J Mol Model (2014) 20:2472
DOI 10.1007/s00894-014-2472-5

ORIGINAL PAPER

Conformational analysis of a polyconjugated protein-binding ligand by joint quantum chemistry and polarizable molecular mechanics. Addressing the issues of anisotropy, conjugation, polarization, and multipole transferability

Elodie Goldwaser · Benoit de Courcy · Luc Demange · Christiane Garbay · Françoise Raynaud · Reda Hadj-Slimane · Jean-Philip Piquemal · Nohad Gresh

Received: 10 March 2014 / Accepted: 21 September 2014 / Published online: 1 November 2014

Abstract We investigate the conformational properties of a potent inhibitor of neuropilin-1, a protein involved in cancer processes and macular degeneration. This inhibitor consists of four aromatic/conjugated fragments: a benzimidazole, a methylbenzene, a carboxythiourea, and a benzene-linker dioxane, and these fragments are all linked together by conjugated bonds. The calculations use the SIBFA polarizable molecular mechanics procedure. Prior to docking simulations, it is essential to ensure that variations in the ligand conformational energy upon rotations around its six main-chain torsional bonds are correctly represented (as compared to high-level ab initio quantum chemistry, QC). This is done in two successive calibration stages and one validation stage. In the latter, the minima identified following independent stepwise variations of each of the six main-chain torsion angles are used as starting points for energy minimization of all the torsion angles simultaneously. Single-point QC calculations of the minimized structures are then done to compare their relative energies ΔE_{conf} to the SIBFA ones. We compare three different methods of deriving the multipoles and polarizabilities of the central, most critical moiety of the inhibitor: carboxythiourea (CTU). The representation that gives the best agreement with QC is the one that includes the effects of the mutual polarization energy E_{pol} between the amide and thioamide moieties. This again highlights the critical role of this contribution. The implications and perspectives of these findings are discussed.

Keywords Polarizable force fields · Multipoles · Conjugation

Electronic supplementary material The online version of this article (doi:10.1007/s00894-014-2472-5) contains supplementary material, which is available to authorized users.

E. Goldwaser · L. Demange · C. Garbay · F. Raynaud · N. Gresh (✉)
Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques, UMR 8601, UFR Biomédicale, Université ParisDescartes, 45, rue des Saints-Pères, 75006 Paris, France
e-mail: nohad.gresh@parisdescartes.fr

E. Goldwaser · B. de Courcy · J.-P. Piquemal (✉)
Laboratoire de Chimie Théorique, Sorbonne Universités, UPMC, Paris 6, case courrier 137, 4, place Jussieu, F75252 Paris, France
e-mail: jpp@lct.jussieu.fr

R. Hadj-Slimane
Tragex Pharma, 29, rue Marcel Dassault, 92100 Boulogne Billancourt, France

Present Address:
L. Demange
Institut de Chimie de Nice (ICN), UMR 7272 CNRS, Université de Nice Sophia-Antipolis, Parc Valrose, 06000 Nice, France

Introduction

Our laboratories are involved in the design of inhibitors of the neuropilin-1 (NRP-1) protein. NRP-1 is a co-receptor of vascular endothelial growth factor (VEGF), which promotes angiogenesis and is involved in several diseases such as cancer and macular denegeracy [1, 2], so inhibitors of NRP-1 could block the development of those pathologies. Virtual screening studies have recently identified a potent inhibitor of NRP-1 which was

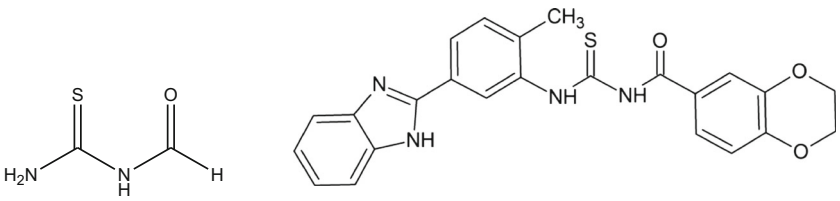
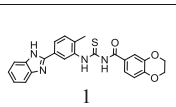
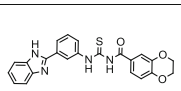
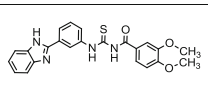
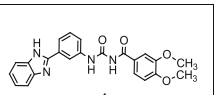


Fig. 1a–b Structures of **a** the carboxythiourea fragment and **b** lig-47

denoted 'lig-47' and shown to possess submicromolar activity on several cancer cell lines. Its structure (Fig. 1) contains the following chemical groups: a benzimidazole, a methylbenzene, a carboxythiourea, and a benzene-substituted dioxane. Lig-47 is a highly conjugated drug, and the connections between its four groups all involve unsaturated atoms with *sp* or sp^2 hybridization. In the present work, we investigate the conformational features of lig-47 in detail, as knowledge of these features is required in order to understand its complexation to NRP-1.

Unraveling the structural and energetic factors involved in NRP-1–lig-47 complexation is a necessary step in the design of new derivatives with enhanced affinities. However, with the exception of complexes of NRP-1 with the naturally occurring peptide tuftsin [3] and small peptide-like ligands [4], no high-resolution structural studies of complexes of NRP-1 with inhibitors or ligands are available. Two such NRP-1 complexes could be used as candidate starting points for lig-47 docking. However, there is a major issue relating to the conformational flexibility of the drug. On the one hand, it could adopt an extended conformation, similar to tuftsin; alternatively, it could adopt a compact conformation, as found from initial studies performed with the SURFLEX software [5] (Borriello et al. is Cancer Letts., submitted). Whereas molecular dynamics (MD) and energy minimization (EM) are likely to identify alternative poses with intermediate drug conformations, it is necessary to be able to correctly rank the energies of all candidate docking poses prior to enhancing the binding affinity by making targeted structural changes. Carboxythiourea (denoted "CTU" hereafter) is a critical building block of lig-47. Determinations of the inhibition of VEGF-A_{165}–NRP-1 binding were recently performed in our laboratories; these studies focused on lig-47 and several structurally related molecules. The inhibition percentages at 10 μM for three representative derivatives are reported in Fig. 2. It is clear that structural changes to the methylbenzene or the benzodioxane ring of lig-47 are not detrimental to its biological activity (24 % and 32 %, as compared to 32 % for lig-47). In marked contrast, substitution of the thiourea substituent by urea (fourth compound) results in complete loss of the antagonist effect. Key features in the affinity of lig-47 for its target are the distribution of electrons over its CTU-centered conjugated backbone and how this impacts its conformational flexibility. We address the latter issue, which must be properly accounted for, in this paper. Overestimating the confomational flexibility would result in an excessively "floppy" ligand and the onset of a collection of unlikely candidate-protein-binding poses. Conversely, underestimating it would give rise to an unrealistically stiff ligand that is unlikely to favorably bind its receptor. The most reliable procedure for computing conformational energy variations is quantum chemistry (QC), but QC calculations are too costly to use to perform an exhaustive exploration of the seven-dimensional energy surface of the drug, especially when it is complexed with its receptor or in solution prior to complexation.

Fig. 2 Sructural analogs of lig-47 and their VEGF-A_{165}–NRP-1 binding inhibition percentages

	 1	 2	 3	 4
VEGF-A165/NRP 1 binding inhibition at 10μM	33%	24%	32%	0.00%

In our laboratories, we are involved in the development of an accurate, QC-grounded molecular mechanics/dynamics potential denoted SIBFA (sum of interactions between fragment ab initio computed), which is used to calculate inter- and intramolecular interactions [6, 7]and has been applied to ligand–protein complexes (reviewed in [8–10]). The vast majority of the conformation changes dealt with so far in this context have involved rotations around single bonds, for which a single threefold torsion barrier with amplitudes in the 0–2.2 kcal/mol range could suffice. By contrast, all of the torsion angles in lig-47 are associated with conjugated bonds. It is thus clear that a threefold (n=3, see Eq. 2 below) barrier is inadequate and that barriers with n=1 and n=2 should be calibrated instead. These barriers alone could ensure coplanarity of the two connected fragments in order to account for conjugation. In their absence, steric repulsions between the connected fragments would result in a mutually perpendicular arrangement. However, a reliable conformation study of lig-47 involves more than just calibrating torsion barriers. This is because atoms in all four of the building blocks of lig-47 can move closer together during EM or MD as a result of simultaneous, and possibly concerted, changes in the conformation of its seven rotatable bonds. In some cases, such proximities may be imposed because the interacting fragments are connected by chemical bonds and are not free to relax fully, unlike in "true" intermolecular interactions between free fragments. This also implies that the intramolecular interfragment interactions must be accurately calibrated and validated. Such calibration should focus on the set of atoms (C, N, O, and S) that belong to the constitutive fragments of lig-47, and should be done for a range of interaction distances, including distances that are shorter than the equilibrium distance (which can occur during EM or MD).

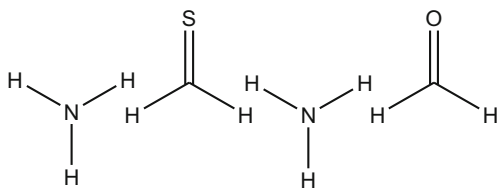


Fig. 3 Diagram showing how carboxythiourea is split into four subfragments

Therefore, the present investigation is carried out in three steps. The CTU fragment ("a") is first represented by considering the fragment in its entirety.

In an initial step, we will utilize electron localization (ELF) topological analysis [11–13] to study the electron density distribution of carboxythiourea, because it is the central, as well as the most polar, moiety of lig-47. One critical issue is the locations of the lone pairs of its thiocarbonyl and carbonyl bonds, since it is essential to correctly represent lone pairs when using SIBFA, as noted in our previous works [6–10]. This is due to their involvement in the short-range exchange-repulsion and charge-transfer contributions, as well as in the lone-pair polarization contribution. Differences from isolated formamide and thioamide entities can be expected, since these entities are connected by a conjugated bond in lig-47.

The ELF analysis of CTU will be done on the fully extended conformation that is stabilized by conjugation. The ELF-derived internal coordinates of the lone-pair centroids will be used as input data for SIBFA. In a second step, we will probe the most important sites on CTU by simulating the approach of a water molecule to the site, which occurs either in-plane or perpendicular to the plane. Such analyses will then be performed for benzimidazole, methylbenzene, and benzene-substituted dioxane.

A proton from water always approaches in the direction of an sp^2 or sp^3 lone pair and, for aromatic and conjugated C atoms and the N atoms of benzimidazole and CTU, along the direction of a π lone pair. The water probe approaches along the NH bonds of CTU. ΔE (QC) will be decomposed into its separate Coulomb,

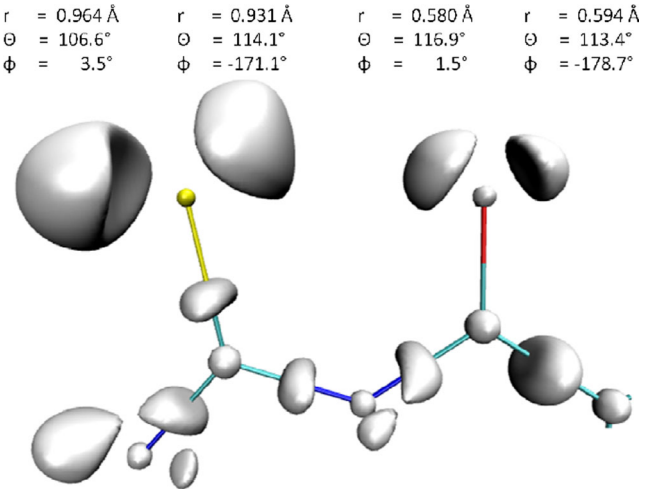


Fig. 4 Representation of the ELF contours on CTU in its extended conformation

short-range repulsion, polarization, and charge-transfer contributions.

We will accordingly readjust the effective van der Waals (vdW) radii used in SIBFA for the respective energy contribution and their possible increments along the angular directions corresponding to the orientations of the individual lone pairs. This is possible because the input includes the internal coordinates and vdW increment (default value is null) of each individual lone pair.

As mentioned below, with the exception of the π lone pairs of sulfur, the vdW increments are readjusted by limited amounts. This should not impair transferability because such increments can be considered to be fragment-specific.

