6	2.0	0.1
δ3(N2,C7,C8,C9)	179.4	179.8	179.8
δ3(N2,C7,C12,C11)	179.1	179.9	179.6
δ5(C1,N1,N2,C7)	179.9	179.8	179.7

QM level of theory: DFT B3-LYP/cc-pVTZ. Bond lengths are in Ångrstrøm. Bond and dihedral are angles in degrees

Table 7 Partial comparison between the experimental (Exp.) [4] and the computed (QM and MM) geometries for trans-1DA-X

Label	Exp. X = H	QM X = H	MM X = SH
b1(N1,N2)	1.247	1.248	1.265
b2(N1,C1)	1.428	1.417	1.421
α1(C2,C1,N1)	115.3	115.5	115.9
α1(C6,C1,N1)	123.7	124.7	124.0
α2(C1,N1,N2)	114.1	115.5	116.0
δ1(C2,C1,N1,N2)	0.0	0.0	0.0
δ2(C1,N1,N2,C7)	180.0	180.0	179.9

QM level of theory: DFT B3-LYP/cc-pVTZ. Bond lengths are in Ångrstrøm. Bond and dihedral angles are in degrees. Terminal group "X" refers to Fig. 1

which is meant to reproduce the most common situation where this force field will be used (i.e., to simulate SAMs). In other words, we consider throughout this article a single molecule inside the simulating box, effectively isolated by its replica thanks to the large box size. At the same time, by using periodic boundary conditions and such large box

Table 8 RESP charges for thio-2-DA

Atom	trans	cis	Atom	trans	cis
N1	−0.209	−0.294	H1	0.077	0.161
N2	−0.202	−0.277	H2	0.118	0.127
C1	0.34	0.497	H3	0.118	0.127
C2	−0.171	−0.271	H4	0.077	0.161
C3	−0.085	−0.141	H5	0.094	0.156
C4	−0.097	0.012	H6	0.119	0.126
C5	−0.085	−0.141	H7	0.105	0.111
C6	−0.171	−0.271	H8	0.119	0.126
C7	0.333	0.427	H9	0.094	0.156
C8	−0.182	−0.230	H10	0.139	0.141
C9	−0.094	−0.159	H11	0.127	0.108
C10	−0.047	0.020	H12	0.142	0.150
C11	−0.094	−0.159	H13	0.142	0.150
C12	−0.182	−0.230	H14	0.127	0.108
C13	0.246	0.100	H15	0.109	0.104
C14	−0.24	−0.143	H16	0.11	0.126
C15	−0.081	−0.118	H17	0.11	0.126
C16	−0.003	−0.025	H18	0.109	0.104
C17	−0.081	−0.118			
C18	−0.24	−0.143			
C19	0.13	0.074			
C20	−0.149	−0.092			
C21	−0.11	−0.174			
C22	−0.105	−0.073			
C23	−0.11	−0.174			
C24	−0.149	−0.092			
S1	−0.198	−0.173			

Atom labels refer to Fig. 2

sizes, we are performing the parameterization procedure with a computational setup as close as possible to the one that will be used to simulate the SAMs. We note that imposing 3D periodic boundary conditions (as opposed to 2D) for a SAM is computationally convenient as it allows to use fast 3D Ewald-based methods for electrostatics; on the other hand, it creates a fictitious periodicity in the perpendicular direction that requires specific corrections [33].

Test MD simulations are run with a time step of 1 fs within the canonical ensemble (NVT). The Nosé-Hoover thermostat is used with a reference temperature of 300 K and a time constant for T-coupling of 200 fs. Constraints to length of bonds involving hydrogens are applied via the LINCS algorithm (standard parameters are used), bonding and non-bonding potential energy functions are the same as the MM-PES calculations, charges are kept fixed during the simulations. All molecular dynamics and molecular mechanics calculations are performed with the GROMACS 4.5.4 package [34].

2.4 Experimental IR spectra

IR spectrum of acetyl-protected thio-2-DA (P-2-DA) was recorded with a PerkinElmer Spectrum GX FT-IR spectrometer using KBr micro-pellet technique.

3 Results and discussion

3.1 QM single molecule calculations

In Tables 2 and 3, we collect the equilibrium values for the main internal coordinates (see Table 1), as obtained after the geometry optimization, in comparison with published data. For the 1-DA geometries, a good agreement is obtained between the predicted geometries and the experimental data when the DFT B3-LYP/cc-pVTZ level of theory is employed. A good matching is also obtained at DFT B3-LYP/cc-pVQZ level but the computational cost is large. Also at B3-LYP/6-31+G(d,p) level, the agreement is satisfactory, especially on bonds distances and angles. We have therefore used DFT B3-LYP/cc-pVTZ level of theory for the next computations, since it provides reliable results and is more computationally convenient than cc-pVQZ.

