Supporting Information for:

Bartender: Martini 3 Bonded Terms via Quantum Mechanics-based Molecular Dynamics

Gilberto P. Pereira[a,b], Riccardo Alessandri[c], Moisés Domínguez[a], Rocío Araya-Osorio[e], Linus Grünewald[f], Luís Borges-Araújo[a,b], Sangwook Wu[g,h], Siewert-Jan Marrink[f], Paulo C. T. Souza*[a,b], Raul Mera-Adasme*

[ea] Laboratoire de Biologie et Modélisation de la Cellule, CNRS, UMR 5239, Inserm, U1293, Université Claude Bernard Lyon 1, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364, Lyon, France.

[b] Centre Blaise Pascal de Simulation et de Modélisation Numérique, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364, Lyon, France.

[c] Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, United States

[d] Departamento de Ciencias del Ambiente, Facultad de Química y Biología, Universidad de Santiago de Chile (USACH), Av. Libertador Bernardo O'Higgins 3363, 9170022 Estacion Central, Chile.

[e] Departamento de Quimica, Facultad de Ciencias, Universidad de Tarapacá, Av. Gral. Velasquez 1775, Arica, Chile.

[f] Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

[g] PharmCADD, Busan, 48792, Republic of Korea.

[h] Department of Physics, Pukyong National University, Busan, 48513, Republic of Korea.

* Paulo Cesar Telles de Souza - email: paulocts@gmail.com, and Raul Mera-Adasme - email: rmeraa@academicos.uta.cl.

Table of contents

1. Scheme illustrating how to compute bonded-distribution overlap

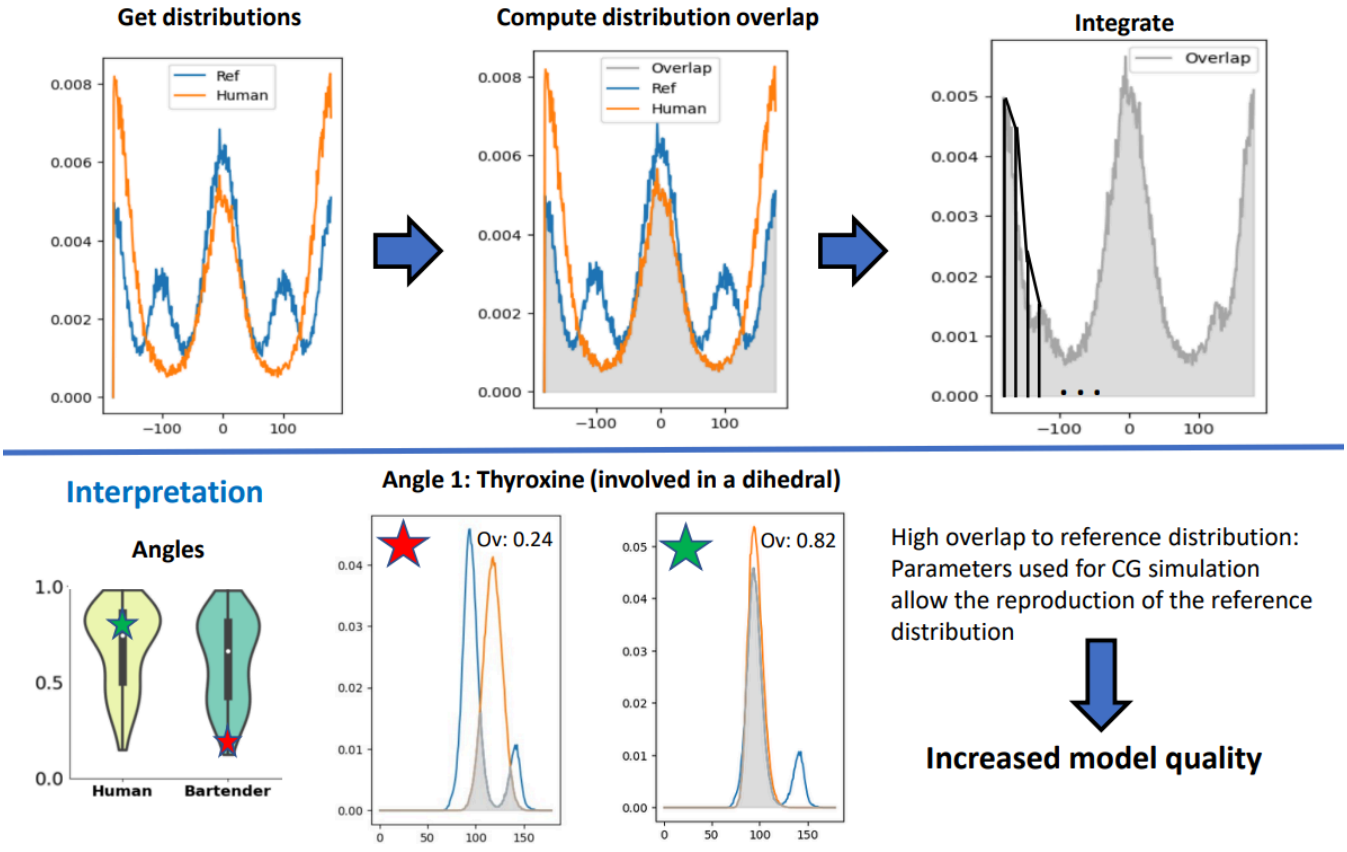


Figure S1 - Scheme illustrating how to compute bonded-distribution overlap. Angle 1 from Thyroxine is used to illustrate both suboptimal and optimal overlap calculations.

2. Martini 3 small molecule dataset: extended data

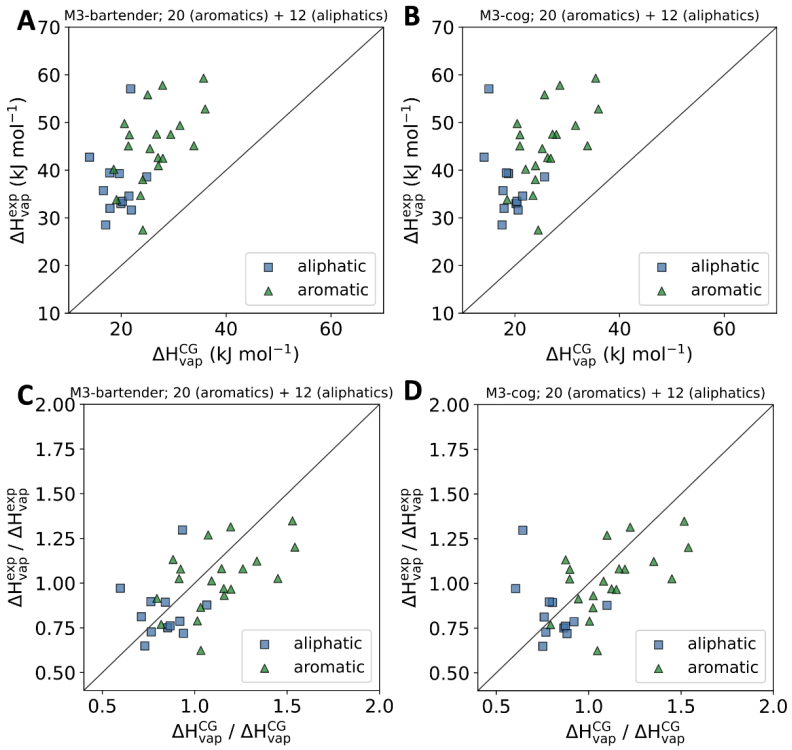
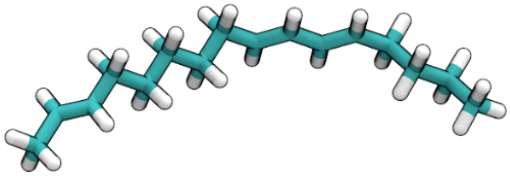


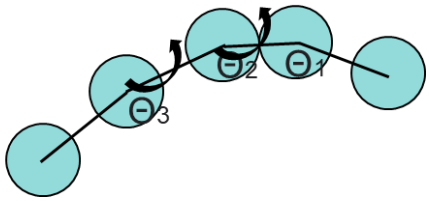
Figure S2. Bartender benchmark on the Martini 3 small molecule dataset bonded parameters: extended data. Correlation of the experimental heat of vaporization and the computed heat of vaporization obtained with the models parameterized using (A) Bartender and (B) human experts.[25] (C) and (D) show the same data as (A) and (B), respectively, but now normalized by dividing for the heat of vaporization of water (43.98 kJ mol^{-1} the experimental one, and 23.35 kJ mol^{-1} the one of Martini 3 water), a comparison already used in Ref 25. In all plots, green triangles denote data points associated with aromatic compounds, while blue squares represent those related to aliphatic compounds. The number of compounds considered correspond to the one for which experimental data are available. Note that there are more data points than in Ref 25.