Following this calibration, we will perform stepwise (15°) variations in torsion angle around each of the rotatable bonds, first for CTU and then for the entire lig-47, and calculate the resulting variations in QC conformational energy. For each of the seven angles, we will then calibrate the amplitude of the torsional potential V_0 for both the one- and the two-fold barriers

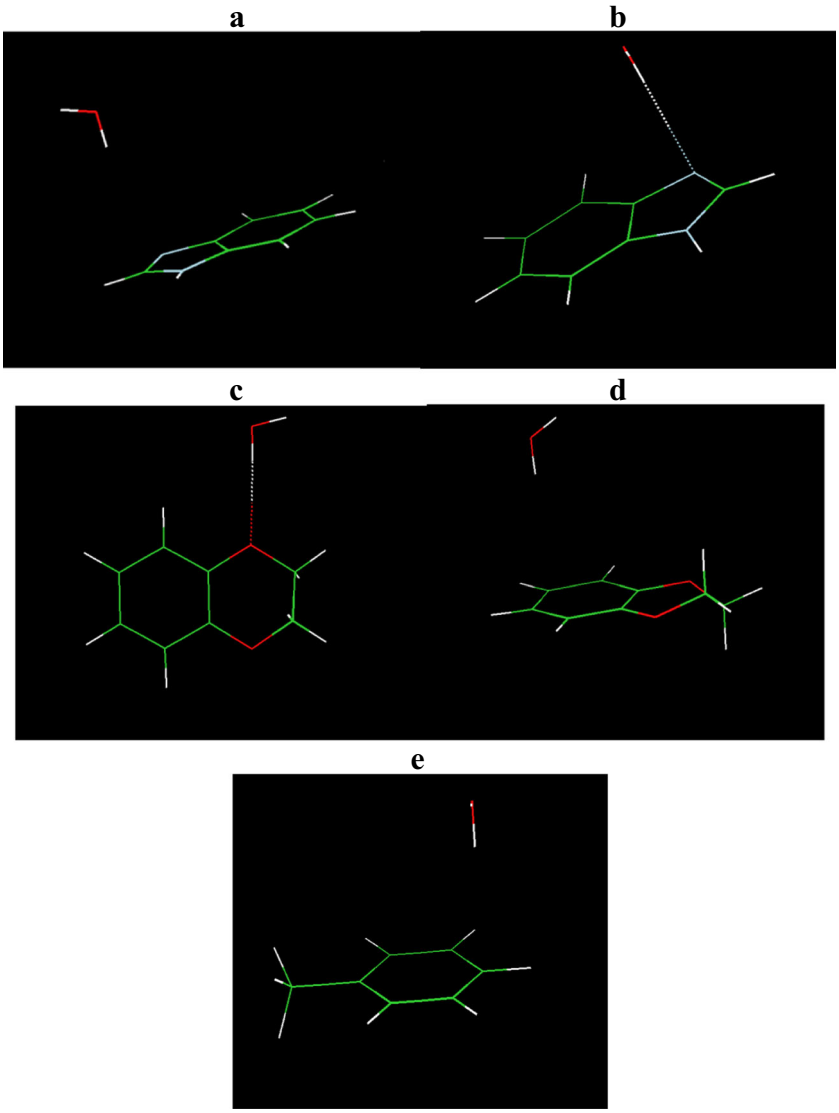


Fig. 5a–n Studied directions of approach of water to the sites of lig-47. **a** Approach perpendicular to the H-linked C atom of the pentacyclic ring of benzimidazole. **b** Approach perpendicular to the purine-like nitrogen of the pentacycle of benzimidazole. **c** Approach via oxygen 11 of the benzodioxane. **d** Approach perpendicular to carbon 2 of the benzodioxane. **e** Approach perpendicular to carbon 2 of methylbenzene. **f** Approach via the outer lone pair of the sulfur atom of CTU. **g** Approach via the inner lone pair of the sulfur atom of CTU. **h** Approach via the inner lone pair of the oxygen atom of CTU. **i** Approach via the outer lone pair of the oxygen atom of CTU. **j** Approach via the π lone pair of the sulfur atom of CTU. **k** Approach via the π lone pair of the oxygen atom of CTU. **l** Approach via the π lone pair of the primary nitrogen atom of CTU. **m** Approach via the π lone pair of the secondary nitrogen atom of CTU. **n** Approach via the hydrogen on the secondary nitrogen of CTU

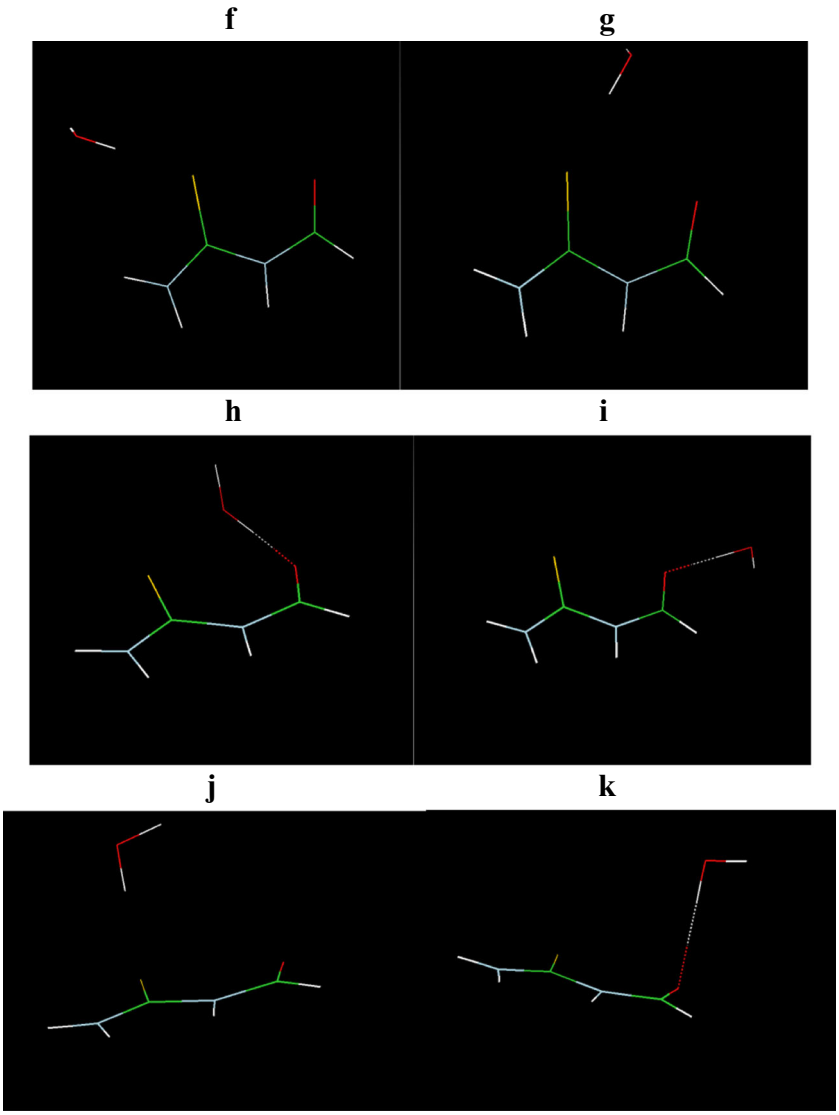


Fig. 5 (continued)

in SIBFA to ensure a good match with the corresponding QC conformational curves.

In the fourth step, we will evaluate the transferability of the approach. To achieve this, we will select minima of the six curves and use them as starting points to perform SIBFA energy minimization. We will then rank the minimized conformations in terms of their relative stabilities and compare the rankings to the corresponding rankings obtained from single-point QC calculations. The influence of the polarization contribution on the energy separations between the minimized conformations will be assessed, and other, more comprehensive, methods of incorporating this contribution will be evaluated.

Finally, we will assess the impact of the correlation/dispersion on the conformations of lig-47 calculated at the QC level, and the extent to which we can account for this using the E_{disp} contribution in SIBFA.

Procedure

Quantum chemistry

The conformational energy variations are obtained via the cc-pvtz(−f) basis set [14, 15] using the G09 software [16]. Analyses of the interaction energies of the water probe with the sites of lig-47 are done with the reduced variational space analysis (RVS) procedure [17], as implemented in GAMESS [18]. The electron distribution is analyzed using the ELF procedure [11] with the TopMod package [12].

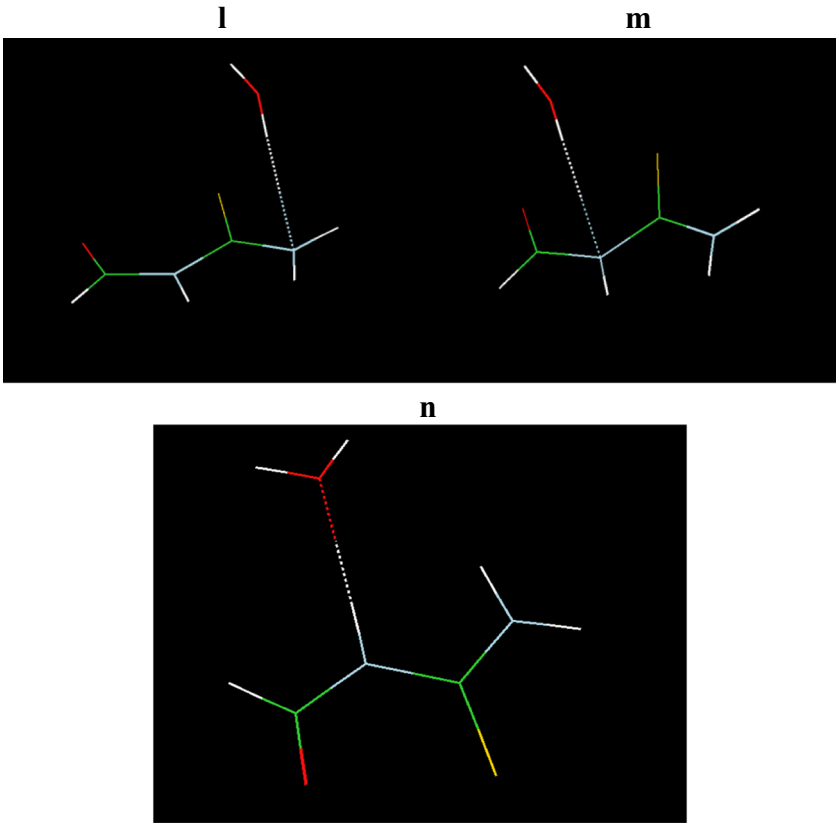


Fig. 5 (continued)

Further details about the applications of ELF to biological systems can be found in [12].

SIBFA

In the SIBFA procedure [6–10], the conformation energy in a flexible molecule is computed as the sum of the intermolecular interactions ΔE between the fragments that comprise the molecule. ΔE is computed as a sum of five contributions:

$$\Delta E = E_{\mathrm{MTP}*} + E_{\mathrm{rep}} + E_{\mathrm{pol}} + E_{\mathrm{ct}} + E_{\mathrm{disp}}, \quad (1)$$

which are, respectively, the electrostatic multipolar, the short-range repulsion, the polarization, the charge-transfer, and the dispersion contributions. $E_{\mathrm{MTP}*}$ is computed with distributed multipoles derived from the ccpvtz(−f) QC molecular orbitals (MOs) of the individual fragments of lig-47 using Stone's analysis [19, 20], and is augmented with a short-range penetration correction [21]. E_{rep} is the sum of the bond–bond, bond–lone pair, and lone pair–lone pair interactions. E_{pol} is computed using distributed anisotropic polarizability tensors located on the chemical bonds and the localized lone-pair hybrids of the fragments, as derived from the procedure reported by Garmer and Stevens [22]. The polarizing field is calculated with the same multipoles as for $E_{\mathrm{MTP}*}$ and is screened by a Gaussian function. E_{ct} is a short-range charge-transfer contribution and E_{disp} is the dispersion contribution. A periodic torsional energy contribution is included via the expression

$$V = \frac{V_{0(1)}}{2}(\cos\Phi + 1) + \frac{V_{0(2)}}{2}(\cos 2\Phi + 1), \quad (2)$$

where φ is the torsion angle. The two terms on the right-hand side express the periodicity, which corresponds to $n=1$ and 2, respectively, for the conjugated bonds.

Lig-47 is constructed from four distinct fragments: benzimidazole, methylbenzene, CTU, and benzene-linked dioxane. It should be recalled that, in the context of SIBFA [6–10], the multipoles of the two junction bonds X–H and H–Y that make up an X–Y

bond are redistributed on atoms X and Y and at the midpoint of the created bond X–Y. E_{MTP*} is not computed between the atoms and bonds belonging to any connecting bond and the two fragments it connects. To allow for rotations around its successive N–C, C–N, and N–C bonds, CTU is split into four fragments (Fig. 3). The first and third are sp^2 amines, the second is thioaldehyde, and the fourth is aldehyde. In this process, fictitious connecting H atoms are created and given null multipoles, and each of the two connecting bonds that make up a junction is given half the multipole and polarizability of the original bond prior to splitting. Thus, the multipoles of the N–H and H–C bonds of the first junction have half the values of the initial N–C bond and are located at its middle, etc. CTU therefore retains its net initial charge of 0. An important point relates to E_{pol} and E_{ct}: these contributions are calculated between all polarizability centers of the interacting fragments, but they are not calculated *between* the four split subfragments that make up CTU. This is because mutual polarization of these subfragments is considered to have initially taken place during the SCF iterations, which resulted in the MOs from which the multipoles were derived. Furthermore, such mutual polarizations would take place between subfragments with non-net charges (0, −1, +1, etc.), which was observed by us to greatly amplify the magnitude and weight of E_{pol}. In addition, in the general case, the multipoles used for E_{pol} and the polarizabilities of the junctions are not redistributed but carried back on the X and Y atoms of X–Y bonds, and the X–H and Y–H bonds are collapsed onto X and Y, respectively. This enables each fragment to retain its original net charge and also prevents two connecting fragments from polarizing one another at distances that would be shorter than the length of the actual X–Y bond. This overall procedure has been found to be effective when compared with QC calculations focusing on ten conformers of Ala tetrapeptides [23] that were designed to benchmark the accuracy of molecular mechanics force fields [24]. A similar treatment is applied to methylbenzene, which is split into a benzene-like and a methane-like fragment.