We remark that according to previous computational studies [35] and experimental characterisations [36, 37], our calculations did provide not equal bond lengths for the two benzene rings. For trans-1-DA, we observe shorter and longer bonds alternating each other. This is predominantly remarked from our calculation, at HF and DFT level, especially employing large basis sets. MP2 does not show such behavior. Concerning with this aspect, CASSCF gives different pictures, depending on the active space: data obtained at CASSCF(10,8)/6-31G(d) level [35] agree with HF and DFT ones, data obtained at CASSCF(14,12)/6-31G(d) level [38] agree with MP2 ones. In the cis conformer, the differences in bond lengths of the same ring are less surprising, since they have been documented at all the different levels of theory applied so far.

We are not aware of any experimental structure for thio-1-DA. Anyway, we are quite confident that the obtained DFT structure is reliable, since other studies [39, 40] on push–pull para-substituted 1-DAs show that the geometry (especially in the central moiety) is not particularly affected by the presence of the substituents [35, 39, 40]. In Table 4, we report the optimized (equilibrium) values for the selected parameter for trans- and cis-thio-1-DA, obtained with three different basis sets.

Let us now consider the 2-DA systems. Table 5 summarizes the selected parameters for the two conformers of the non-thiolated and thiolated-2-DA. The experimental structure of 2-DA and thio-2-DA is not available; however, a structure for 2-DA substituted with thioacetate (X = SOCH3) has been reported [4]. Table 6 collects some of the results. (Further data are available in the "Online Resource 1," Table S.5). Some important points may be highlighted. First of all, as assumed for thio-1-DA, also in this case, the presence of the SH group does not remarkably affect the geometry of thio-2-DA in comparison with 2-DA. In general, we can say that the optimized geometry is in good agreement with the experimental one. Actually, the average percent error on the bond lengths (angles) is about 0.9 % (0.6 %), that is, absolute errors are of the order of 0.015 Å (0.5°). Also, the dihedral angles are well reproduced from a computational point of view, at the selected level of theory, with exception for phenyl–phenyl dihedrals. We actually see from the experimental data that the two phenyls are predicted to be coplanar, while the theoretical structure provides non-null dihedral angles. Since 1-DA and the other derivatives present an inherent helicity expressed by the cis-azobenzene isomers which may be of energetically equivalent P- or M- kind, in this QM study, we have considered only cis-P-atropisomers. Moreover, during the investigation of the azobenzene derivatives properties, we observed that enlarging the system size by adding phenyls, the molecular geometry deviates from the planarity because of the torsions of the external phenyls with respect to the central moiety. This torsion causes the molecules 2-DA and 3-DA (bis[(1,1′:4′,1′′-terphenyl)-4-yl]diazene, not presented in this work), in both trans and cis forms, to assume a spiral shape, giving rise to different rotamers. There exists four (2^2) stable rotamers for trans- and eight ($2*2^2$) stable rotamers for cis-2-DA, four deriving from cis-P-1-DA, and four from cis-M-1-DA. For 3-DA, there exists eight (2^3) stable rotamers for trans and sixteen ($2*2^3$) stable rotamers for cis-3-DA. Anyhow, up to now, we have only investigated the all-R rotamers. The experimental structure, on the opposite, shows the two phenyls as coplanar, with small dihedral angles, likely because the packing in the crystal favors such conformation [41]. Our calculations show that only 8 kJ/mol is needed to make the biphenyl rings coplanar. It might also be that in the X-ray structure, an equidistributed statistical average of R- and S-rotamers is pictured. As for 1-DA, also for 2-DA shorter and longer bonds alternate each other in phenyl rings, especially for the innermost rings.

3.2 MM parameterization

The force field parameterization, based on QM-PES scans, has been first performed for 1-DA. Then, the resulting parameters were used for the core part of 2-DA (i.e., that is common to 1-DA), while the other 2-DA parameters ("shell" parameters) have been optimized by using the 2-DA QM-PES scans (Fig. 3). Nearly, all internal coordinates are scanned twice, in one case starting from the cis equilibrium conformation and in the second case, starting

from the trans equilibrium conformation (the CCS angle and the CCSH dihedral are the only exceptions). On this base, a configuration-specific set of core parameters is optimized for both the cis- and the trans-minima. These parameters are most suitable to describe situations where the interconversion between cis and trans form does not take place in the timescale of the simulation. Moreover, a single set of parameters able to describe both trans and cis species has been obtained exploiting all the collected 1-DA QM-PES scans. Finally, an analogous procedure is applied to the 2-DA system leading to both conformer-specific and a single common set of shell parameters. The overall procedure is summarized in Fig. 3. The final parameter sets are tested against different sets of charges (see "Online Resource 1"): the agreement between QM-PES samples and MM-PES is not significantly altered once the MM energy is shifted by a proper constant.