3. Molecules used in the Flexible Molecules dataset

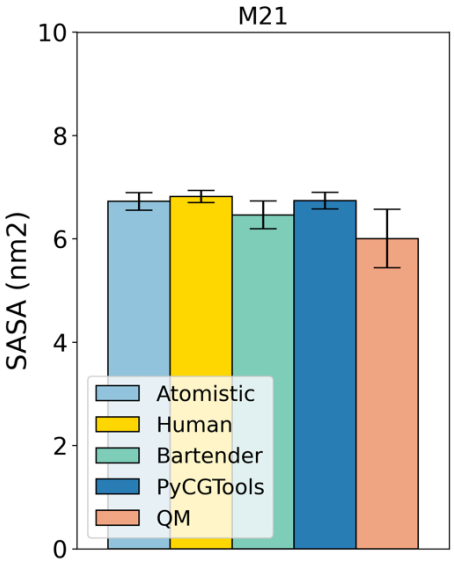
Molecule 1



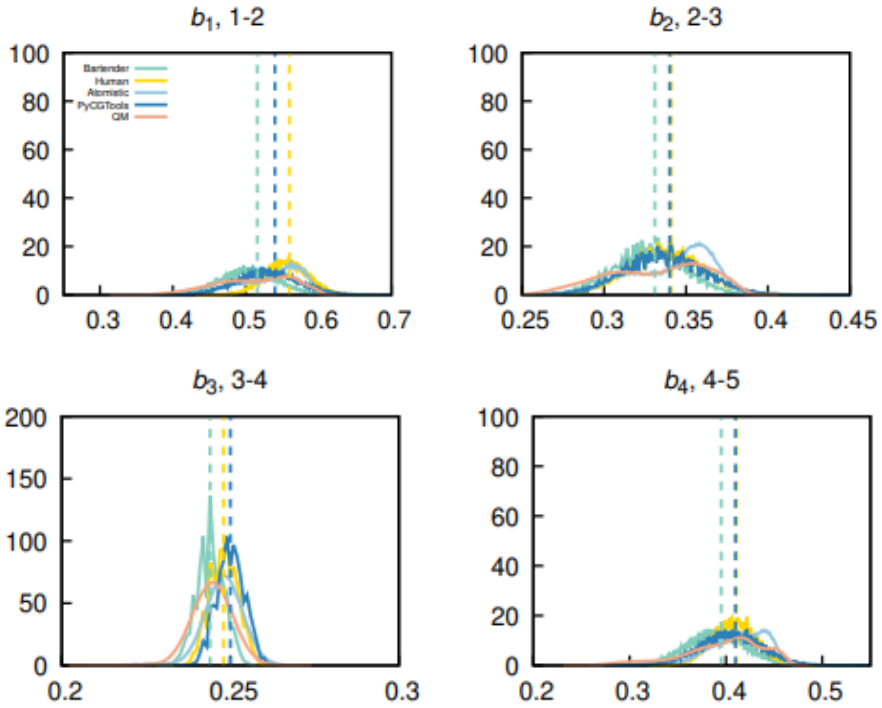
Bonded network



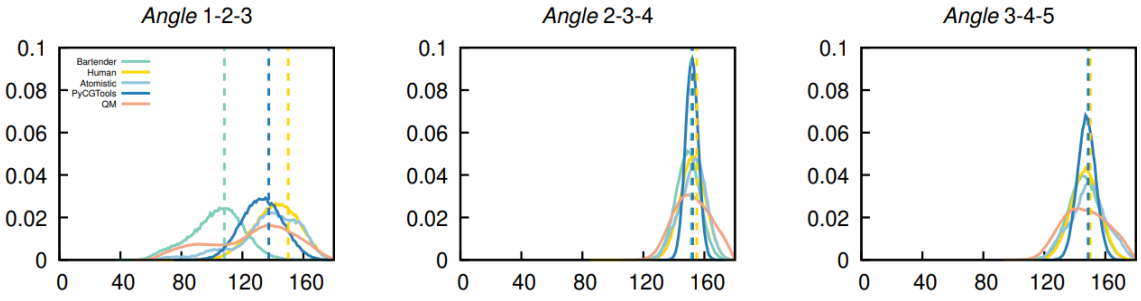
SASA



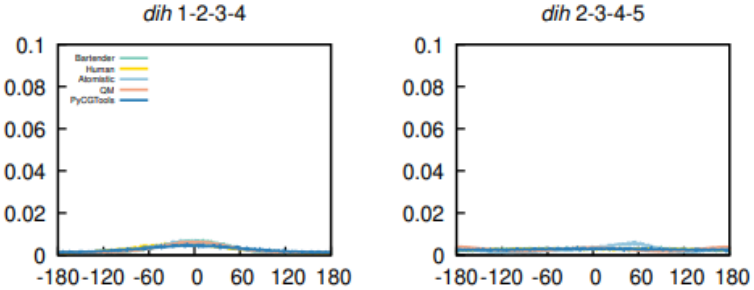
Bonds



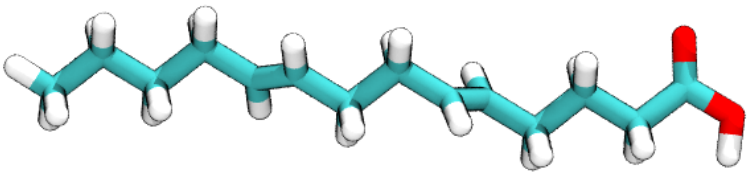
Angles



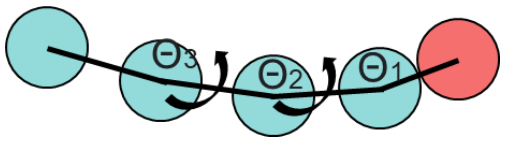
Dihedrals



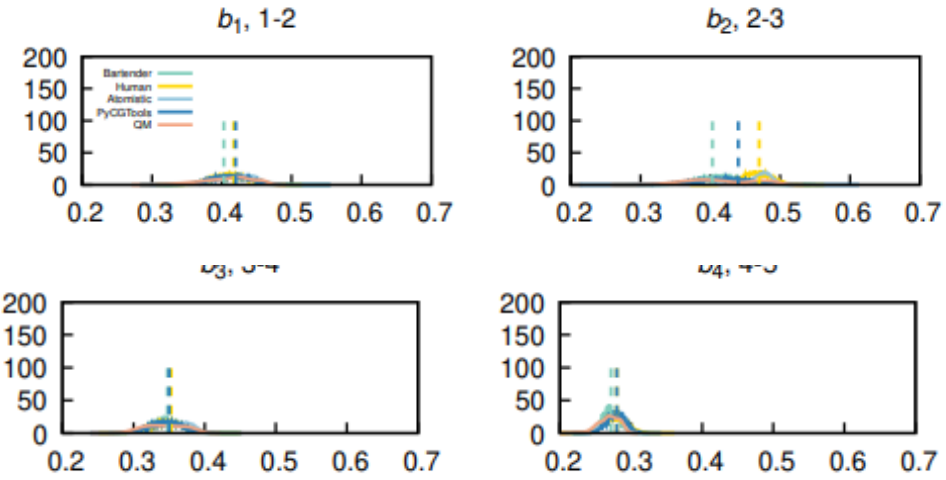
Molecule 2



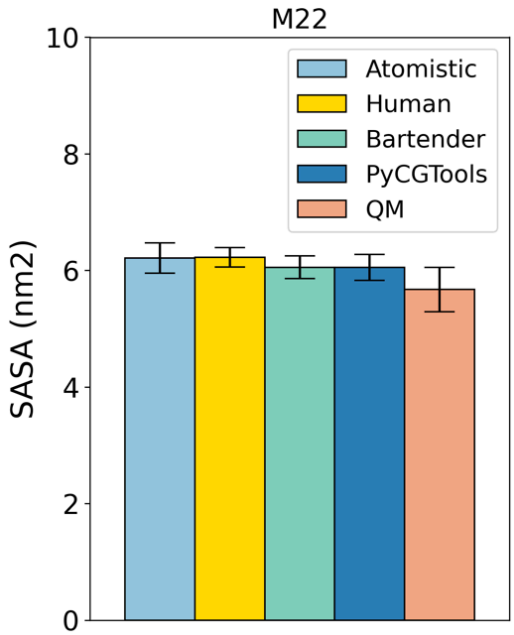
Bonded network



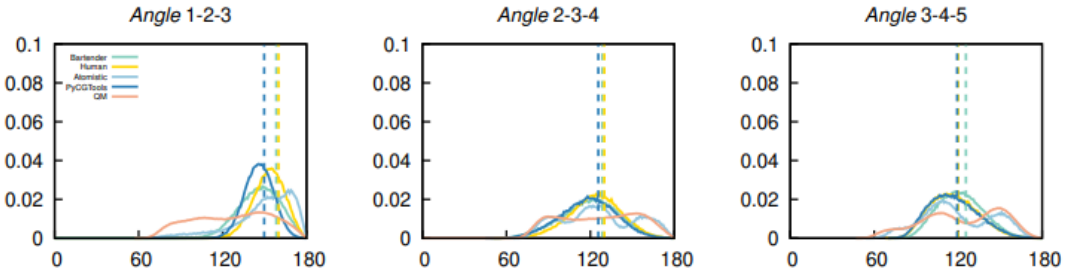
Bonds



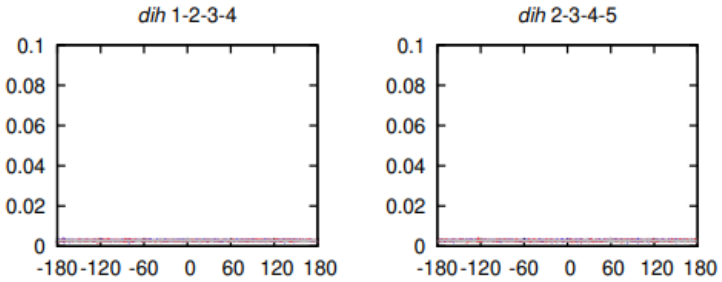
SASA



Angles

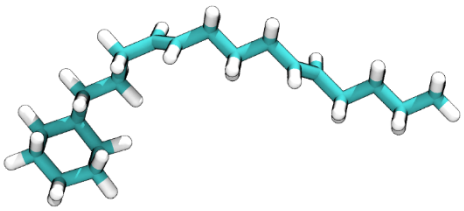


Dihedrals

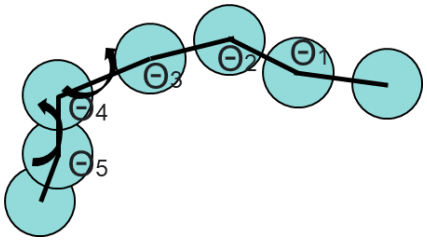


No meaningful dihedrals required.