Energy minimization of selected conformers was done with the Merlin [25] minimizer.

Results and discussion

ELF analysis of CTU

Figure 4 highlights the ELF function distribution around the sp^2 lone pairs of CTU. The internal coordinates that provide the positions of their attractors (analogous to the centroids in the localized MO picture) with respect to the atoms that bear them are listed in this figure.

Table 1 Values (distances in Å, angles in degrees) of the vdW parameters for the H, C, N, O, and S atoms and of the internal coordinates and vdW increments/decrements of the inner sp^2, outer sp^2, and π lone pairs

Atom	VdW (repulsion)	Atom	VdW (penetration)	
C (benzene)	1.82	C (benzene)	1.6	
O	1.40	O	1.41	
O (ether)	1.448	O (ether)	1.1	
N (sp^2)	1.48	N (sp^2)	1.61	
S (CTU)	1.85	S (CTU)	1.6	
O (carbonyl)	1.41	O (carbonyl)	1.44	
H	1.24	H	1.1	
H (amide)	1.30	H (amide)	1.3	
N (sp^3)	1.72	N (sp^3)	1.45	
	θ	φ	r	Increment
CTU O (inner)	116.9	0.00	0.59	−0.03
CTU O (outer)	113.4	180.00	0.58	−0.03
CTU O (π)	90.00	90.00	0.5	0.00
CTU O (π)	90.00	90.00	0.5	0.00
CTU S (outer)	106.6	0	0.93	0.08
CTU S (inner)	114.1	180.00	0.93	−0.05
CTU S (π)	90.00	90.00	0.5	0.60
CTU S (π)	90.00	90.00	0.5	0.60
CTU N_{prim} (π)	90.00	90.00	0.25	−0.15
CTU N_{prim} (π)	90	−90	0.25	−0.15
CTU N_{sec} (π)	90.00	90.00	0.25	−0.20
CTU N_{sec} (π)	90	−90	0.25	−0.20

Probing lig-47 and CTU using one water molecule

The different approach directions of the water probe are illustrated in Fig. 5. Water approaches via a proton perpendicular to a C atom of benzimidazole, methylbenzene, or benzene-linked dioxane; via the N atom of benzimidazole; or via any of the N, O, or S atoms of CTU. It approaches in-plane through its H atom to the external bisector of one ether O of benzene-linked dioxane and along two possible directions to the C=S and C=O bonds of CTU. Upon binding to the "inner" S or O sp^2 lone pair, the angle θ (C=X–Hw) was set to 150°. It was set at 120° upon approaching the "outer" lone pairs. A larger value of θ for the inner approach to one C=X bond allows the steric repulsion between water and the other bond to be reduced. Finally, water approaches in-plane along the external bisector of its O atom to the central amide H of CTU.

Table 2 Values (kcal/mol) of the intermolecular interaction energies at the equilibrium distance and their contributions to the binding of the probe water molecules with the lig-47 fragments

Benzimidazole C7 d=3.4 Å	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}
RVS	−0.2	0.1	−0.1	0	−0.2
SIBFA	−0.2	0.1	−0.1	0	−0.2
Benzimidazole N6 d=2.7 Å	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}
RVS	−1.4	1	−0.4	−0.1	−0.9
SIBFA	−1.4	0.9	−0.1	−0.1	−0.7
Dioxane O11 d=2.2 Å	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}
RVS	−3.9	2.7	−0.6	−0.4	−2.2
SIBFA	−4.2	2.6	−0.4	−0.2	−2.2
Dioxane C2 d=2.8 Å	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}
RVS	−1.9	1.1	−0.3	−0.2	−1.3
SIBFA	−1.5	1.1	−0.2	0	−0.6
Methylbenzene C2 d=2.8 Å	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}
RVS	−1.8	1.1	−0.3	−0.2	−1.2
SIBFA	−1.6	1.1	−0.2	0	−0.7

CTU	O (inner) d=2.9 Å	O (outer) d=2.1 Å	O (π) d=2.6 Å	S (outer) d=2.7 Å	S (inner) d=3.0 Å	S (π) d=2.7 Å
E_C	−3.9	−6.7	−2.1	−5.2	−1.8	−4.4
E_{mtp}	−4.2	−6.1	−2	−4.9	−2	−4.2
E_{exc}	2.2	4	0.9	3.1	0.6	2.7
E_{rep}	2.2	3.8	0.8	2.8	0.6	2.6
E_{pol}	−0.4	−1	−0.4	−0.6	−0.3	−0.6
E_{pol}	−0.4	−0.9	−0.2	−0.7	−0.3	−0.3
E_{ct}	−0.3	−0.5	−0.2	−0.5	−0.1	−0.6
E_{ct}	−0.2	−0.3	−0.1	−0.3	−0.1	−0.2
ΔE_{RVS}	−2.4	−4.3	−1.8	−3.3	−1.6	−2.9
E_{tot}	−2.6	−3.5	−1.5	−3.1	−1.8	−2.1

CTU N_{prim}	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}
RVS	−1.3	0.7	−0.3	−0.1	−1
SIBFA	−1.3	0.8	−0.2	−0.1	−0.8
CTU N_{sec}	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}
RVS	−0.5	0.1	−0.1	0	−0.6
SIBFA	−0.5	0.2	−0.1	0	−0.4
CTU NH	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}
RVS	−5.8	2.5	−0.8	−0.3	−4.6
SIBFA	−5.9	2.2	−1	−0.4	−5.1

For the SIBFA calculations, the centroids of the sp^2 lone pairs were located at positions consistent with the ELF results. The vdW radii for all atoms and contributions are listed in Table 1, as are the internal coordinates of the lone pairs and increments. As mentioned above, we have calibrated the increments/decrements of the S and O vdW radii along each individual lone pair so that E_{rep} (SIBFA) matches the radial behavior of $E_{exch\text{-}rep}$ (RVS). This was done by decreasing the vdW radius of the inner sp^2 lone pair of S by 0.05 Å while increasing that of the outer lone pair by 0.08 Å, yielding a difference of 0.13 Å. For O, both lone pairs underwent similar decreases in their vdW

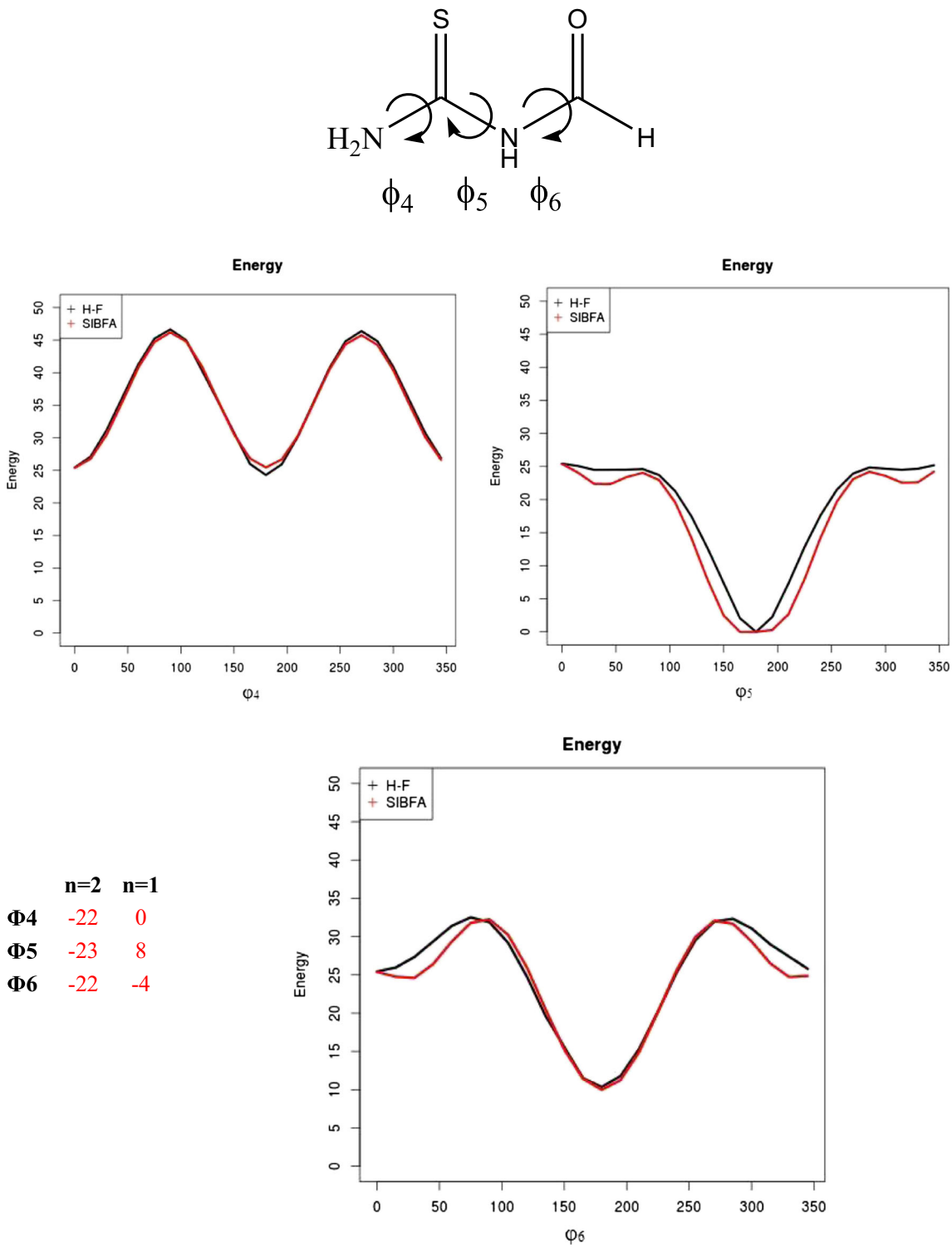


Fig. 6 Variations in the conformational energy of CTU as functions of the torsion angles $\varphi 4$, $\varphi 5$, and $\varphi 6$

radii, amounting to 0.03 Å. To correctly account for the magnitudes of E_{rep} and E_{ct} in the vertical complexes of water with the C=S bond, large increments (0.60 Å) were found to be necessary for the vdW radii of the π lone pairs on S, while no increments were necessary for the corresponding complex with the C=O bond.

Table 2 lists the values at the equilibrium distance of the intermolecular interaction energies and their contributions to the binding of probe water molecules to the following non-CTU sites: benzimidazo C_7 and N_6; dioxane O and C; methylbenzene C. It also lists the values for the CTU sites, i.e., the carbonyl O for inner

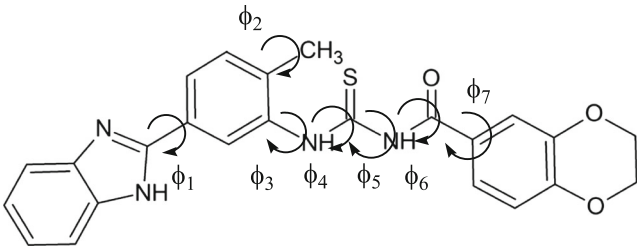


Fig. 7 Diagram of the torsion angles of lig-47

and outer in-plane approaches by the water (on the same side and *trans* to the C=S bond, respectively); the carbonyl O for a perpendicular approach; the thiocarbonyl S for inner, outer, and perpendicular approaches; the first and second amide nitrogens for a perpendicular approach; and the amide H for an in-plane approach. The results for the binding of water to CTU at a range of distances shorter than the equilibrium distance are given in Table S1 of the "Electronic supplementary material," ESM. Table S1 shows that the overall agreement between the SIBFA and RVS values is satisfactory, but needs to be validated at the level of each contribution.

Monitoring the conformation behavior of CTU and lig-47

Having calibrated the vdW radii on the basis of the QC energy decomposition analyses for intermolecular interactions, we then calibrated the amplitudes of the torsional barriers to intramolecular interactions. This was first done on the sole CTU moiety by varying the three torsion angles shown in Fig. 6 in a stepwise manner such that the variations in the sum of the SIBFA energy without dispersion plus the torsional contribution match those obtained from QC calculations done at the HF level. For consistency with the subsequent lig-47 conformation study, we denote the torsion angles around the N1–C2, C2–N2, and N2–C3 bonds (the atoms are numbered in succession from left to right) as $\varphi 4$, $\varphi 5$, and $\varphi 6$, respectively. The φ angles are numbered in ascending order along the main chain. The starting conformation is the fully extended one, so the φ values are 180°. The fitted V_0 values corresponding to $n=1$ and 2 are also given in Fig. 6. The SIBFA curves are seen to closely match the QC ones, with deep minima corresponding to the extended conformations.