3.2.1 *Parameterization of 1-DA*

3.2.1.1 Distinct sets of parameters for each isomer The following optimization steps are performed for each of the two conformers by exploiting the 1-DA QM-PES scans. As a first step, the parameters for the NN, CN, CS, and CCN function types are optimized starting from the literature values as guesses [19]. Since the corresponding internal coordinates are scanned through a frozen procedure (see Sect. 2.2), these parameters are optimized first and then they are used in the following parameterization steps.

The second step is the optimization of the parameters for the CNN, CNNC, CCNN, CCS, and CCSH potential functions (the last two functions only for the trans set, assuming that for cis, they do not change significantly). This is done by first optimizing the parameters for each of these function types alone, then a refinement is achieved by a fitting procedure where all the 1-DA QM-PES are joined together and all the parameters are optimized at once.

3.2.1.2 Set of parameters common to both isomers ("Mixed") The single set of common parameters is obtained by repeating the two-step procedure on the unified (i.e., cis + trans) QM-PES scans. Non-bonding function parameters are modified as follows: common RESP charges are obtained from the arithmetic mean of the conformer-specific ones, whereas for the LJ parameters, a geometric mean is used (the only parameter that differs from cis to trans in the literature parameterization is the "ε" value of the carbon atoms).

Two optimization procedures are performed, in one case starting from the trans and in the other case from the cis set of optimized parameters as initial guess (see Fig. 3). At the end of each step, the best subset of parameters (in terms of mean error per structure) is chosen.

3.2.2 *Parameterization of 2-DA*

3.2.2.1 Distinct sets of parameters for each isomer Regarding the set of conformer-specific shell parameters, a two-step procedure is applied as for 1-DA. First, the parameters

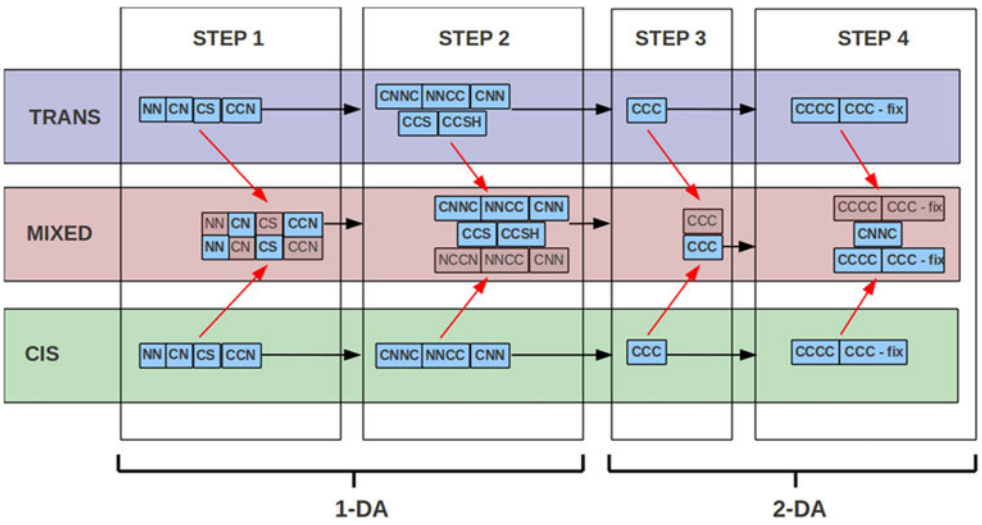

Fig. 3 Sketch of the parameterization procedure. Each *rectangular box* refers to a parameter type as defined in Table 1. *Black arrows* stand for the inter-step transfer of optimized parameters that will be kept fixed in the subsequent step, while red arrows indicate the transfer of parameters to be used as the initial guess for the subsequent minimization process. The label "CCC-fix" in the "STEP 4" refers to the choice of adding the PES sample relative to the CCC harmonic function in the minimization procedure without changing the corresponding parameters (an explanation to this choice is given in the text). *Blue boxes* refer to the final optimized set of parameters described in tables 9 and 10, and *red boxes* refer to the rejected set of parameters within mixed optimization procedure