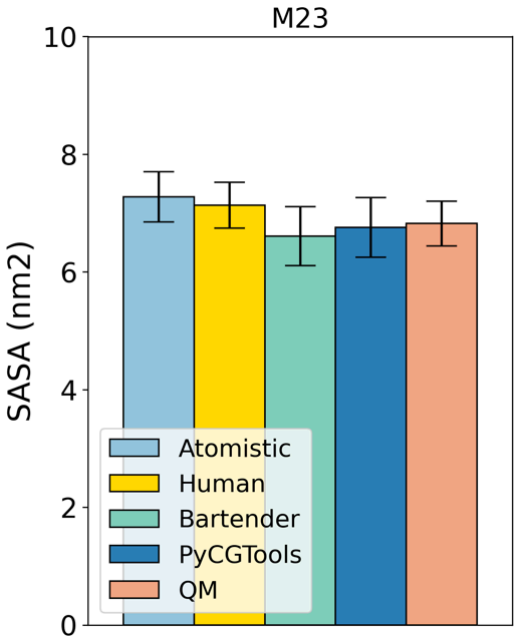
Molecule 3



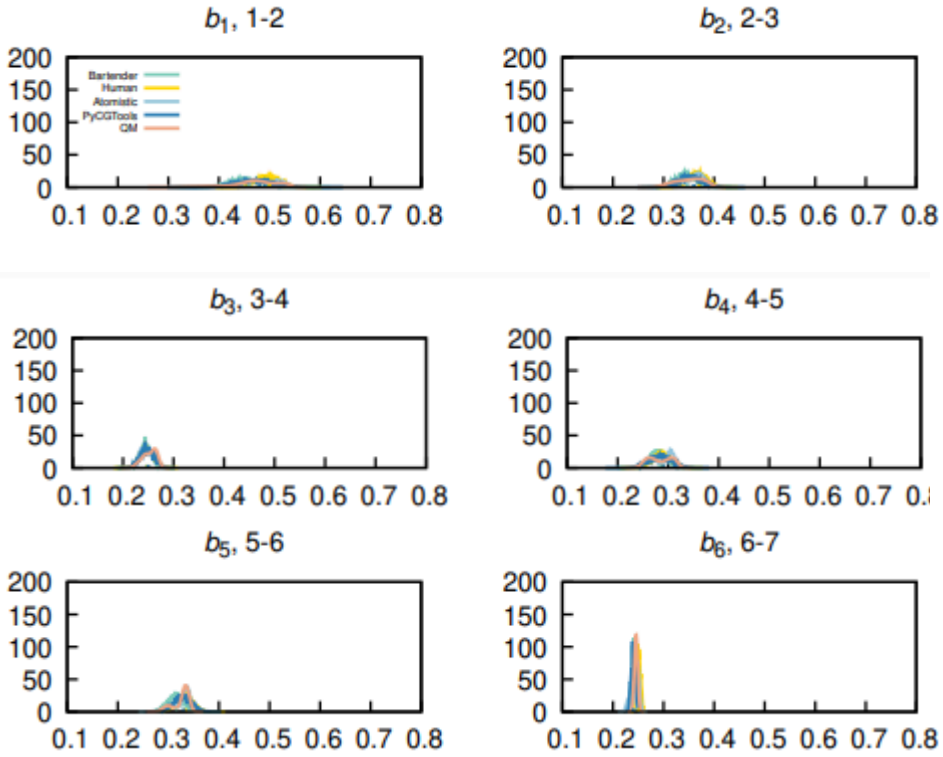
Bonded network



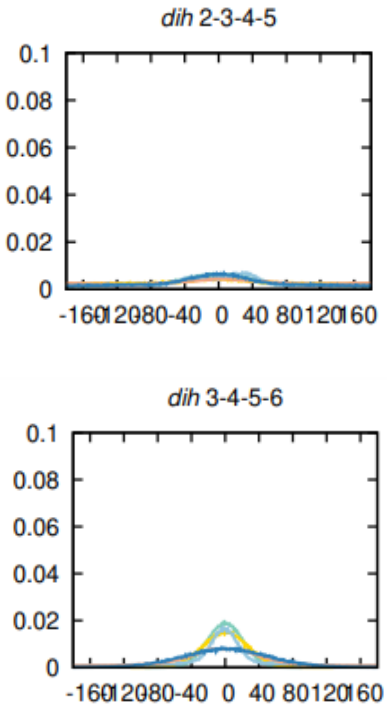
SASA



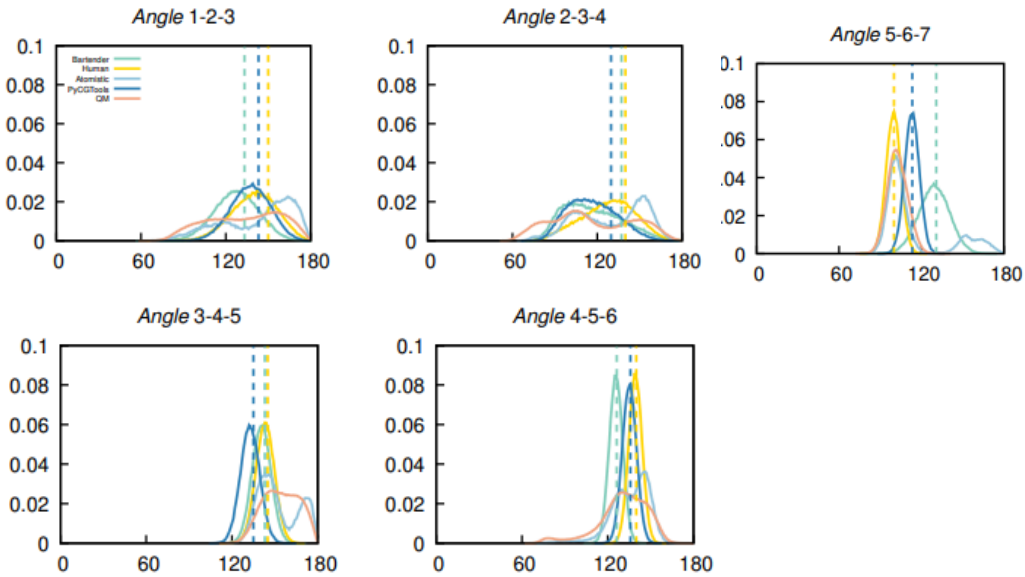
Bonds



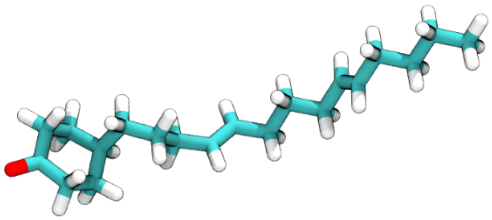
Dihedrals



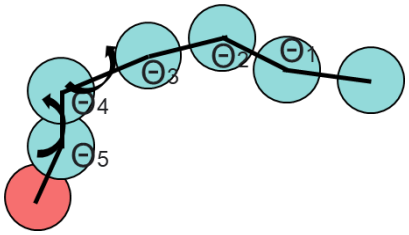
Angles



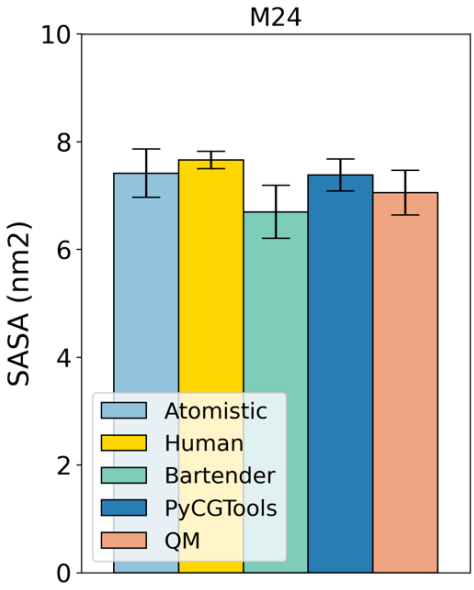
Molecule 4



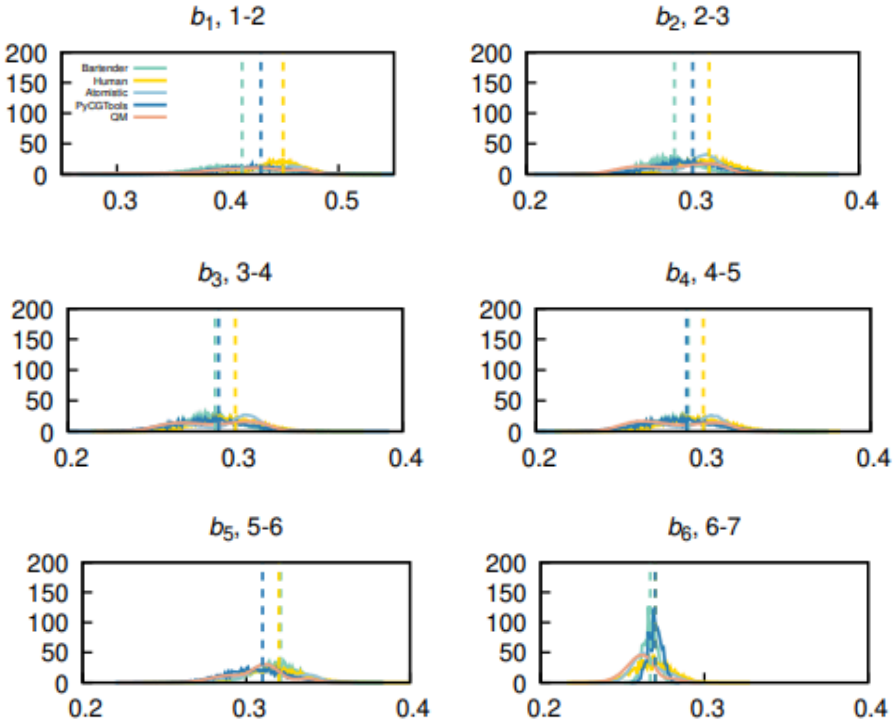
Bonded network



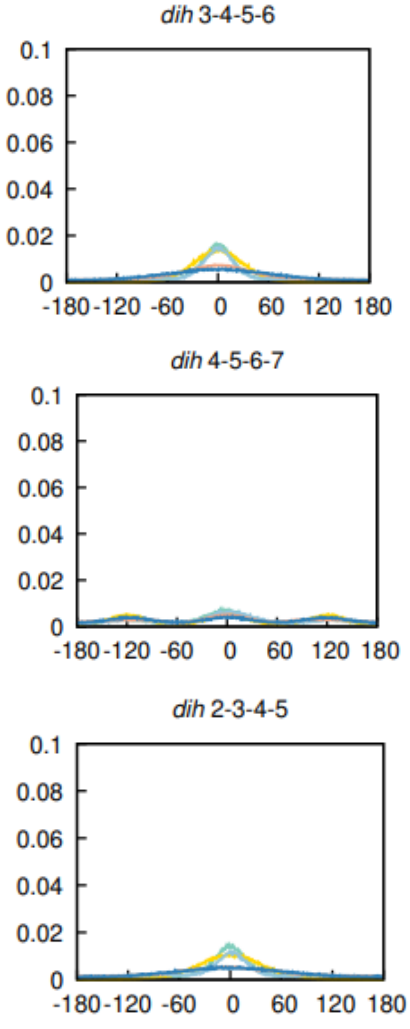
SASA



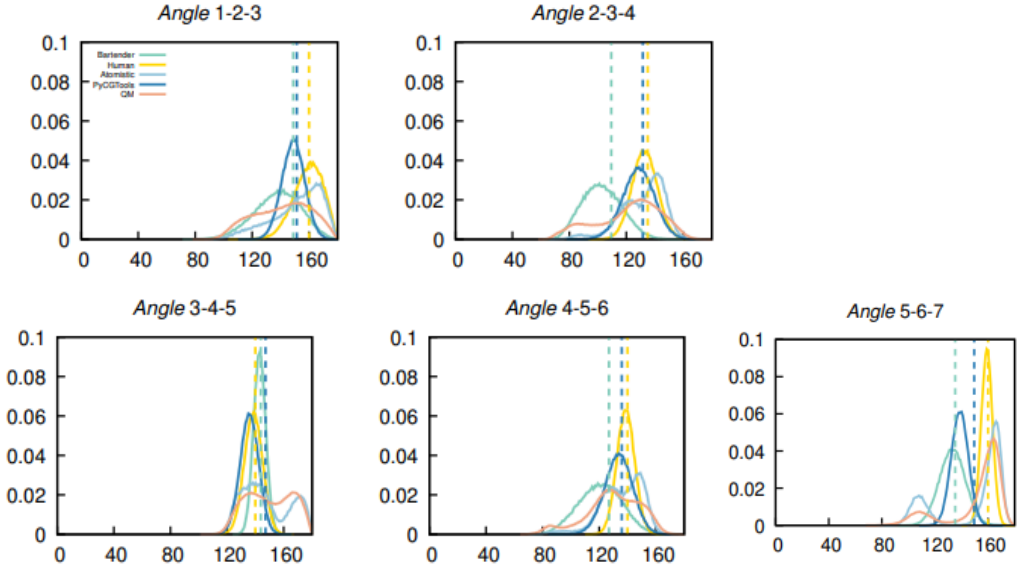
Bonds



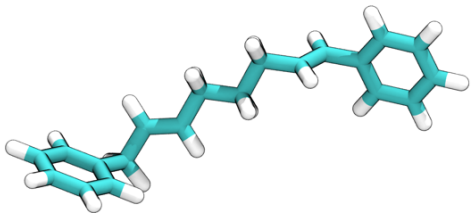
Dihedrals



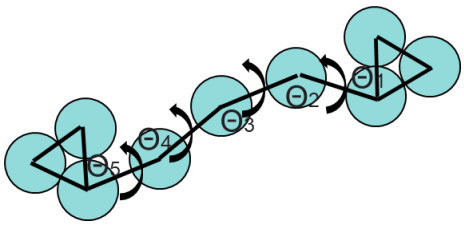
Angles



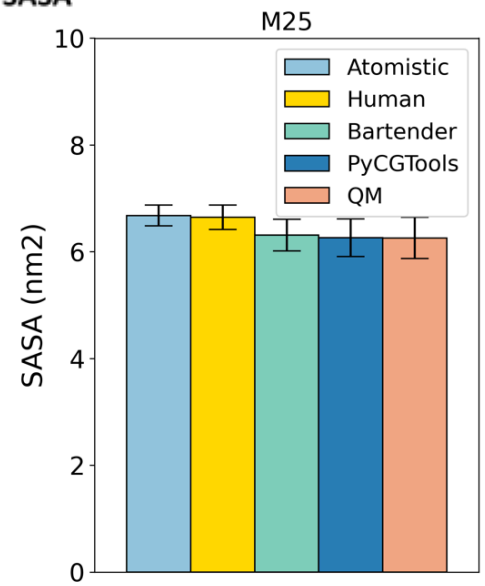
Molecule 5



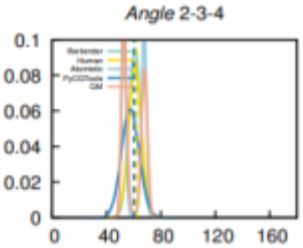