We then considered lig-47 in its entirety. The labels for the eight torsional angles and the fitted V_0 values are given in Fig. 7. $\varphi 1$ denotes the torsion angle around the bond separating the benzimidazole and methylbenzene rings. No rotations are considered around $\varphi 2$ for the methyl substituent because of free rotation. The starting value of $\varphi 3$ was 240° instead of 180° because steric repulsion occurs between the sulfur and the benzene H atom *ortho* to both junctions connecting benzene with benzimidazole and CTU when $\varphi 3$ is 180°. Retaining a $\varphi 3$ value of 180° when performing all eight stepwise variations would shift all of the curves other than that for $\varphi 3$ to higher energies—a bias that could mask the possible occurrence of energy-relevant conformations in such curves. The curves (see Fig. 8) are drawn assuming that the energy of the lowest-lying conformer, which occurred with $\varphi 5=180°$ (curve 4), has an energy of zero. There is a good match between the QC and SIBFA curves in all three plots, and all six curves are consistent with each other. The most conspicuous deviations between the QC and SIBFA curves, corresponding to errors of >5 kcal/mol, occur at the rightmost parts of the $\varphi 3$ and $\varphi 5$ curves and at the leftmost part of the $\varphi 6$ one, but these relate to unstable conformations for which δE_{conf} >20 kcal/mol. On the other hand, it is important to note that the minima in the QC and SIBFA curves for $\varphi 1$–3 and $\varphi 5$–6 are shifted by similar amounts with respect to the global minimum of curve $\varphi 4$.

Energy minimization of the lowest-lying conformers

The previous curves were drawn by varying each torsion angle individually while the other angles were maintained at values corresponding to an extended conformation. This obviously prevents lig-47 from folding and thus attractive intramolecular interactions from occurring between well-separated atoms in lig-47. It was therefore critical to evaluate if the good match between the SIBFA and QC curves still holds when all of the torsion angles are relaxed simultaneously namely, upon performing energy minimization. Note that free rotation around the methyl group should not impact the main-chain conformations. EM was done starting from selected minima in the conformational energy curves. To make the evaluation more comprehensive, we also included energy-minimized conformers that were obtained at an earlier stage of this study, with a less refined set of vdW and V_0 values.

The three-dimensional structures resulting from minimization of the four most relevant conformers are shown in Fig. 9a–d. The values of the torsion angles of the 15 conformers are reported in Table S2 of the ESM. Their conformational energies and their

	n= 2	n= 1
Φ1	-10	-1
Φ3	-17	-5
Φ4	-28	-5
Φ5	-23	5
Φ6	-24	5
Φ7	-19	0

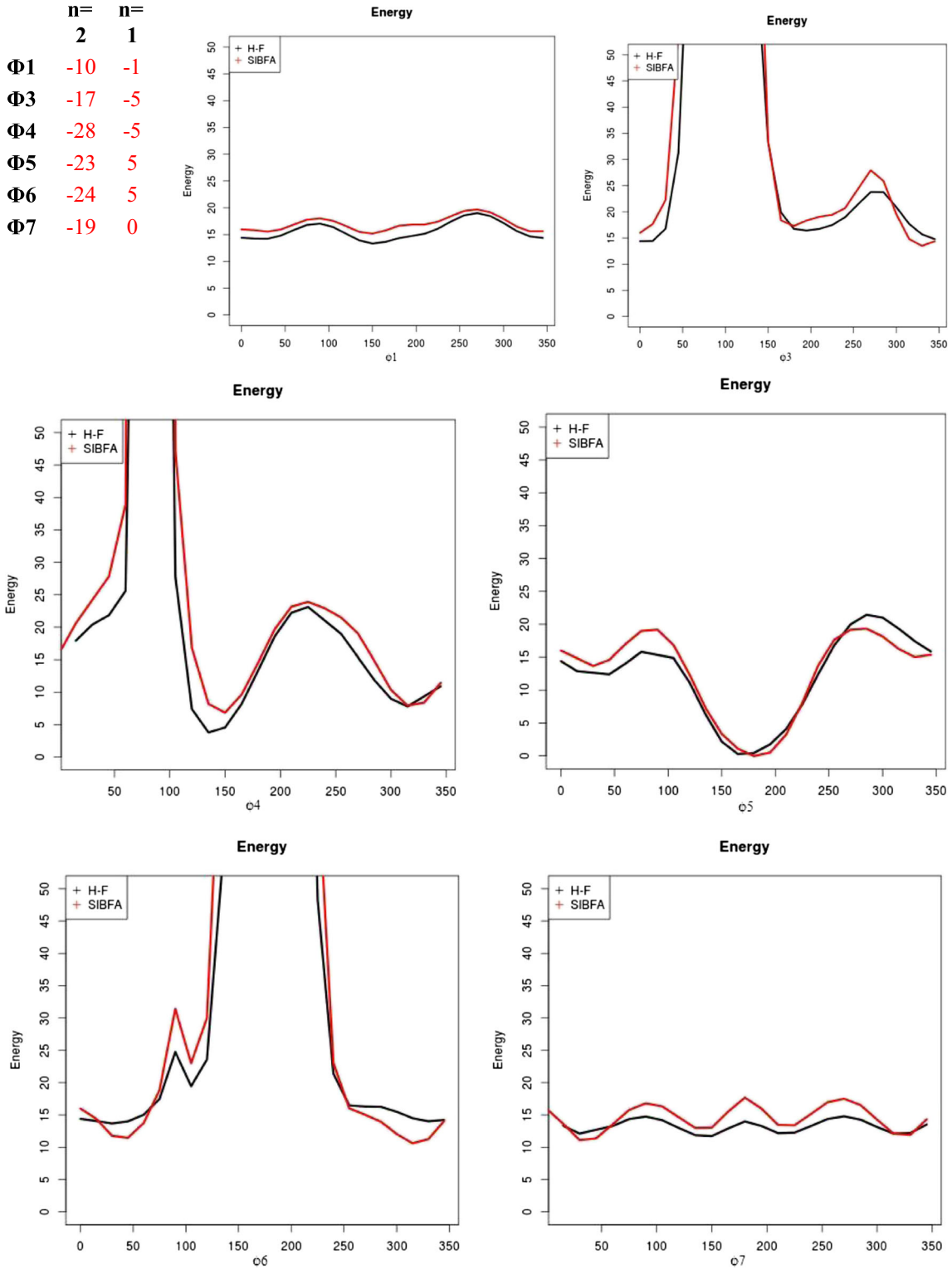
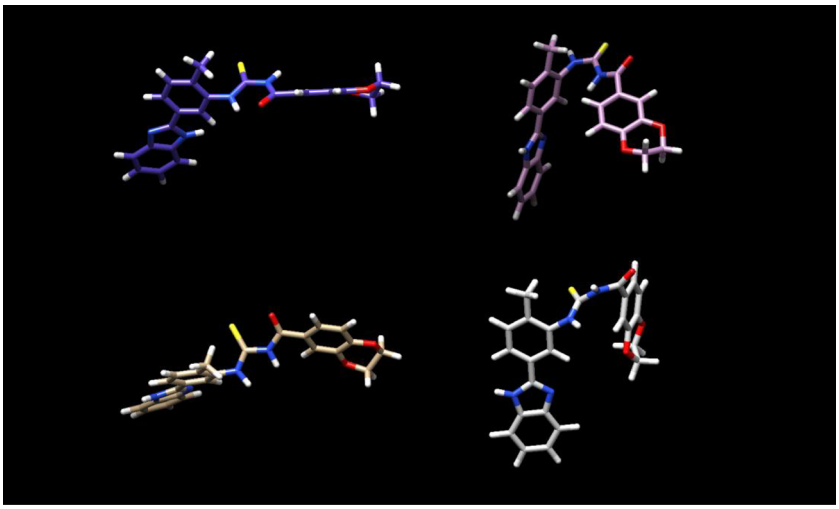


Fig. 8 Variations in the conformational energy of lig-47 as functions of the torsion angles $\varphi 1$ and $\varphi 3$–$\varphi 7$

contributions are listed in Table 3, where the energy of the most stable conformer was taken to be zero. Single-point QC calculations were performed to allow for a comparison of the relative SIBFA and QC



1b (topleft), 2a (topright), 0a (below left), 4a (below right)

Fig. 9a–d Three-dimensional structures of the four most relevant conformations (**a**–**d**) of lig-47

conformational energy. Two related conformers, A and A′, were found to be the most stable in both the SIBFA and QC calculations, while all the remaining conformers have comparable relative conformational energies, whichever method (SIBFA or QC) is employed. However, A and A′ differ in energy from the other conformers by a much smaller amount in SIBFA than in QC (12 kcal/mol as compared to 20).

Table 3 Relative QC and SIBFA conformational energies (kcal/mol) of all conformers along with the various energies that contribute to the SIBFA energy. 0a and 0b are the conformations resulting from the energy minimization of the initial conformation used for the curve. The other conformations result from energy minimizations of the structures corresponding to the minima of the conformational curves. 1a and 1b result from energy minimization of the minimum of the $\varphi 1$ curve and correspond to $\varphi 1 = 150°$. 2a and 2b correspond to an initial angle $\varphi 3$ of 180°. 3a and 3b correspond to an initial angle $\varphi 4$ of 135°. For 3c, $\varphi 4 = 315°$. For 4a and 4b, $\varphi 5 = 165°$. For 5a, $\varphi 6 = 260°$. For 6a, $\varphi 7 = 30°$. For 6b and 6c, $\varphi 7 = 150°$

Conformer	HF	$E_{1order}+E_{2order}+E_{tor}$	E_{mtp}	E_{rep}	E_{1order}	E_{pol}	E_{ct}	$E_{1order}+E_{pol}+E_{ct}$	E_{tor}
0a	20.97	12.40	9.70	4.1	13.80	−3.40	0.00	10.30	2.10
0b	21.22	13.30	12.00	0.30	12.30	−4.40	0.10	7.80	5.50
1a	0.17	2.10	−2.20	8.20	6.00	0.40	0.00	6.30	−4.20
1b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2a	17.80	15.30	7.80	−40.50	−32.70	−5.90	−2.40	6.50	8.80
2b	19.49	11.50	11.20	−1.40	9.80	−3.60	0.00	6.10	5.40
3a	18.42	11.20	7.80	−40.50	−32.70	−8.10	−2.60	2.70	8.50
3b	18.53	11.50	7.80	4.40	12.20	−8.00	−2.60	2.80	8.70
3c	20.97	12.40	9.70	−38.60	−28.90	−3.40	0.00	10.30	2.10
4a	22.39	17.20	4.90	−43.40	−38.50	−1.30	0.30	1.70	15.50
4b	21.36	17.50	4.90	−7.30	−2.40	−2.30	0.10	−4.80	22.30
5a	21.93	15.70	3.90	−44.40	−40.50	−1.40	0.30	0.60	15.10
6a	16.78	11.20	7.30	−41.00	−33.70	−6.80	−2.40	2.50	8.70
6b	20.12	13.30	9.30	3.20	12.50	−1.90	0.00	10.50	2.80
6c	19.22	13.00	11.00	−0.30	10.70	−1.90	0.10	8.70	4.30