of the CCC harmonic function are optimized starting from a reasonable guess and employing the QM-PES scans relative to the α4-type (C14C13C4, C13C4C3, C11C10C19, and C10C19C14) internal coordinates; after that the parameters for the CCCC dihedral potential are optimized by employing both the α4 and δ4 internal coordinates scans and replacing the CCC guess parameters with the ones optimized in the previous step. We chose this strategy because of the large difference in the total number of scan points for the two internal coordinate types (α4 scan has 20 points and δ4 scan has 96 points). We have verified that a parameterization step with all the parameters optimized at once would have led to a reasonable set of parameters for the CCCC function and a distorted set of parameters for the CCC function.

3.2.2.2 Set of parameters common to both isomers ("Mixed".) As for 1-DA, the common set of shell parameters is obtained by repeating the two-step procedure described above with an unified collection of QM-PES scans and an additional one regarding the δ2 internal coordinate for both the conformers. The need for a further scan of this internal coordinate relies in an inaccurate description of the cis–trans energy difference for the 2-DA system by the isomer-specific parameters. The inclusion of the core (i.e., 1-DA) parameters of the CNNC function in both of the steps of the minimization procedure and the use of the additional δ2 scan allows us to reproduce the cis–trans minimum energy difference by properly modifying the core CNNC dihedral parameters. The remaining core parameters are taken from the set of common parameters optimized with the 1-DA PES scans. As pointed out for 1-DA, in order to obtain a full common set of parameters, other non-optimized force field parameters that differ from cis to trans both in the starting point parameterization [2], and in the present work calculations are arithmetical averaged (RESP charges and bonding parameters) or geometrical averaged (non-bonding parameters).

Tables 9 and 10 show the optimized set of core and shell parameters, together with the mean error per scan point. The comparison of the QM-PES and MM-PES, given in the "Online Resource 1" (Fig. S.1-8), shows the good quality of the fitting.

3.3 Comparison of MM results with experimental data

3.3.1 Structure of 1-DA

The geometrical parameters obtained from the MM force field are reported in Table 7. The agreement is good with the targeted DFT calculations and, in turn, with experimental derived parameters. Differences in the bond lengths ($\sim$0.02 Å, comparable to X-ray—electronic diffraction discrepancies) and angles ($\sim$1.6°), confirm the quality of the MM description.

3.3.2 Structure of 2-DA and thio-2-DA

As stated above, no experimental data are available on the structure of 2-DA or its thiolated form. However, the protected form of trans-thio-2-DA has been studied by X-ray. We report in Table 6 experimental data compared to MM results. Again, as for 1-DA, the comparison shows the good quality of the force field. The most notable discrepancy is on the dihedral angles NNCC, which is however relatively flexible and may be affected by the crystal environment in the experiment. For 1-DA, the same angle has a zero value also in the experiments.

3.3.3 Vibrational spectrum of trans-thio-2-DA

Harmonic vibrational frequencies for the thio-2-DA molecule are obtained within the classical framework by diagonalizing the hessian matrix (MMh in the following). We also calculated anharmonic vibrational frequencies obtained from a dynamical trajectory for the single molecule by Fourier transforming the correlation function of several internal coordinates (MMc). The diagonalization procedure is performed with the GROMACS tools [47], whereas correlation functions are calculated for a 0.5 ns long simulation in vacuum, after 1 ns of equilibration at room temperature, with parameters specified in Sect. 3.2; timestep was 0.5 fs, and the time constant for T-coupling was 100 fs.

Measuring the vibrational spectrum of thio-2-DA is impractical since it rapidly undergoes oxidation in air. Therefore, we have measured the vibrational spectrum of the acetyl-protected form of thio-2-DA (P-2-DA—namely thioacetic acid-S-(4′-{[(1,1′-biphenyl)-4-yl]diazenyl}-[(1,1′-biphenyl)-4-yl]) ester), the same for which the X-ray structure was determined. By comparing the experimental spectrum with the one at the QM level (see "Online Resource 1"), we assign a normal mode to each of the selected experimental bands; selected absorptions are chosen such that to be easily associable to normal modes localized on the azobenzene skeleton; furthermore, they show a signal intensity of at least 30 % of the most intense peak. Then, a one-to-one correspondence is also found between QM normal modes of P-2-DA and MM normal modes of thio-2-DA in order to compare experimental vibrational frequencies of P-2-DA and classical vibrational frequencies of thio-2-DA. Data are presented in Table 11.