Bonded network

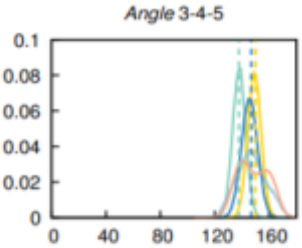


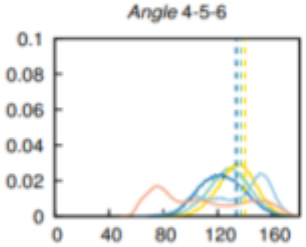
SASA

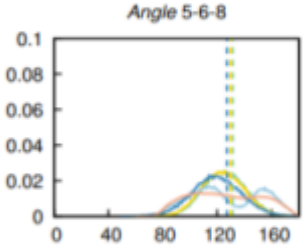


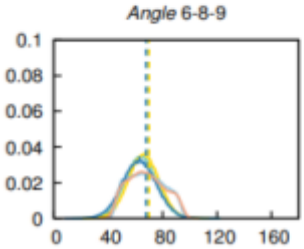
Angles



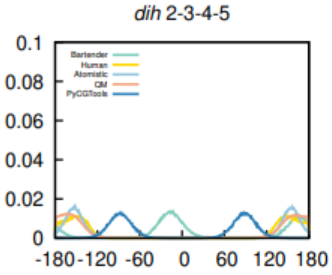


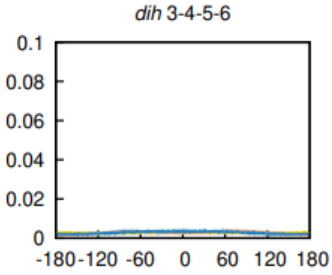


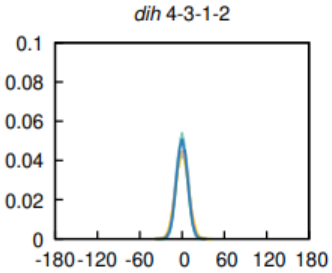


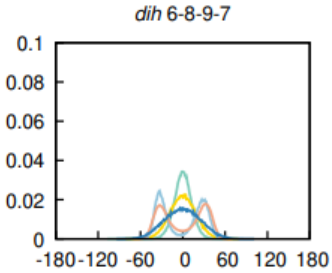


Dihedrals

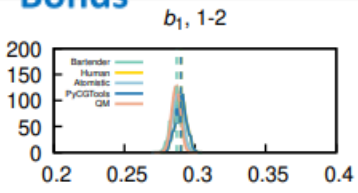


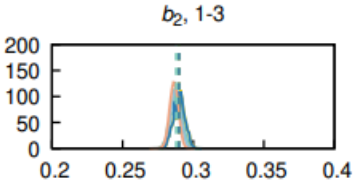


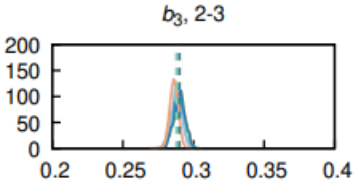


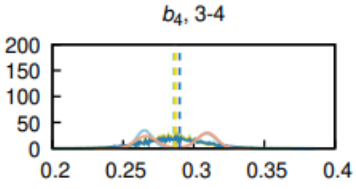


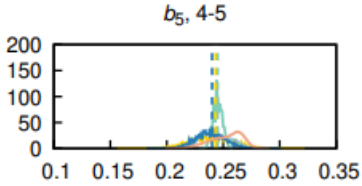
Bonds

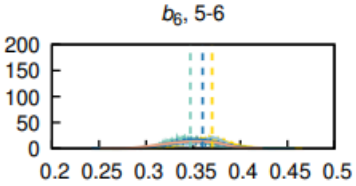


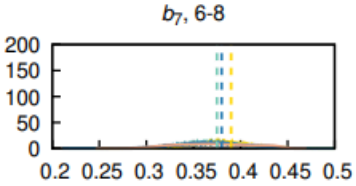


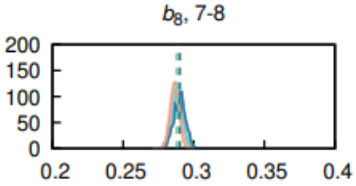