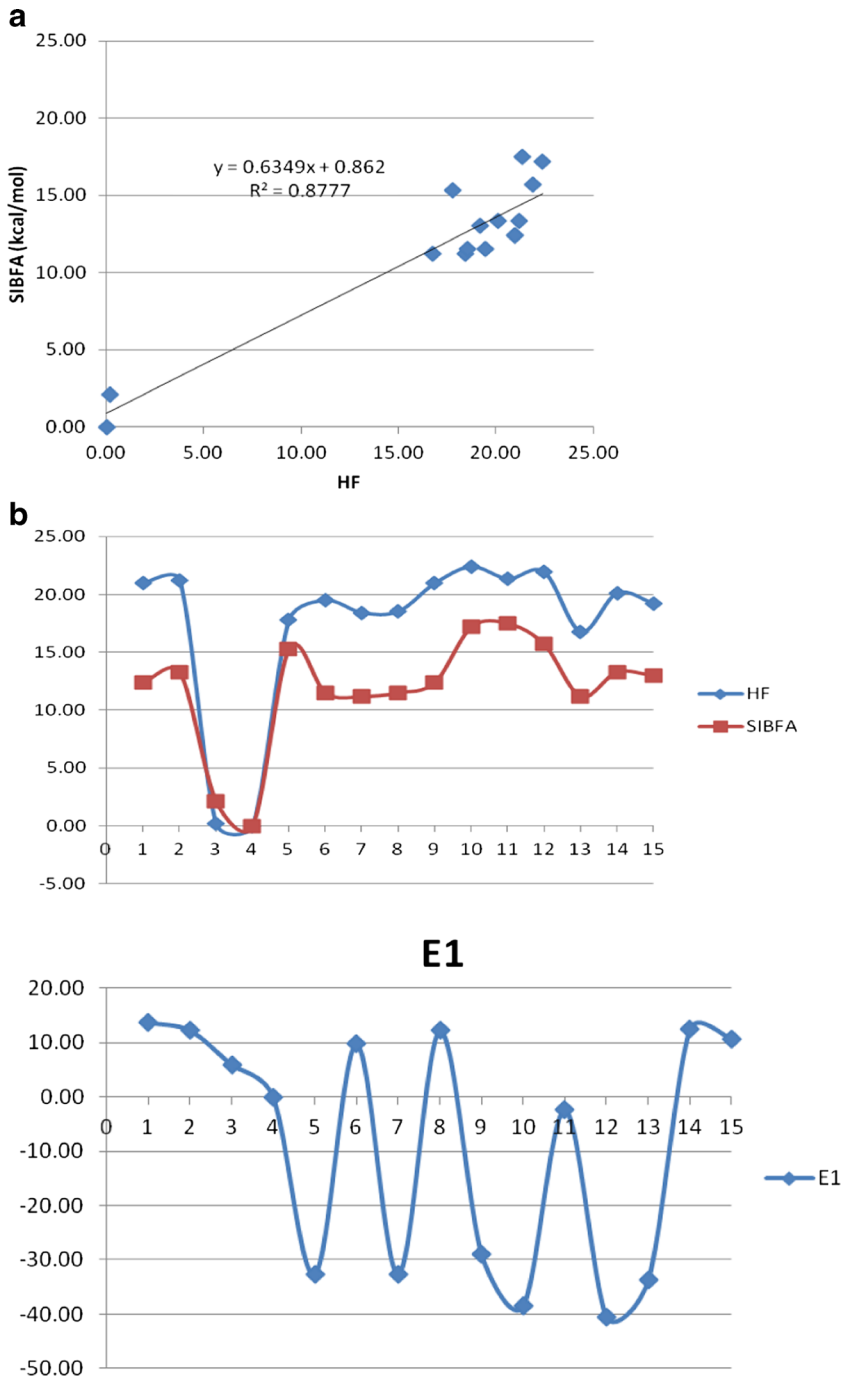


Fig. 10 **a** Regression curve illustrating how closely the curves for δE_{conf} (SIBFA) and δE_{conf} (HF) match. **b** Comparison of the evolutions of δE_{conf} (QC), δE_{conf} (SIBFA), and the components of δE_{conf} (SIBFA) as functions of the conformer

The match between δE_{conf} (SIBFA) and QC is illustrated by the regression curve shown in Fig. 10a, which has a slope of 0.63 and yields a correlation of r^2=0.88. The evolution of δE_{conf} when either SIBFA or QC is used, as well as the evolutions of some SIBFA contributions, are shown as functions of the conformer considered in Fig. 10b. In Fig. 10, as well as in Fig. 12, the fifteen conformers are connected by a

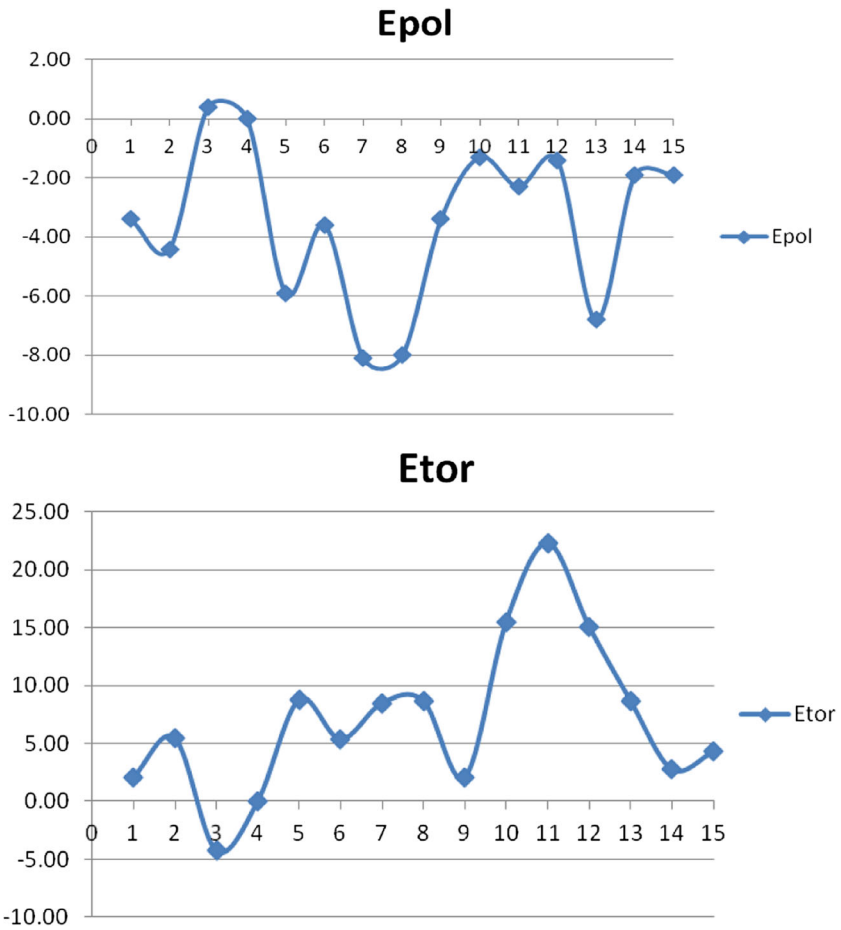


Fig. 10 (continued)

curve in order to guide the eye. The underestimation of δE_{conf} by SIBFA could have adverse consequences upon docking lig-47 to its target protein. Indeed, the high-lying conformers can be stabilized more by favorable intermolecular interactions with the receptor than by the energy minimum when the differences between conformers are in the range 12–20 kcal/mol. In this case, the high-lying conformers are erroneously predicted to bind more favorably than conformers related to the lowest-lying ones. Imbalanced inter- and intramolecular interactions is a caveat common to all docking procedures [26–32].

It should be noted that the two energy minima are the only ones that are stabilized by an intramolecular H-bond between the amide H of CTU connected to methylbenzene and its carbonyl oxygen. This stabilization could be underestimated since the mutual polarization of these two groups is prevented within CTU. We must therefore consider alternative methods of constructing CTU. One involves assembling CTU from thioamide and amide fragments, which are then split again into sp^2 amine, thioaldehyde, and aldehyde subfragments. This enables the mutual polarization of thioamide and amide to be calculated, but it could misrepresent the electronic distribution (and hence $E_{\mathrm{MTP}*}$) as the conjugation that is actually present in CTU is missing. A second alternative method involves assembling CTU using the multipoles and polarizabilities of the amine, thioaldehyde, and aldehyde fragments, all of which are calculated independently. This enables the polarization between all four fragments to be computed, but at the cost of a further loss of accuracy in the representation of the electron density. We assess the advantages and limitations of these two approaches below.

Table 4 Construction of CTU by assembling thioamide and amide fragments. Values of the intermolecular interaction energies (kcal/mol) at the equilibrium distances and their contributions to the binding of probe water molecules with the sites on CTU are shown

CTU	O (inner) d=2.9 Å	O (outer) d=2.1 Å	O (π) d=2.6 Å	S (outer) d=2.7 Å	S (inner) d =3.0 Å	S (π) d=2.7 Å
E_C	−3.9	−6.7	−2.1	−5.2	−1.8	−4.4
E_{mtp}	−4.9	−6.9	−2.9	−4.4	−2.3	−4
E_{exc}	2.2	4	0.9	3.1	0.6	2.7
E_{rep}	2.2	3.9	0.9	2.8	0.6	2.6
E_{pol}	−0.4	−1	−0.4	−0.6	−0.3	−0.6
E_{pol}	−0.5	−0.8	−0.3	−0.4	−0.3	−0.4
E_{ct}	−0.3	−0.5	−0.2	−0.5	−0.1	−0.6
E_{ct}	−0.2	−0.3	−0.1	−0.3	−0.1	−0.2
ΔE_{RVS}	−2.4	−4.3	−1.8	−3.3	−1.6	−2.9
E_{tot}	−3.4	−4.1	−2.4	−2.3	−2.1	−2
CTU N_{prim} d=2.8 Å	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}	
RVS	−1.3	0.7	−0.3	−0.1	−1	
SIBFA	−0.7	0.7	−0.2	−0.1	−0.3	
CTU N_{sec} d=3.3 Å	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}	
RVS	−0.5	0.1	−0.1	0	−0.6	
SIBFA	−0.8	0.1	−0.1	0	−0.8	
CTU NH d=2.3 Å	E_C/E_{mtp}	E_{exc}/E_{rep}	E_{pol}	E_{ct}	E_{tot}	
RVS	−5.8	2.5	−0.8	−0.3	−4.6	
SIBFA	−6.6	2.1	−1.2	−0.4	−6.1	

Assembling CTU from thioamide and amide fragments (representation "b")

The binding energies of water with CTU at the equilibrium distance are reported in Table 4. For completeness, they are also given in Table S3 of the ESM for a range of distances that are shorter than the equilibrium distance. Not unexpectedly, the agreement between these results and the QC results is less than the agreement between the QC results and those obtained when treating CTU as an integral unit. The errors in ΔE can exceed 1 kcal/mol in some cases, for instance when inner water binds to the carbonyl O or to the inner NH group. These errors can be traced back essentially to the electrostatic contribution. This may imply that alternative ways to derive the multipolar expansion on highly conjugated *and* flexible fragments while retaining the net charge on the subfragments must be sought (see [33] for an example). Work will have to be considered along these lines. It should be noted that this issue does not apply in the general case of the SIBFA library of fragments, which essentially consists of saturated fragments or rigid conjugated or aromatic ones.

Using this set of multipoles, we recalibrated the V_0 values for the eight torsion angles of lig-47. Those values are shown in Fig. 11, which displays the new conformational energy curves. The V_0 values are seen to vary by only small amounts from the values obtained for the integral CTU, with the notable exception of n=1 for $\varphi 5$ (V_0 for n=1 is −15 kcal/mol, as compared to 6 when calculated for the integral CTU). This is an interesting case of the possible impact of the representation of conjugation on E_{MTP*}, which can be corrected for by using a different sign and amplitude of V_0. The eight new curves have similar overlaps to the QC ones, with the possible exception of $\varphi 5$, for which the 50–100° region has δE_{conf} (SIBFA) values that are up to 5 kcal/mol higher than the QC ones (i.e., they are in the 15–20 kcal/mol range, in contrast to the δE_{conf} (QC) plateau at 15 kcal/mol).

The corresponding values of the conformational energy of lig-47 are reported in Table 5, along with the individual contributions to the SIBFA energy. The regression graph and the evolutions of δE_{conf} (QC), δE_{conf} (SIBFA), and the various contributions to δE_{conf} (SIBFA) as functions of the conformer are reported in Fig. 12. It should be noted that these

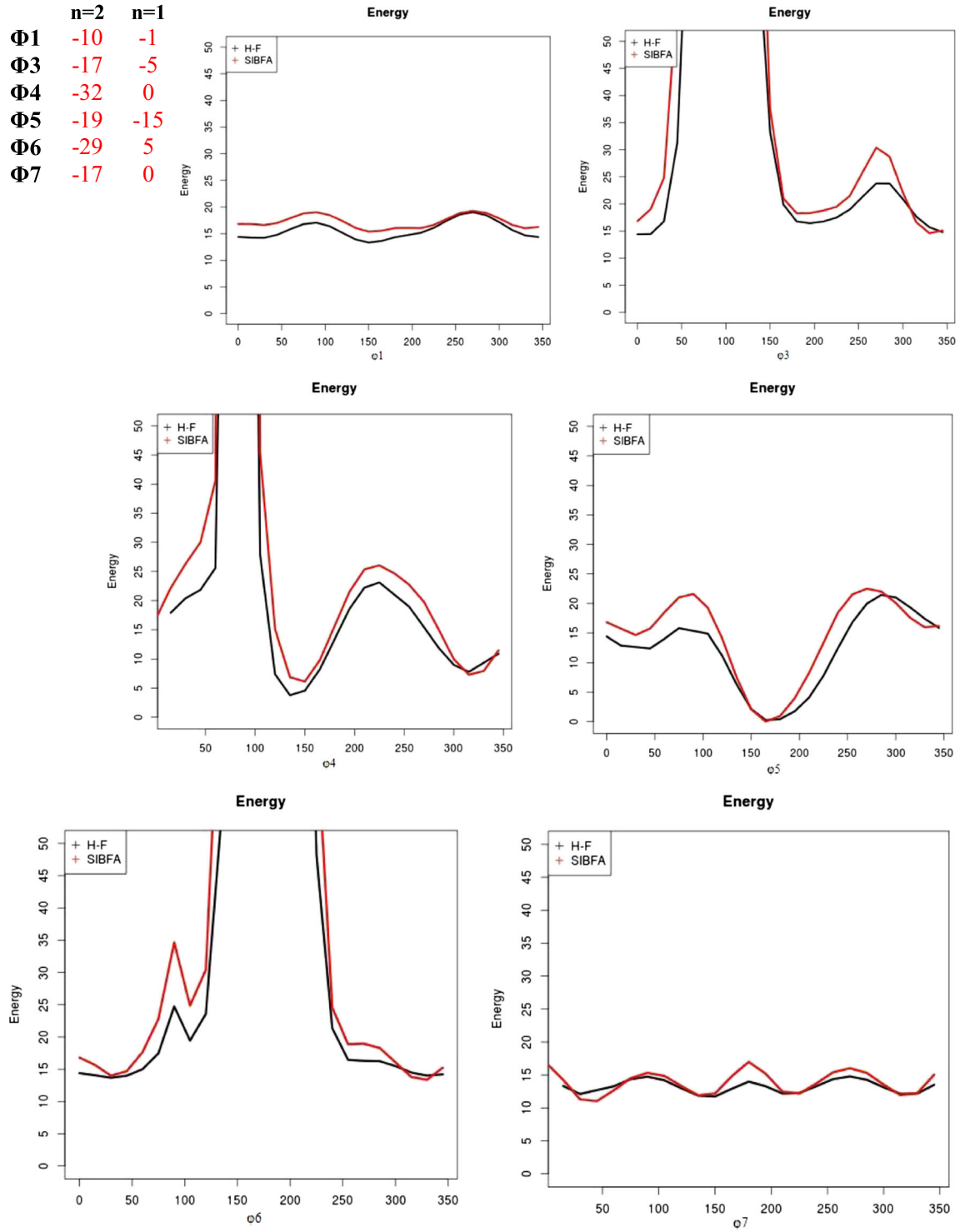


Fig. 11 Construction of CTU by assembling thioamide and amide fragments. Variations in the conformational energy of lig-47 as functions of the torsion angles $\varphi 1$ and $\varphi 3$–$\varphi 7$ are shown