Frequency values related to experimental and MMc spectra are referred to the maximum of the absorption peak. Concerning experimental data, where multiple absorptions are expected, a deconvolution is performed

Table 9 1-DA parameters sets. Conformer-specific (CIS, TRANS) and common (MIXED) sets

	NN			CN		
	b_0	k_0	Err	b_0	k_0	Err
TRANS	0.12539	592,426.73	1.36	0.14022	268,915.69	0.62
MIXED	0.12459	579,520.93	2.92	0.14124	237,626.87	1.49
CIS	0.12393	600,439.37	1.29	0.14253	233,213.57	0.53

	CS			CCN			CNN	
	b_0	k_0	Err	α_0	k_0	Err	α_0	k_0
TRANS	0.17780	191,498.66	0.30	119.974	444.717	0.62	110.364	625.080
MIXED	0.17796	193,810.82	0.32	119.127	364.862	0.77	108.613	375.861
CIS	0.17815	192,554.56	0.30	115.803	303.860	0.49	114.740	431.584

	CCS		CCNN				
	α_0	k_0	C1	C2	C3	C4	Err
TRANS	119.580	483.925	−0.77	19.20	−1.93	−0.69	0.63
MIXED	115.354	486.495	−0.43	18.38	−9.02	0.00	1.35
CIS			−0.01	19.94	−9.82	0.04	0.86

	CNNC				CCSH				
	C1	C2	C3	C4	C1	C2	C3	C4	Err
TRANS	0.03	138.91	−0.13	−9.98	0.13	6.09	0.60	0.01	0.63
MIXED	16.64	132.05	−2.42	−9.54	1.18	5.87	0.40	0.07	1.35
CIS	0.00	153.22	29.81	3.28					0.86

The Error (Err—kJ mol^{-1}) is a mean unsigned error per PES sample point. Values of k_0 in kJ mol^{-1} nm^{-2} for bonds and kJ mol^{-1} rad^{-2} for angles, b_0 in nm, α_0 in °, C1-4 in kJ mol^{-1}

Table 10 Parameters sets for 2-DA coordinates not present in 1-DA. Conformer-specific (CIS and TRANS) and common (MIXED) sets

	CCC		CCCC				
	α_0	k_0	C1	C2	C3	C4	Err
TRANS	117.516	471.971	1.62	4.17	−0.11	−1.60	0.14
MIXED	117.271	431.966	0.42	4.00	0.07	−1.65	0.19
CIS	121.901	428.482	0.50	3.88	0.08	−1.66	0.12

	CNNC				Err
	C1	C2	C3	C4	
TRANS					
MIXED	3.22	98.66	−0.01	−2.97	0.59
CIS					

The Error (Err—kJ mol^{-1}) is a mean error per PES sample point. Values of k_0 in kJ mol^{-1} rad^{-2}, α_0 in °, C1-4 in kJ mol^{-1}

using Gaussian functions; otherwise, the definition of MMc frequency in case of multiple bands is done by exploiting absorption spectra addressed to different internal coordinates.

A graphical representation of data shown in Table 11 is presented in Fig. 4, and a mean error (RMS error) is reported for QM, MMc, and MMh calculation methods in Table 12, together with linear fitting information.

A good agreement is found between QM and experimental data, and both the employed classical approaches are providing the same results: calculated absorption frequencies are correlated with the experimental ones with some deviations from the ideal behavior. The average error (around 90 cm^{-1}) is that expected for a class I force field [21]. The major deviations (higher than 150 cm^{-1}) are encountered for modes 23, 24, and 25, and regression data

Table 11 Selected vibrational frequencies of trans-2-DA (cm^{-1}) and normal mode description

Nr	Exp	QM	MMh	MMc	Description*
1	501	495	478	477	CCN, CNN, W16b
2	530	525	519	516	W17b, W16b, CCN
3	534	532	529	528	CCN, CNN, CCC
4	549	549	549	540	W17b, W16b, CNN
5	691	695	688	686	W4
6	720	725	777	777	W6a, CNN, CC
7	727	732	733	727	W5, W4
8	737	749	699	695	W5
9	769	769	785	779	W5
10	821	828	893	883	W16b
11	828	834	898	889	W10a
12	846	856	916	903	W17b
13	857	867	925	912	W17b
14	1,001	994	1,031	1,031	W12, W18a
15	1,012	996	1,059	1,060	W12, W18a
16	1,077	1,074	1,171	1,173	CS, W18a
17	1,147	1,125	1,248	1,250	CN, CNN, CCC, CCH
18	1,163	1,146	1,124	1,137	W9a, CN
19	1,182	1,172	1,115	1,126	W9a
20	1,192	1,175	1,112	1,118	W9a
21	1,233	1,224	1,265	1,265	CN, W18a
22	1,479	1,475	1,535	1,535	W19a
23	1,593	1,590	1,873	1,876	W8a CC
24	1,595	1,595	1,758	1,767	W8a
25	1,599	1,599	1,767	1,758	W8a