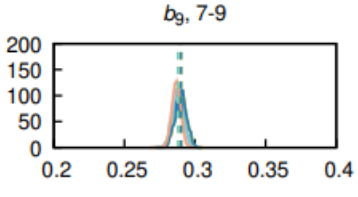


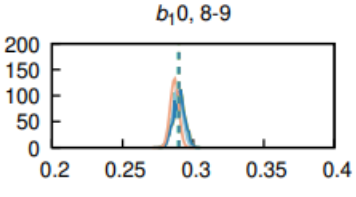




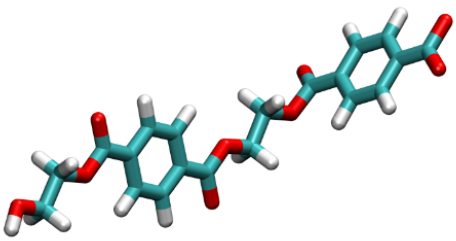




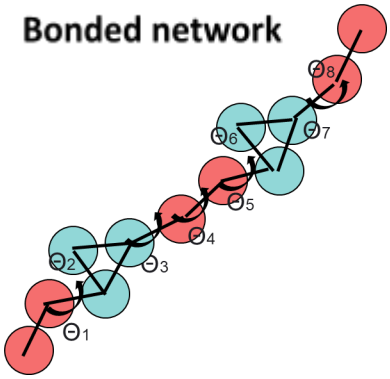




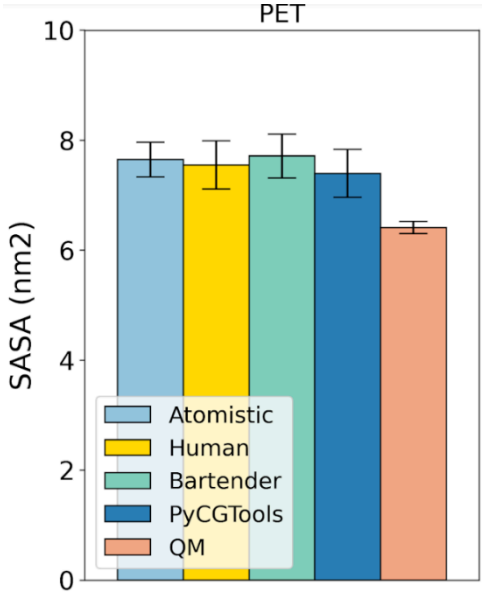
Molecule 6



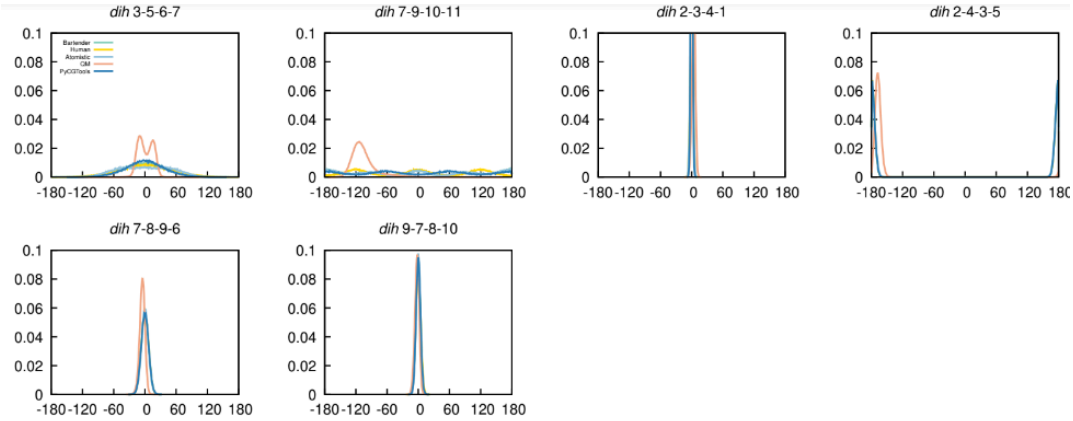
Bonded network



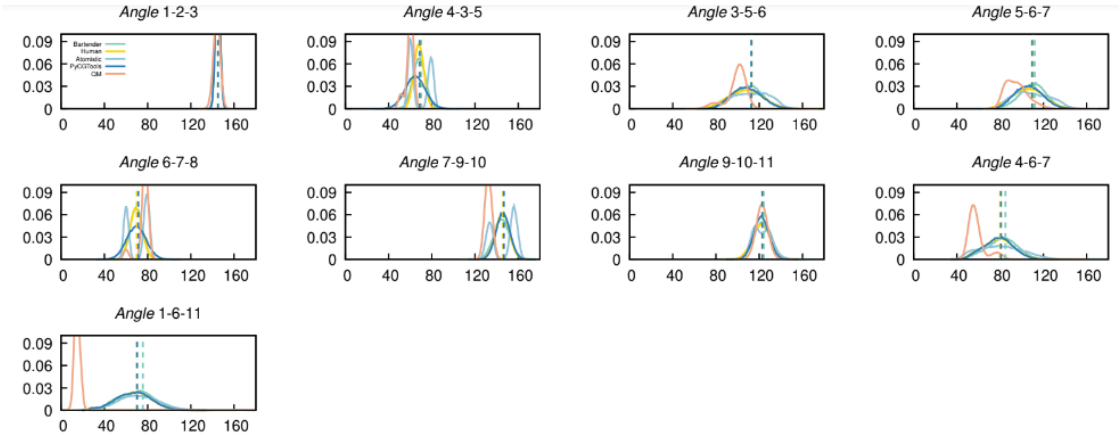
SASA



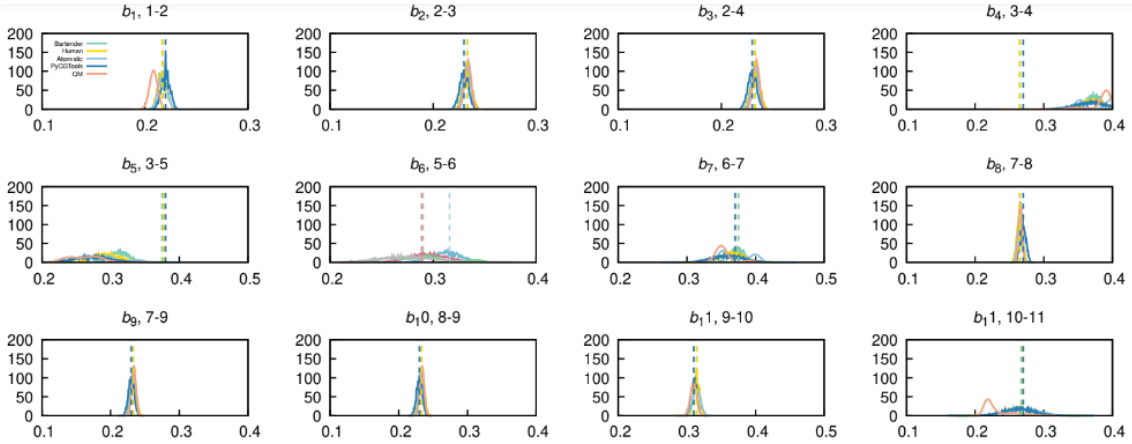
Dihedrals



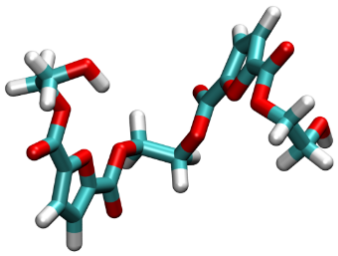
Angles



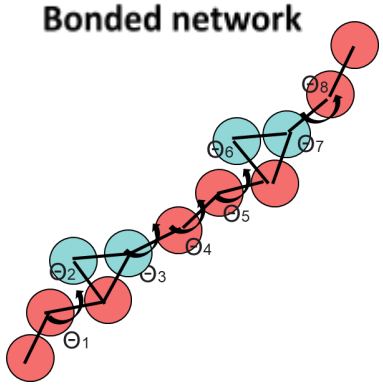
Bonds



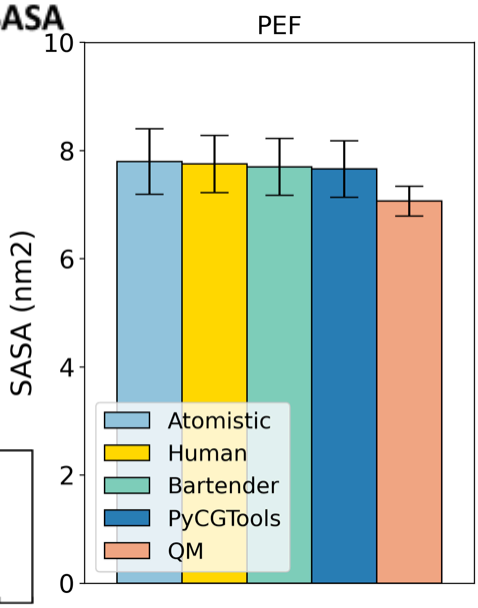
Molecule 7 - PEF



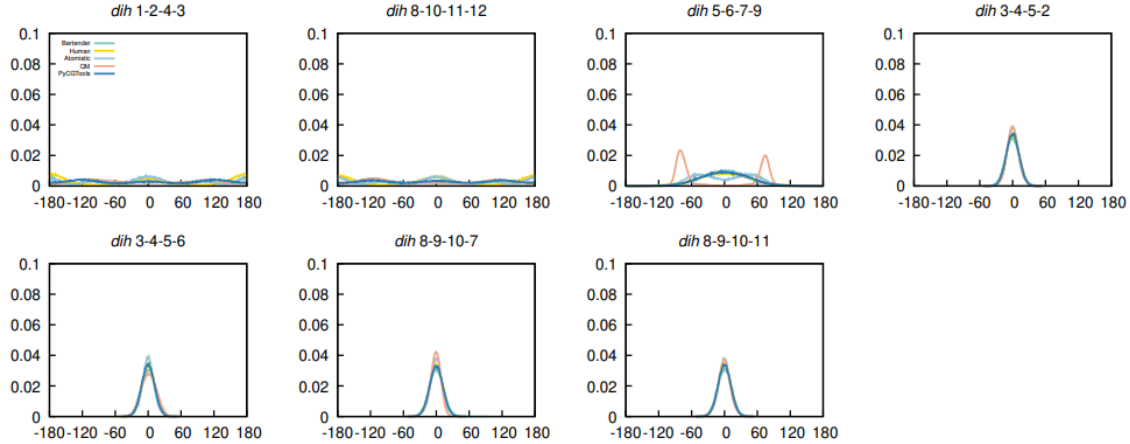
Bonded network



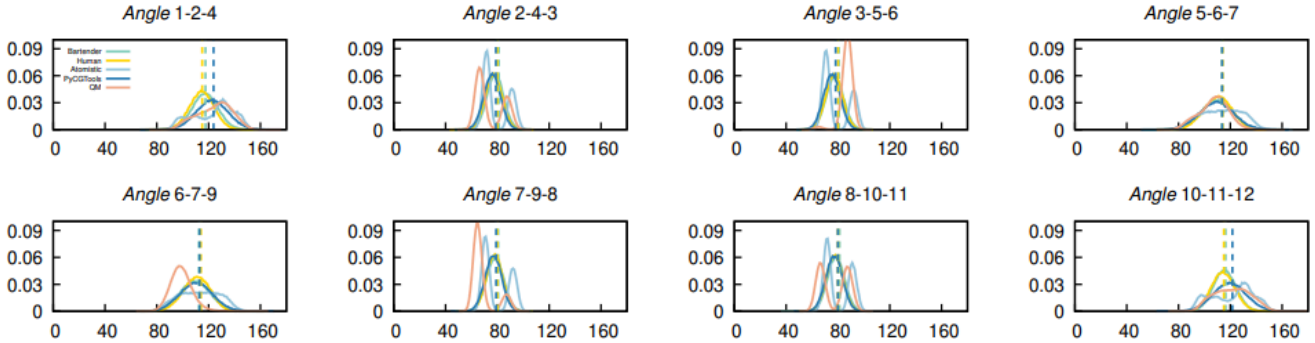
SASA



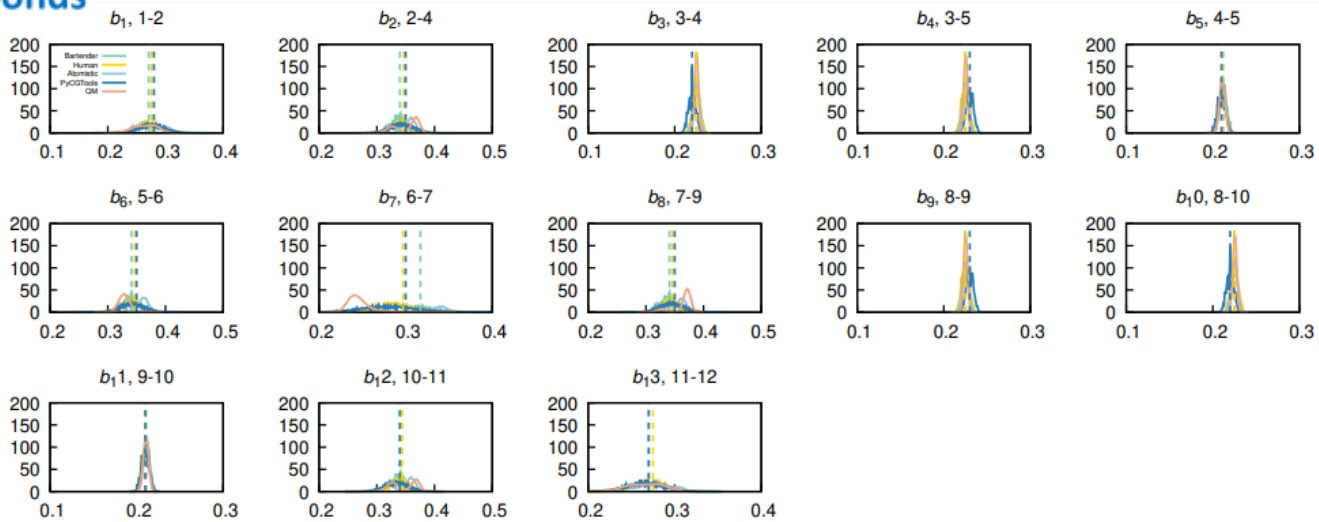
Dihedrals