results show very significant improvements in both the r^2 value (0.96 as compared to 0.88) and the slope (0.94 as compared to 0.63) compared to the results for the integral CTU. Table 5 shows that the δE_{conf}

Table 5 Construction of CTU by assembling thioamide and amide fragments. The relative QC and SIBFA conformational energies (kcal/mol) of all conformers are shown, along with the contributions to the SIBFA energy

Conformer	HF	$E_{1order}+E_{2order}+E_{tor}$	E_{mtp}	E_{rep}	E_{1order}	E_{pol}	E_{ct}	$E_{1order}+E_{pol}+E_{ct}$	E_{tor}
0a	20.97	19.50	20.30	−3.9	16.40	15.50	0.80	32.7	−13.2
0b	21.22	19.70	22.50	−7.7	14.80	14.40	0.90	29.9	−10.2
1a	0.17	0.00	0.00	0	0.00	0.00	0.00	0	0
1b	0.00	1.20	3.40	−8	−4.60	1.60	0.20	−2.8	4
2a	17.80	19.80	16.00	−1.7	14.30	18.00	−0.50	31.7	−11.9
2b	19.49	18.90	21.70	−9.3	12.40	16.00	0.80	29	−10.1
3a	18.42	18.10	15.90	−3.2	12.70	18.10	−0.60	30	−11.9
3b	18.53	18.30	16.00	−3.2	12.80	18.10	−0.70	30.1	−11.8
3c	20.97	19.50	20.30	−3.9	16.40	15.50	0.80	32.7	−13.2
4a	22.39	21.70	10.80	−10.1	0.70	19.20	1.10	21	0.7
4b	21.36	23.90	10.60	−15.3	−4.70	19.70	0.90	15.9	8
5a	21.93	19.30	10.00	−10.2	−0.20	18.40	1.10	19.2	0.1
6a	16.78	17.00	14.80	−4.3	10.50	18.50	−0.50	28.5	−11.5
6b	20.12	19.50	19.50	−4.8	14.70	16.30	0.90	31.7	−12.2
6c	19.22	18.80	21.10	−8.2	12.90	16.10	0.90	29.8	−11

(SIBFA) values are very close to the corresponding QC ones, and the energy separation between the lowest-lying minima and the other minima is now in the correct range of 17–24 kcal/mol, instead of 11-15 kcal/mol as previously. E_{pol}, now involving the NH group of thioamide and the CO group of formamide, is the decisive contributor that sets the two minima apart from the rest.

Assembling CTU from isolated amine, thioaldehyde, and aldehyde fragments (representation "c")

A "minimalist" representation of CTU using the multipoles and polarizabilities of isolated amine, thioaldehyde, and aldehyde fragments further degrades the agreement with the QC results for the interaction of the water probe with the CTU sites (the RVS and SIBFA results are reported in Table S4 of the ESM). This is a consequence of the loss of conjugation, as reflected in some poor values of E_{MTP*} as compared to E_C, particularly upon binding to the S atom. The six conformational energy curves of lig-47 are shown in Table S5 of the ESM. They are similar to their QC counterparts (as also seen for the previous representations of CTU). However, this was only true following a significant recalibration of V_0 for $\varphi 4$, $\varphi 5$, and $\varphi 6$. Thus, for $n=1$, the V_0 values are −22, −2, and −11 kcal/mol, respectively; they were 0, −15, and 5 kcal/mol for the previous representation. The conformational energy results are given in Table 6. Figure 13a and b show the regression graph and the evolution of δE_{conf} as a function of conformer number, respectively. The r^2 value is less (0.80) than in the previous case (0.96), whereas the slope is the same (0.94). Generally speaking, δE_{conf} (SIBFA) has comparable values to δE_{conf} (QC), but there are two conspicuous outliers with values of 27.8 and 29.6 kcal/mol as compared to 21.5 and 21 kcal/mol, respectively, from QC.

It can be concluded from this section that, even though it gives relative δE_{conf} values that generally agree with those afforded by QC, the minimalist approach is unreliable, as indicated by the presence of outliers and a poor match to ΔE (RVS) in some cases.

Possible impact of correlation/dispersion

The previous results were obtained at the Hartree–Fock level, since we were seeking a "proof of principle" procedure to justify the construction of lig-47. The impact of correlation/dispersion was further evaluated by recomputing δE_{conf} (QC) at correlated levels, namely MP2, B3LYP [34, 35], and B97-D [36]. These results are reported in Table 7, along with the SIBFA ones from representations a–c, which now include E_{disp}. MP2 is the QC approach that leads to the largest reduction in δE_{conf}. This reduction is in

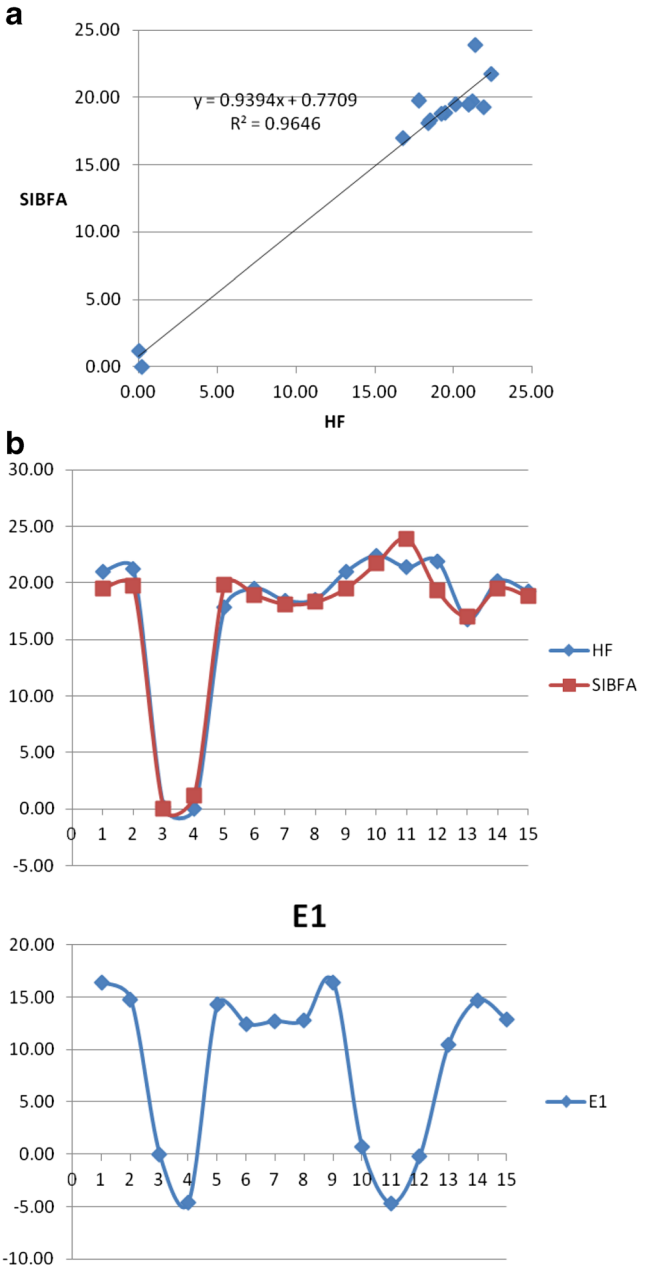


Fig. 12a–b Construction of CTU by assembling thioamide and amide fragments. **a** Regression graph showing how closely δE_{conf} (SIBFA) and δE_{conf} (HF) match. **b** Comparison of the evolutions of δE_{conf} (QC), δE_{conf} (SIBFA), and the components of δE_{conf} (SIBFA) as functions of the conformer

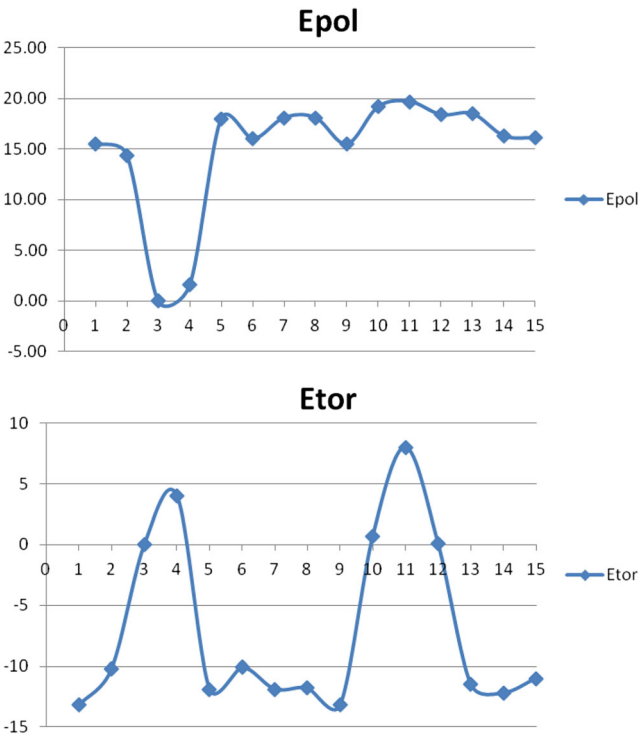


Fig. 12 (continued)

fact more pronounced than the reduction which occurs upon including E_{disp} in SIBFA. MP2 was previously shown to overestimate the stability of folded conformers via intramolecular basis set superposition error calculations [37, 38] (e.g., for 1a, 1b, 2b, 3a, and 6b), and this can explain their excessive decreases in δE_{conf}. The results from B3LYP, which includes correlation but does not incorporate dispersion effects, are very close to the δE_{conf} (SIBFA) results for representation b (r^2=0.98), even though the multipoles and polarizabilities were derived at the HF level. δE_{conf} (SIBFA) also agrees closely with B97-D, which includes an explicit dispersion term, although this agreement is not as strong as it is with B3LYP (r^2=0.95 instead of 0.98). Representations a and c lead to significantly reduced agreement with the QC calculations. In this context, several papers have showed the importance of including correlation effects when estimating the rotational barriers around conjugated bonds [39–43]. This led us to recalibrate V_0 (n=1) and V_0 (n=2) against B97-D conformation–energy curves. A comparison of the evolutions of the B97-D and SIBFA conformational energies is provided in Fig. 14, which also shows the recalibrated V_0 barriers. The most important changes concern the two central torsion angles, φ4 and φ5 of the thioamide moiety. A substantial improvement with respect to B97-D is apparent. The conformational energies are listed in the penultimate column of Table 7. Those results imply that in the context of polarizable MM, the

Table 6 Construction of CTU by assembling sp^2 amine, thioaldehyde, and aldehyde fragments. Relative QC and SIBFA conformational energies (kcal/mol) of all conformers are shown, along with the contributions to the SIBFA energy

Conformer	HF	$E_{1order}+E_{2order}+E_{tor}$	E_{mtp}	E_{rep}	E_{1order}	E_{pol}	E_{ct}	$E_{1order}+E_{pol}+E_{ct}$	E_{tor}
0a	20.97	21.30	3.80	3.80	7.60	20.70	0.30	28.7	−7.4
0b	21.22	20.50	6.10	−0.20	5.90	18.50	0.30	24.8	−4.3
1a	0.17	3.30	3.00	8.40	11.40	−0.70	0.10	10.7	−7.4
1b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0
2a	17.80	21.40	3.30	6.40	9.70	−6.10	−0.20	3.5	17.9
2b	19.49	21.80	5.50	−1.90	3.60	22.60	0.30	26.5	−4.7
3a	18.42	23.50	4.80	4.90	9.70	−4.30	−0.20	8.3	15.2
3b	18.53	19.60	4.90	4.90	9.80	−4.20	−1.40	4.2	15.4
3c	20.97	21.30	3.80	3.80	7.60	20.70	0.30	28.7	−7.4
4a	22.39	18.40	−0.30	−2.50	−2.80	−3.50	0.30	−5.9	24.3
4b	21.36	27.10	−5.00	−7.90	−12.90	3.70	0.10	−9.2	36.3
5a	21.93	25.40	−1.30	−2.70	−4.00	7.40	0.30	3.6	21.8
6a	16.78	20.70	2.70	3.80	6.50	−1.10	−0.10	5.3	15.4
6b	20.12	15.90	3.50	2.90	6.40	13.20	0.30	19.9	−4
6c	19.22	14.00	5.50	−2.60	2.90	13.70	0.30	16.9	−2.9

inclusion of an explicit dispersion contribution could allow the changes in δE_{conf} (QC) upon passing from the HF to the correlated level to be accounted for, with an additional improvement being afforded by a preliminary recalibration of V_0. However, this approach could be more problematic when using classical molecular mechanics potentials in which the $1/R^6$ vdW contribution contains more than simply the dispersion.