* Benzene ring vibration is labeled as starting with "W" according to Wilson's classification [48, 49]. Two letters (e.g., CC) refer to bond stretching, Three letters (e.g., CNN) refer to angle bending. "Exp" and "QM" refer to P-2-DA (X = $SOCH_3$ in Fig. 1), and MMh and MMc refer to thio-2-DA (X = SH in Fig. 1)

in Table 12 suggest an overestimation of the frequencies in the high-frequency part of the spectrum.

We point out that the frequency of some normal modes of conjugated molecules are not well reproduced by harmonic bond potentials, without cross-terms between internal coordinate (like that we are using). This is the case of Kekulè-like modes of benzene rings [50] and generally of modes that involve the out-of-phase stretching of vicinal bonds in a delocalized framework.

Normal modes 23, 24, and 25 belong to this class; in fact, these vibrations involve the out-of-phase CC stretching intra-ring and inter-rings (only mode 23). An analogous explanation is valid for the overestimation of the frequency of the mode 17 in which the CN stretching are involved, the overall effect is anyhow less pronounced. Table 13 shows the error and the fitting information for a set of data that do not include frequencies of normal modes 23, 24, and 25.

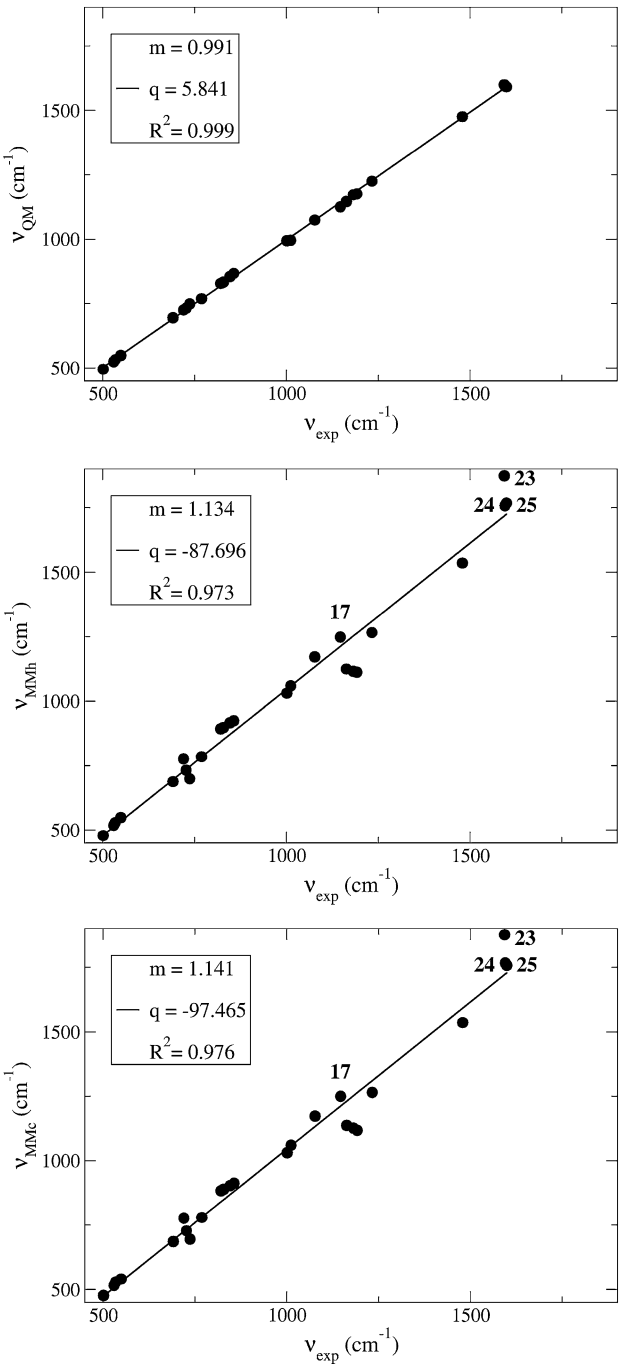

Fig. 4 Comparison between calculated and experimental frequencies. From top to bottom: QM, MMh, MMc. The *line* shows a linear fitting of the data; m: slope; q: intercept; R2: square of the correlation coefficient