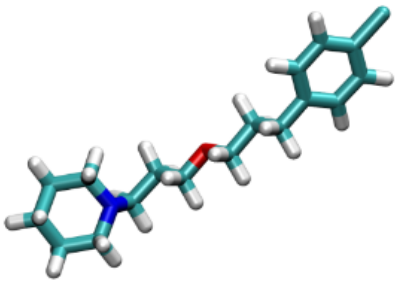
Angles



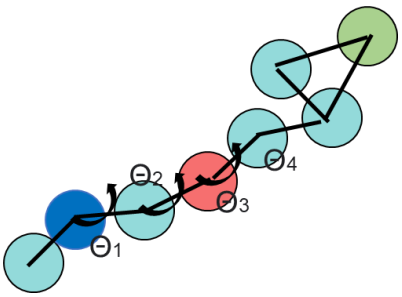
Bonds



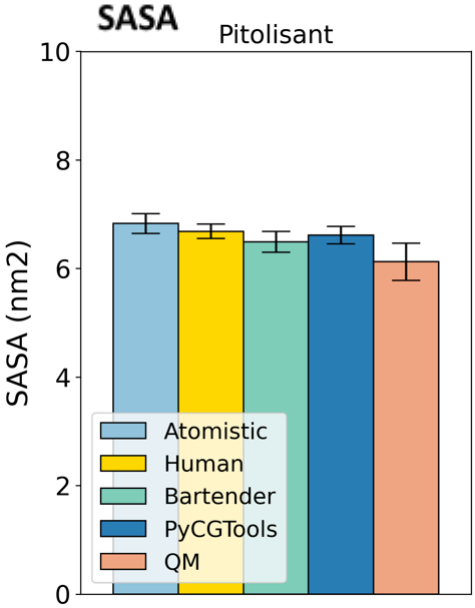
Molecule 8 - Pitolisant



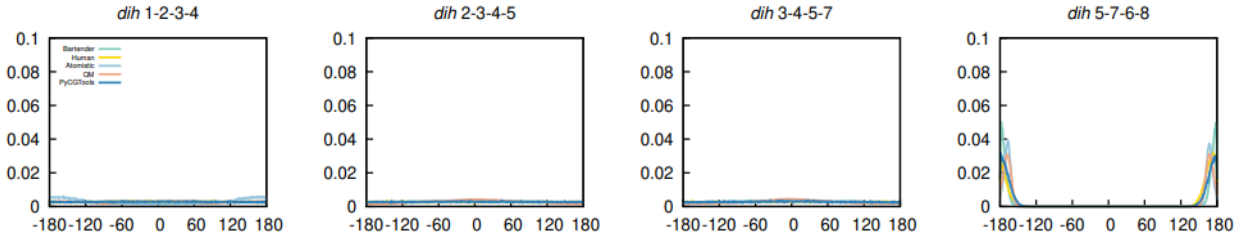
Bonded network



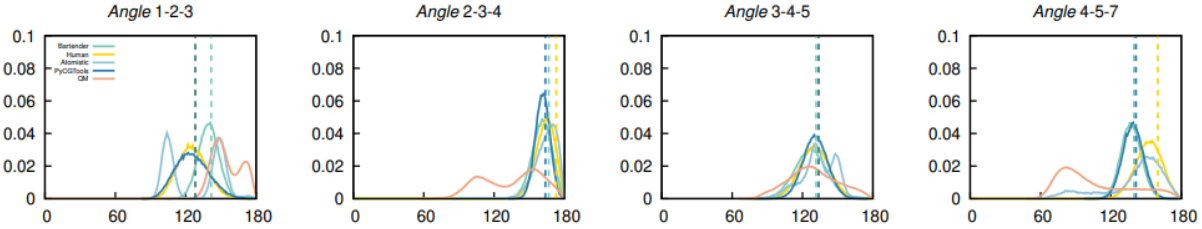
SASA

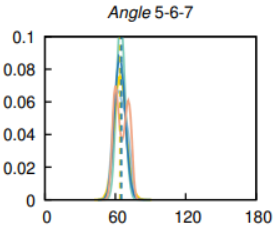


Dihedrals

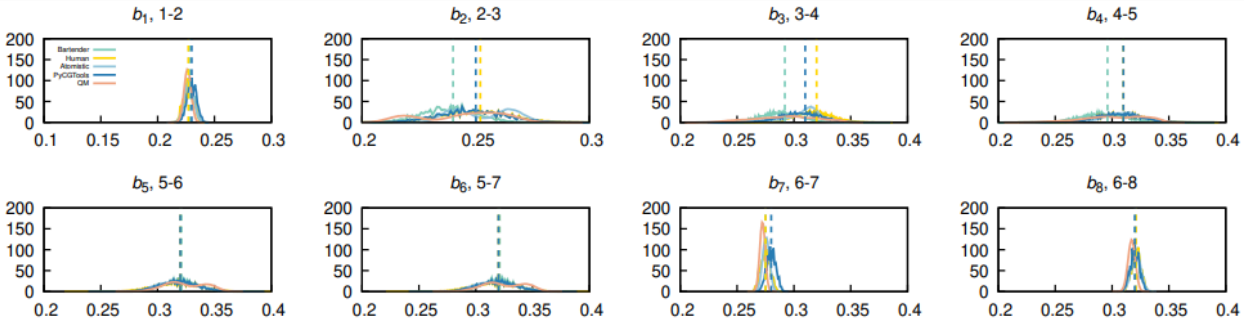


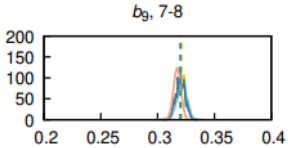
Angles



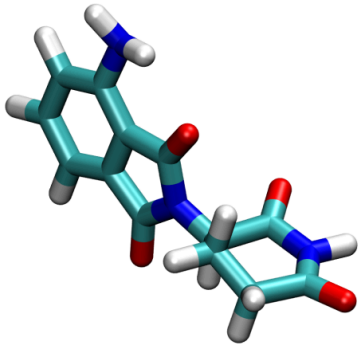


Bonds

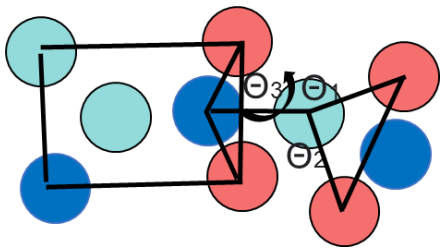




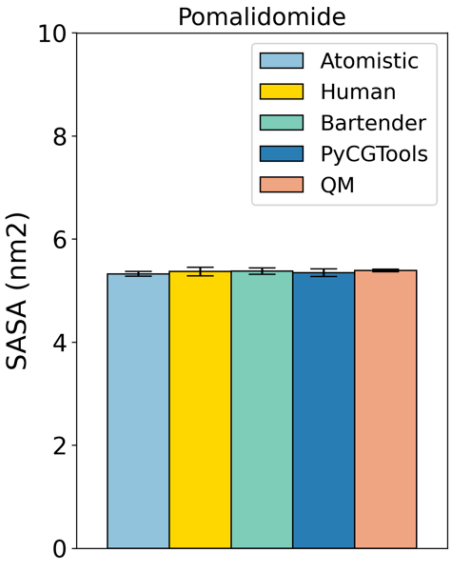
Molecule 9 - Pomalidomide



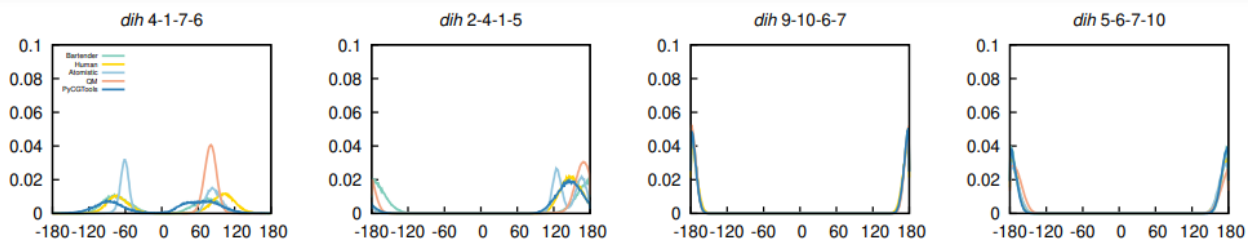
Bonded network



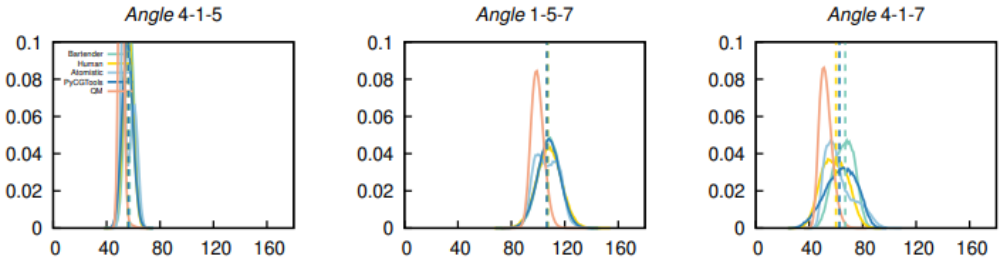
SASA



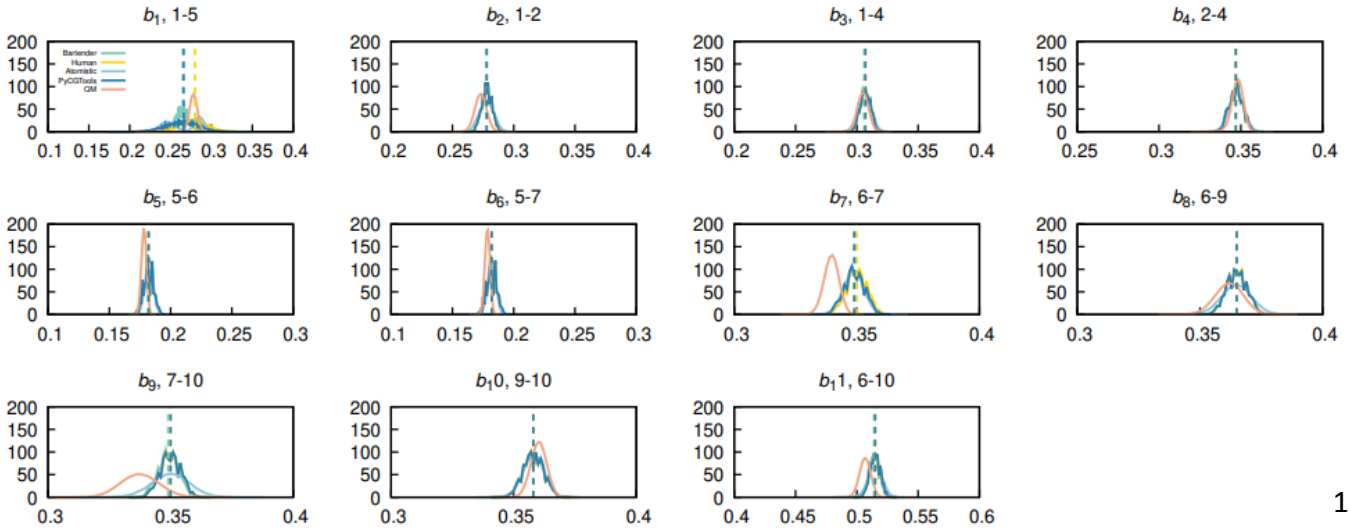
Dihedrals