Conclusions

Rotations around bonds connecting unsaturated atoms govern the conformations of numerous pharmacological ligands. This is exemplified by lig-47, a molecule that was selected via virtual screening and is active both in vitro and in cellulo against neovascularization [1, 2]. Lig-47 is constructed from a benzimidazole, a methylbenzene, a carboxythiourea, and a benzene-linked dioxane. In the context of molecular mechanics/dynamics, it is essential to reliably account for the variations in conformational energy that result from rotations around its connecting bonds in order to preserve the proper balance of intramolecular interactions on the one hand and intermolecular interactions (i.e., ligand–receptor and ligand–water) on the other. Both of these types of interactions can only be validated through comparison with the results from high-level ab initio QC computations. Calibrating the one- and two-fold rotational barriers of lig-47 is obviously necessary but is also far from sufficient. The presence of a highly conjugated polarizable yet flexible central carboxythiourea (CTU) entity led us to address the following issues in succession. To our knowledge, there is no precedent for such analyses.

– *Anisotropy*. The energy contributions to SIBFA are all anisotropic (see [7–10] for discussions). The anisotropy of the overlap-dependent terms E_{rep} and E_{ct} stems from the explicit representation of the heteroatom lone pairs. However, there is an extra amount of anisotropy that could be due to the unequal spatial extensions and lone pairs that are centered on a given atom. Regarding the 'inner' sp^2 lone pairs of the thiocarbonyl and carbonyl bonds, this could be due to the mutual repulsion of these lone pairs. We have accordingly modified the vdW radii of the CTU S, O, and N lone pairs. QC energy decomposition analysis (EDA) of the interaction of a water probe with CTU enabled us to calibrate the amounts by which the vdW radii of the inner lone pairs decrease and the outer lone pairs and π lone pairs increase. The modified vdW radii led to a significant improvement in the reproduction of the

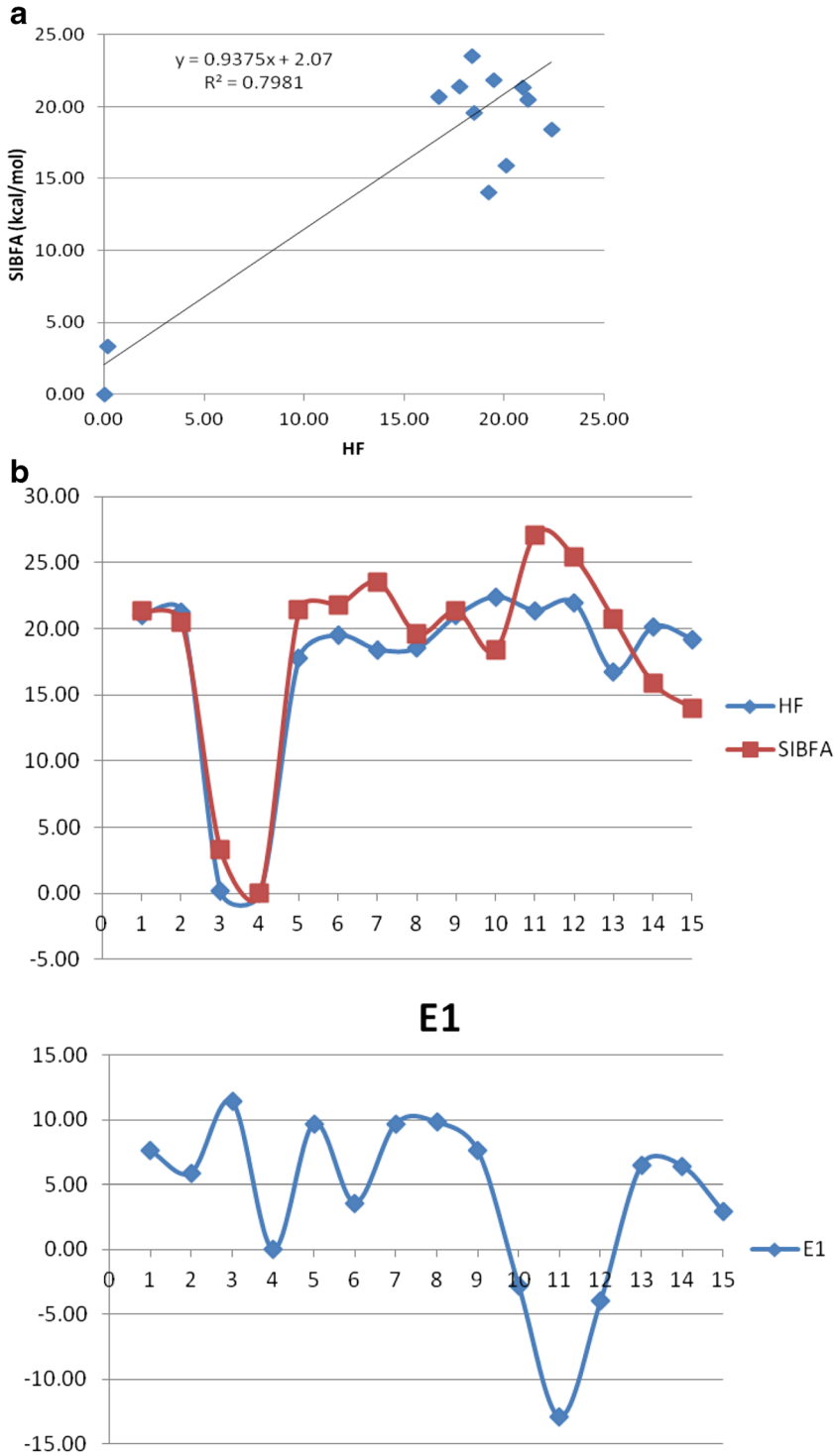


Fig. 13a–b Construction of CTU by assembling sp^2 amine, thioaldehyde, and aldehyde fragments. **a** Regression curve illustrating how closely δE_{conf} (SIBFA) and δE_{conf} (HF) match. **b** Comparison of the evolutions of δE_{conf} (QC), δE_{conf} (SIBFA), and the components of δE_{conf} (SIBFA) as functions of the conformer

QC exchange-repulsion term from EDA by E_{rep} (SIBFA). These tuned vdWs were then used in the second step of this study, namely the calibration of the torsional barriers.

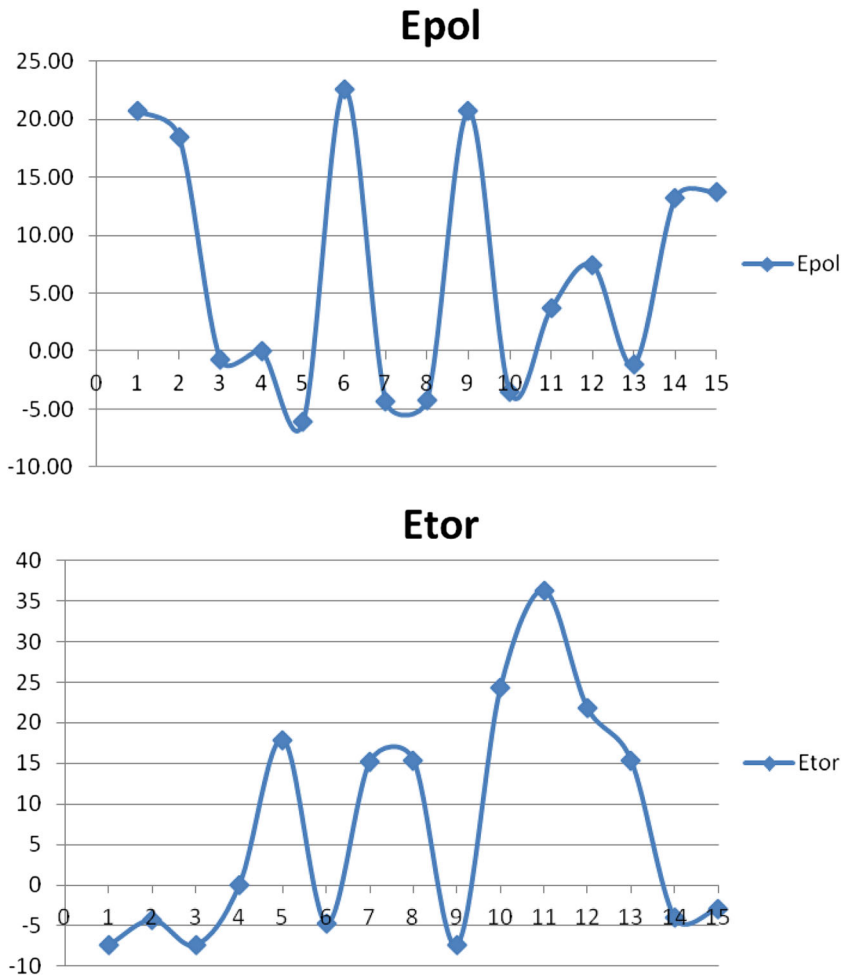


Fig. 13 (continued)

– *Conjugation*. In the context of MM/MD, conjugation raises two issues. The first is the need to calibrate onefold ($n=1$) and twofold ($n=2$) V_0 rotation barriers around each bond connecting sp/sp^2 atoms. The second is the requirement of a satisfactory representation of the electronic distribution of an "integral" conjugated fragment such as CTU. Using the multipolar distribution and the distributed polarizabilities of CTU split into four non-mutually-polarizable subfragments enabled us to calibrate the two rotation barriers around each of the three connecting bonds of CTU and to accurately reproduce the QC conformational energy curve. The study was then extended to lig-47 by calibrating the barriers of the three additional torsional bonds. For all six torsion angles, the δE_{conf} (SIBFA) values again closely matched their QC counterparts. Could such a representation of lig-47 be used in actual simulations? This led us to consider the next issue.

– *Polarization and multipole transferability*. The energy of lig-47 was subsequently minimized by simultaneously relaxing its six main-chain torsion angles using selected minima from the six conformation–energy curves as starting points. The two most strongly bound minima were stabilized by an intramolecular H-bond within CTU between the carbonyl O and the thioamide NH proton. However, their relative stabilizations were underestimated with respect to QC. Noting the prominent role of E_{pol} in stabilizing H-bonded networks [44, 45], and in an attempt to facilitate its involvement in bonding between the thioamide and amide moeities, we considered an alternative method of

Table 7 Comparison of δE_{conf} (QC) values obtained at the MP2, B3LYP, and B-97D levels with the corresponding δE_{conf} (SIBFA) values that incorporate the dispersion contribution. r^2 regression coefficients are shown at the bottom of the table. Energies are in kcal/mol

Conformer	MP2	B97d	B3LYP	E_{tot} (SIBFA) a	E_{tot} (SIBFA) b	SIBFA (B97D)	E_{tot} (SIBFA) c
0a	18.4	19.1	19.5	11.6	19.4	19	41.4
0b	19.1	19.9	19.9	12.9	19.9	20	19
1a	0	0	0	1.4	0	0	23.5
1b	1.11	1.2	1	0	2	2.5	0
2a	13.9	14.8	16.8	13.5	18.6	16.1	40.5
2b	18.2	19	18.9	11.3	19.4	20	20.7
3a	15.1	15.8	17.6	9.7	17.3	14.6	39.9
3b	15.2	15.9	17.7	10	17.5	14.8	17.7
3c	18.4	19.1	19.5	11.6	19.4	18.9	41.4
4a	15	16.9	20.7	15.5	20.6	15.7	37.6
4b	12.5	14.3	20.2	14.5	21.5	17.2	23
5a	14.2	16.4	20.8	13.1	17.4	13.5	43.7
6a	13.8	14.4	16.2	10	16.4	13.7	40.4
6b	17.3	17.9	18.4	12.8	19.6	18.8	36.3
6c	17.4	18	18	12.8	19.2	19.1	34.7
				MP2	B97D	B3LYP	
a				0.84	0.88	0.95	
b				0.93	0.95	0.98	
c				0.52	0.54	0.57	

constructing CTU: assembling it from the thioamide and amide entities, which were then further split into amine and thioaldehyde on the one hand and amine and aldehyde on the other. There is thus a trade-off between the lost conjugation between these two entities (which is present in integral CTU) and the increase in their mutual polarization. Following a new recalibration step for the vdWs and the torsional barriers, the δE_{conf} values at the energy minima were recomputed. This led to a marked improvement in the δE_{conf} values such that they were now in near-quantitative agreement with the QC ones, with an r^2 regression coefficient of 0.97 and a slope of 0.98. E_{pol} was observed to be the main contributor to the δE_{conf} (SIBFA) values. When the mutual polarization of thioamide and amide entities of CTU was accounted for, the evolution of E_{pol} as a function of conformer number was found to give the closest match to the corresponding evolution of δE_{conf} (QC). Finally, a minimalist representation in which CTU was constructed from the multipoles and polarizabilities of independent sp^2 amine, thioaldehyde, and aldehyde fragments was analyzed. It presented significantly weaker agreement with the QC values in terms of both CTU interactions with a water probe and δE_{conf}.