Table 12 Complete set, error, and linear fitting data

	RMS error (cm^{-1})	Slope (m)	Intercept (q; cm^{-1})
QM	9.44	0.991 ± 0.005	6 ± 6
MMh	88.9	1.13 ± 0.04	−88 ± 41
MMc	87.5	1.14 ± 0.04	−97 ± 39

Table 13 Reduced set, error, and linear fitting data

	RMS error (cm^{-1})	Slope (m)	Intercept (q; cm^{-1})
MMh	54.0	1.02 ± 0.04	−1 ± 39
MMc	50.6	1.04 ± 0.04	−14 ± 36

The set of parameters common to both isomers (Sect. 3.2.2) reproduces the selected vibrational frequencies of the trans-2-DA at the same level of accuracy of the dedicated set. In fact, the RMS error calculated within the hessian approach (MMh) for the reduced set is 55.2 cm^{-1}, very similar to the RMS errors reported in Table 13 for the specific trans parameters. A detailed list of vibrational frequencies calculated with the set of parameters common to both isomers (Table S8) and further information on the fitting data can be found in the Online Resource 1.

4 Conclusions

In this article, we have presented an ab initio-based derivation of a classical force field for different azobenzene derivatives. The targeted molecules are important photoswitches and have been used to translate optical into mechanical energy (i.e., optomechanical applications). We have used a parameterization procedure based on the extensive scans of the PES of such molecules at the DFT level. The force field that we are proposing reproduces the experimental structures of the target molecules. Moreover, experimental vibrational frequencies (presented here for a protected form of thio-2-DA) were reproduced to the expected degree of accuracy (although not at the spectroscopic accuracy level). In producing the force field for the various molecules, attention has been paid not only to reproducing well the QM-PES of each single molecules, but also to provide rather transferable parameters by defining parameters type, each corresponding to different internal coordinates, or to the same internal coordinate in different molecules. This allows to treat derivatives of the molecules studied here (e.g., including ring substituents) with a reasonable reliability without having to repeat the cumbersome parameterization of the backbone of the azobenzene molecules. Thanks to the parameters presented here; we have provided the tools to simulate the behavior of azo-SAMs at the classical level and to study their photophysics with QM/MM methods.

A possible line of evolution of this work would be the parameterization of a coarse-grained force field in order to simulate azobenzene SAMs behavior with a lower detail of description but at longer timescales. At this level of description, azobenzene rods can be also simulated using two Gay-Berne disks [51] and the isomerization reaction can be introduced as a cyclic switching between two force field models as proposed by Duchstein et al. [52]. Such development would enrich the modeling of azobenzene rods of a further level beside the ab initio and the classical atomistic MD presented here.

Acknowledgments We acknowledge computational time and support from CINECA under the ISCRA initiative. Funding from EU NanoSciE + project under the Transnational grant Maecenas is gratefully acknowledge. We thank M.A. Rampi, M. Persico, and G. Granucci for useful discussions.

References

1. Balzani V, Venturi M, Credi A (2004) Molecular devices and machines: a journey into the Nanoworld. Wiley, Weinheim
2. Browne WR, Feringa BL (2006) Nat Nanotechnol 1:25–35
3. Badjic JD, Balzani V, Credi A, Silvi S, Stoddart JF (2004) Science 303:1845–1849
4. Pace G, Ferri V, Grave C, Elbing M, Zharnikov M, Mayor M, Rampi MA, Samorì P (2007) Proc Natl Acad Sci USA 104:9937–9942
5. Ferri V, Elbing M, Pace G, Dickey MD, Zharnikov M, Samorì P, Mayor M, Rampi MA (2008) Angew Chem Int Ed 47:3407–3409
6. Bléger D, Yu Z, Hecht S (2011) Chem Commun 47:12260–12266
7. Bléger D, Ciesielski A, Samorì P, Hecht S (2010) Chem Eur J 16:14256–14260
8. Gahl C, Schmidt R, Brete D, McNellis ER, Freyer W, Carley R, Reuter K, Weinelt M (2010) J Am Chem Soc 132:1831–1838
9. Mar W, Klein ML (1994) Langmuir 10:188–196
10. Gannon G, Greer JC, Larsson JA, Thompson D (2010) ACS Nano 4:921–932
11. Tirosh E, Benassi E, Pipolo S, Mayor M, Valášek M, Frydman V, Corni S, Cohen SR (2011) Beilstein J Nano 2:834–844
12. Tiberio G, Muccioli L, Berardi R, Zannoni C (2010) ChemPhysChem 11:1018–1028
13. Ciminelli C, Granucci G, Persico M (2008) Chem Phys 349:325–333
14. Cusati T, Granucci G, Persico M (2011) J Am Chem Soc 133:5109–5123
15. Böckmann M, Peter C, Delle Site L, Doltsinis NL, Kremer K, Marx D (2007) J Chem Theory Comput 3:1789–1802
16. Böckmann M, Doltsinis NL, Marx D (2010) Angew Chem Int Ed 49:3382–3384
17. Turanský R, Konôpka M, Doltsinis NL, Stich I, Marx D (2010) Phys Chem Chem Phys 12:13922–13932
18. Böckmann M, Marx D, Peter C, Site LD, Kremer K, Doltsinis NL (2011) Phys Chem Chem Phys 13:7604–7621
19. Schäfer LV, Müller EM, Gaub HE, Grubmüller H (2007) Angew Chem Int Ed 46:2232–2237
20. Jorgensen WL, Tirado-Rivers J (1988) J Am Chem Soc 110:1657–1666
21. Palmo K, Mannfors B, Mirkin N, Krimm S (2003) Biopolymers 68:383–394
22. Maple J, Dinur U, Hagler A (1988) Proc Natl Acad Sci USA 85:5350–5354
23. Dasgupta S, Yamasaki T, Goddard W III (1996) J Chem Phys 104:2898–2920
24. Cacelli I, Prampolini G (2007) J Chem Theory Comput 3:1803–1817
25. Cacelli I, Cimoli A, Lavotto PR, Prampolini G (2012) J Comput Chem 33:1055–1067