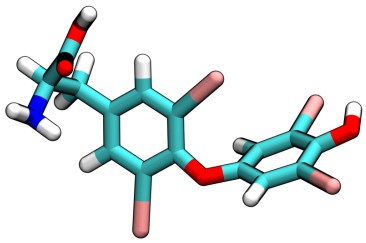
Angles



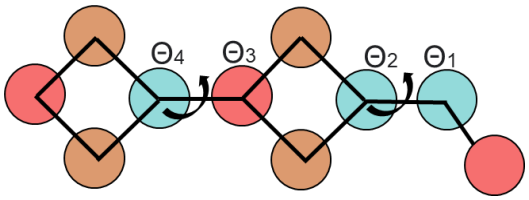
Bonds



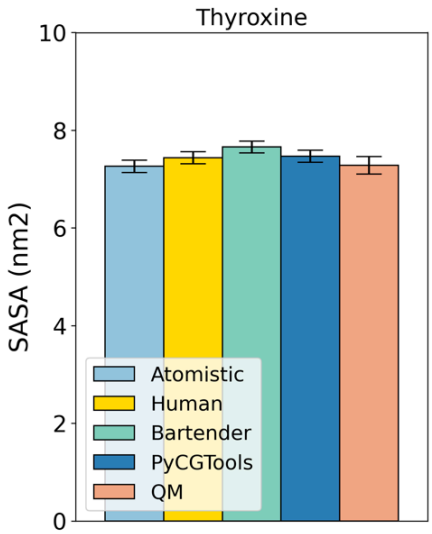
Molecule 20 - Thyroxine



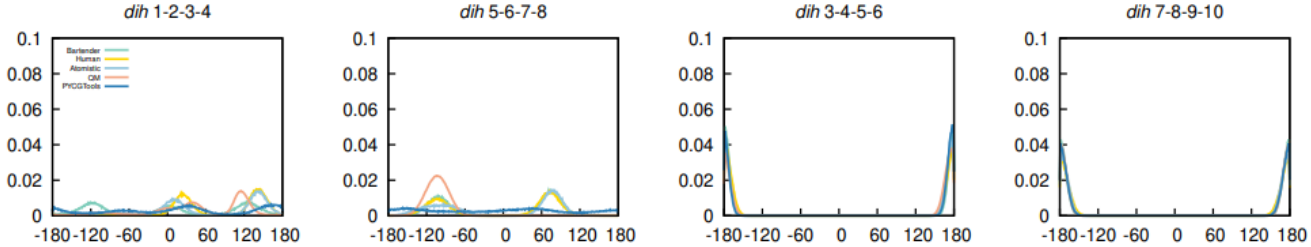
Bonded network



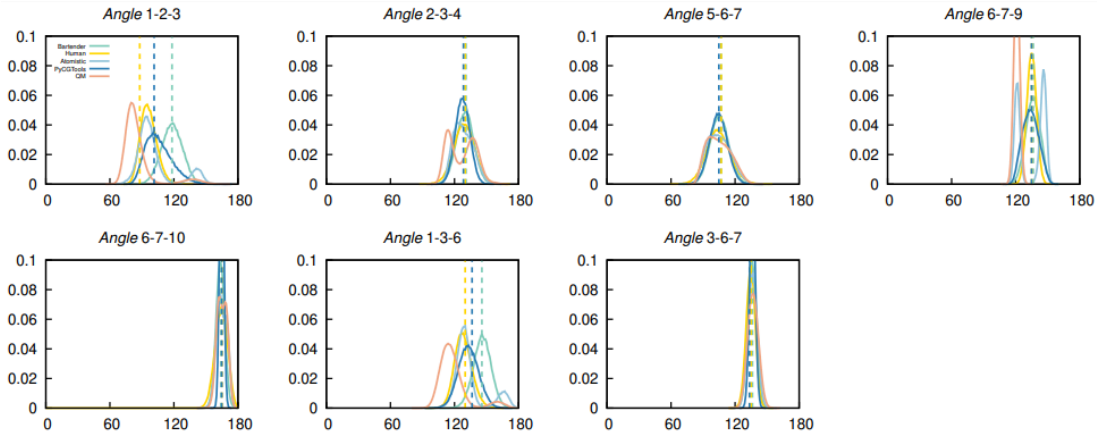
SASA



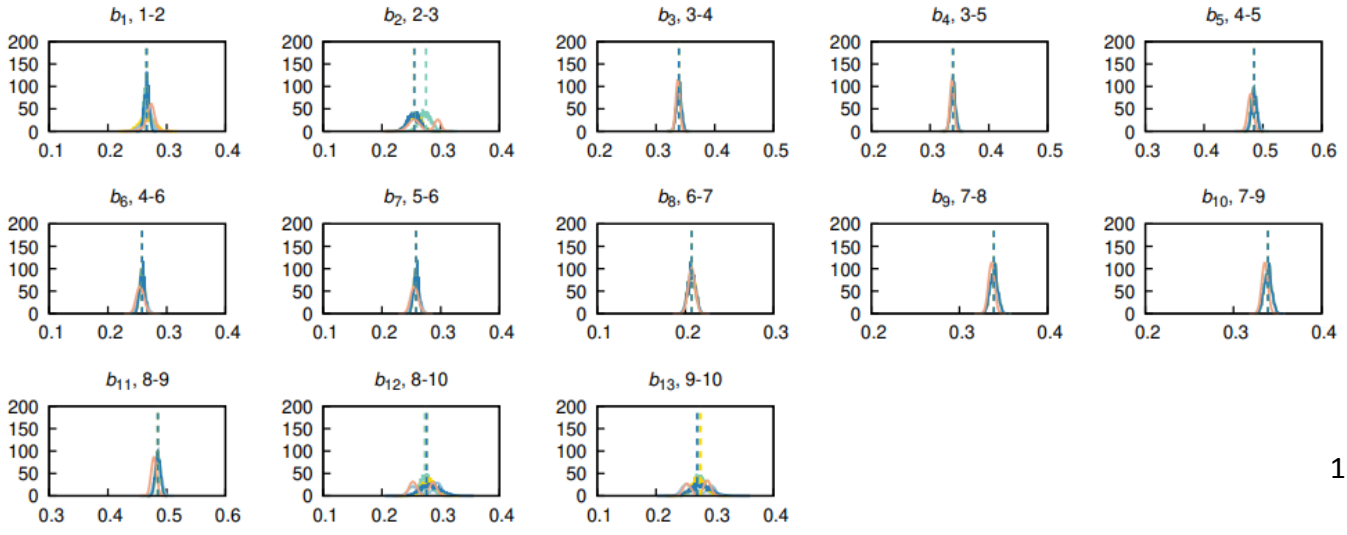
Dihedrals



Angles



Bonds



4. Table S1 - List of molecules from the Martini 3 small molecule dataset considered

uID	Name	SMILES
1MIMI	1-METHYLIMIDAZOLE	Cn1ccnc1
2MPYR	2-METHYL-PYRIDINE	Cc1ccccn1
2NIMX	2-NITRO-m-XYLENE	Cc1c(c(ccc1)C)[N+](=O)[O-]
2NITL	2-NITROTOLUENE	Cc1ccccc1[N+](=O)[O-]
3HT	3-HEXYL-THIOPHENE	CCCCCCc1ccsc1
3PT	3-PROPYL-THIOPHENE	CCCc1ccsc1
4MIMI	4-METHYLIMIDAZOLE	[nH]1cc(nc1)C