Along these lines, several papers have mentioned the need to introduce coupling between the torsional degrees of freedom when treating polyconjugated ligands [46, 47]. However, those studies were performed in the context of nonpolarizable molecular mechanics. It is likely that such couplings indirectly accounted for many-body effects that are otherwise absent from such approaches, while they are inherently included in the E_{pol} contribution.

The present results again underline the essential role of polarization, as found in our previous conformational studies of peptides [23], conformation-dependent divalent cation binding [48–50], ligand–macromolecule complexes [51–53], and networks of H-bonded complexes [44, 45]. They are fully consistent with the conclusions drawn about another polarizable multipolar force field, AMOEBA [54–57], although AMOEBA does not incorporate anisotropy at the level of the short-range repulsion contribution.

	n= 2	n= 1
Φ1	-10	-1
Φ3	-20	-9
Φ4	-32	3
Φ5	-16	-15
Φ6	-25	8
Φ7	-17	0

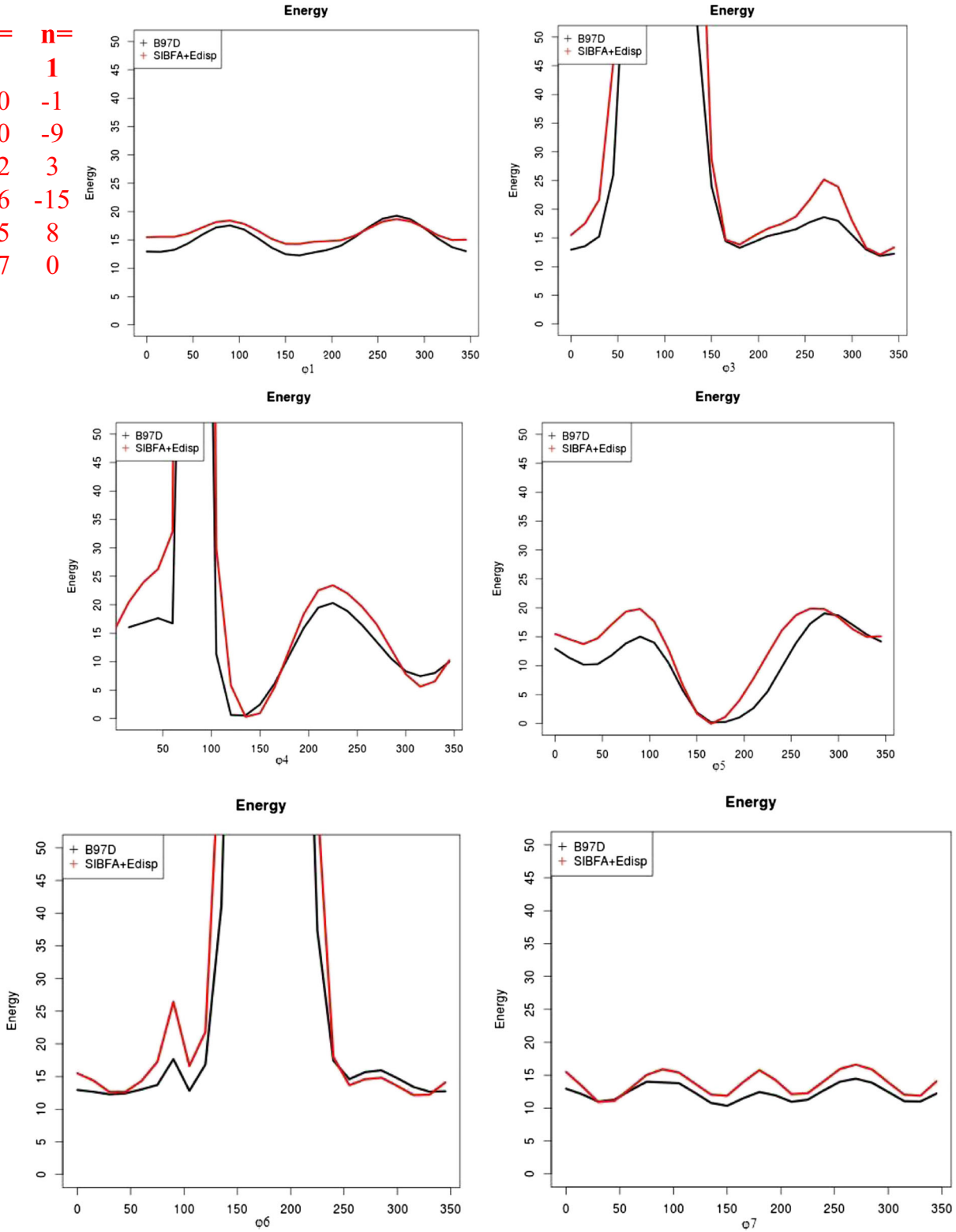


Fig. 14 Construction of CTU by assembling thioamide and amide fragments. Variations in the conformational energy of lig-47 as functions of the torsion angles $\varphi 1$ and $\varphi 3$–$\varphi 7$

We are presently using the minimized conformers of lig-47 as starting points for docking simulations of the target protein neuropilin-1, and the results will be reported separately.

Acknowledgments We wish to thank the Grand Equipement National de Calcul Intensif (GENCI): Institut du Developpement et des Ressources en Informatique Scientifique (IDRIS), Centre Informatique de l'Enseignement Superieur (CINES), France, project no. x2009-075009), and the Centre de Ressources Informatiques de Haute Normandie (CRIHAN, Rouen, France), project 1998053.

We wish to acknowledge a CIFRE grant allotted to Elodie Goldwaser in the course of her Ph.D. thesis.

We are pleased to thank Drs. Lucia Borriello and Pascal Dao for enriching discussions during the course of this work.

References

1. Djordjevic S, Driscoll PC (2013) Drug Discov Today 18:44
2. Allain B, Jarray R, Borriello L, Leforban B, Dufour S, Liu W, Pamonsinlapatham P, Bianco S, Larghero J, Hadj-Slimane R, Garbay C, Raynaud F, Lepelletier Y (2012) Cell Signal 24:214
3. Van der Kooi CW, Jusino MA, Perman B, Neau DB, Bellamy HD, Leahy D (2007) J Proc Natl Acad Sci USA 104:6152
4. Jarvis A, Allerston CK, Jia H, Herzog B, Garza-Garcia A, Winfield N, Ellard K, Aqil R, Lynch R, Chapman C, Hartzoulakis B, Nally J, Stewart M, Cheng L, Menon M, Tickner M, Djordjevic S, Driscoll PC, Zachary I, Selwood DL (2010) J Med Chem 53:2215
5. Jain A (2003) J Med Chem 46:499
6. Gresh N, Claverie P, Pullman A (1984) Theor Chim Acta 66:1
7. Gresh N (1995) J Comput Chem 16:856
8. Gresh N (2006) Curr Pharm Des 12:2121
9. Gresh N, Cisneros GA, Darden TA, Piquemal J (2007) J Chem Theory Comput 3:1960
10. Piquemal J-P, Chevreau H, Gresh N (2007) J Chem Theory Comput 3:824
11. Silvi B, Savin A (1994) Nature 371:683
12. Piquemal J-P, Pilme J, Parisel O, Gerard H, Fourre I, Berges J, Gourlaouen C, De La Lande A, Van Severen M-C, Silvi B (2008) Int J Quantum Chem 108:1951
13. Chaudret R, Gresh N, Cisneros GA, Scemama A, Piquemal J-P (2013) Can J Chem 91:1
14. Dunning TH (1989) J Chem Phys 90:1007
15. Feller D (1996) J Comput Chem 17:1571
16. Frisch MJ, Trucks GW, Schlegel HB et al. (2009) Gaussian 09, revision A.1. Gaussian, Inc., Wallingford
17. Stevens WJ, Fink W (1987) Chem Phys Letts 139:15
18. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347
19. Stone A (1981) J Chem Phys Letts 83:233
20. Stone AJ, Alderton M (1985) Mol Phys 56:1047
21. Piquemal J-P, Gresh N, Giessner-Prettre C (2003) J Phys Chem A 107:10353
22. Garmer DR, Stevens WJ (1989) J Phys Chem A 93:8263
23. Gresh N, Kafafi SA, Truchon J-F, Salahub DR (2004) J Comput Chem 25:823
24. Beachy MD, Chasman D, Murphy RB, Halgren TA, Friesner RA (1997) J Am Chem Soc 119:5908
25. Evangelakis GA, Rizos JP, Lagaris IE, Demetropoulos IN (1987) Comput Phys Comm 46:401
26. McInnes C (2007) Curr Op Chem Biol 11:494
27. Cavasotto CN, Orry A (2007) J Curr Top Med Chem 7:1006
28. Kroemer RT (2007) Curr Protein Pept Sci 8:312
29. Irwin J (2008) J Comp-Aided Mol Des 22:193
30. Sotriffer CA, Sanschagrin P, Matter H, Klebe G (2008) Proteins 73:395
31. Gilson MK, Zhou HX (2007) Ann Rev Biophys Biomol Struct 36:21
32. Schneider G (2010) Nat Rev Drug Discov 9:273
33. Ren P, Ponder JW (2003) J Phys Chem B 107:5933
34. Lee C, Yang W, Parr RG (1988) Phys Rev B37:785
35. Becke A (1993) J Chem Phys 98:5648
36. Grimme S (2006) J Comput Chem 27:1787
37. Gloaguen E, de Courcy B, Piquemal J-P, Pilme J, Parisel O, Pollet R, Biswal HS, Piuzzi F, Tardivel B, Broquier M, Mons M (2010) J Am Chem Soc 132:11860
38. van Mourik T (2008) J Chem Theory Comput 4:1610
39. Head-Gordon M, Pople JA (1993) J Phys Chem 97:1147
40. Head-Gordon M, Pople JA (1993) J Phys Chem 97:10250
41. Meier RJ (1993) J Phys Chem 97:10248
42. Meier RJ (2011) J Phys Chem 115:3604
43. Klug R, Burcl R (2010) J Phys Chem A 114:6401
44. Guo H, Gresh N, Roques BP, Salahub DR (2000) J Phys Chem B 104:9746
45. Piquemal J-P, Chelli R, Procacci P, Gresh N (2007) J Phys Chem A 111:8170
46. Zheng J, Yu T, Papajak E, Alecu IM, Mielke S, Truhlar DG (2011) Phys Chem Chem Phys 13:10885
47. Tafipolsky M, Schmid R (2005) J Comput Chem 26:1579
48. Rogalewicz F, Ohanessian G, Gresh N (2000) J Comput Chem 21: 963
49. Tiraboschi G, Fournié-Zaluski M-C, Roques B-P, Gresh N (2001) J Comput Chem 22:1038
50. Gresh N, Shi GB (2004) J Comput Chem 25:160
51. Antony J, Piquemal J-P, Gresh N (2005) J Comput Chem 26:1131
52. Courcy B, Piquemal J-P, Garbay C, Gresh N (2010) J Am Chem Soc 132:3312
53. Gresh N, Courcy B, Piquemal J-P, Foret J, Courtiol-Legourd, Salmon L (2011) J Phys Chem B 115:8304
54. Jiao D, Golubkov PA, Darden TA, Ren P (2008) Proc Natl Acad Sci USA 105:6290
55. Jiao D, Zhang JJ, Duke RE, Li GH, Schnieders MJ, Ren P (2009) J Comput Chem 30:1701
56. Ren P, Wu C, Ponder JW (2011) J Chem Theory Comput 7:3027
57. Zhang J, Yang W, Piquemal J-P, Ren P (2012) J Chem Theory Comput 8:1314