26. Moré J (1978) The Levenberg-Marquardt algorithm: implementation and theory. In: Watson GA (ed) Numerical analysis, vol 630. Springer, Berlin, p 105
27. Moré J, Wright S (1993) Optimization software guide, In: Frontiers in applied mathematics, Number 14, SIAM. (ISBN: 978-0-898713-22-0; Link to SIAM title listing); MINPACK-1, available from netlib. http://www.netlib.org/minpack
28. Sinha P, Boesch SE, Gu C, Wheeler RA, Wilson AK (2004) J Phys Chem A 108:9213–9217
29. Becke AD (1993) J Chem Phys 98:5648–5652
30. Frisch MJ, Trucks GW, Schlegel HB et al (2009) *Gaussian 09*, Revision A.02, Inc., Wallingford
31. Cieplak P, Cornell WD, Bayly C, Kollman PA (1995) J Comput Chem 16:1357–1377
32. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) J Comput Chem 25:1157–1174
33. Frenkel D, Smit B (2002) Understanding molecular simulations. Academic Press, San Diego
34. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) J Chem Theory Comput 4:435–447
35. Wang L, Xu W, Yi C, Wang X (2009) J Mol Graph Mod 27:792–796
36. Bouwstra JA, Schouten A, Kroon J (1983) Acta Crystallogr Sect C 39:1121–1123
37. Traetterberg M, Hilmo I, Hagen K (1977) J Mol Struct 39:231–239
38. Cembra A, Bernardi F, Garavelli M, Gagliardi L, Orlandi G (2004) J Am Chem Soc 126:3234–3243
39. Wang L, Wang X (2007) J Mol Struct (Theochem) 847:1–9
40. Wang L, Xu J, Zhou H, Yi C, Xu W (2009) J Photochem Photobiol A Chem 205:104–108
41. Corish J, Morton-Blake DA, O'Donoghue F, Baudour JL, Bénière F, Toudic B (1995) J Mol Struct (Theochem) 358:29–38
42. Wei-Guang Diau E (2004) J Phys Chem A 108:950–956
43. Ishikawa T, Noro T, Shoda T (2001) J Chem Phys 115: 7503–7512
44. Cattaneo P, Persico M (1999) Phys Chem Chem Phys 1:4739–4744
45. Fliegl H, Köhn A, Hättig C, Ahlrichs R (2003) J Am Chem Soc 125:9821–9827
46. Mostad A, Romming C (1971) Acta Chem Scand 25:3561–3568
47. http://www.gromacs.org/Documentation/How-tos/Normal_Mode_Analysis
48. Wilson EB (1934) Phys Rev 45:706–714
49. Palafox MA (2000) Int J Quant Chem 77:661–684
50. Schrerer JR, Overend J (1961) Spectrochim Acta 17:719–730
51. Ilnytskyi JM, Neher D, Saphiannikova M (2011) J Chem Phys 135:044901–044913
52. Duchstein P, Neiss C, Görling A, Zahn D (2012) J Mol Model 18:2479–2482