4NIAN	4-NITROANISOLE	COc1ccc(cc1)[N+](=O)[O-]
ACPH	ACETOPHENONE	CC(=O)c1ccccc1
ANIL	ANILINE	Nc1ccccc1
BEAL	BENZALDEHYDE	O=Cc1ccccc1
BENZ	BENZENE	c1ccccc1
BRA	4-BROMOANISOLE	COc1ccc(cc1)Br
BRBZ	BROMOBENZENE	c1ccc(cc1)Br
BZDOL	1,3-BENZENEDIOL	Oc1cc(ccc1)O
BZNI	BENZONITRILE	N#Cc1ccccc1
BZQU	para-BENZOQUINONE	c1cc(=O)ccc1=O
BZTF	BENZOTRIFLUORIDE	c1ccc(cc1)C(F)(F)F
CHEX	CYCLOHEXANE	C1CCCCC1
CHEXE	CYCLOHEXENE	C1CCccC1
CLBZ	CHLOROBENZENE	Clc1ccccc1
CLPR	CHLORPROPHAM	CC(C)OC(=O)Nc1cc(ccc1)Cl
CLTL	2-CHLOROTOLUENE	Cc1c(cccc1)Cl
CUME	CUMENE	CC(C)c1ccccc1
CYPE	CYCLOPENTANE	C1CCCC1
CYPO	CYCLOPENTANONE	C1CCC(=O)C1

CYPOL	CYCLOPENTANOL	C1CCC(C1)O
DBRBZ	1,2-DIBROMOBENZENE	c1ccc(c(c1)Br)Br
DCLBZ	1,2-DICHLOROBENZENE	Clc1cccc(Cl)c1
DIOX	1,4-DIOXANE	O1CCOCC1
DMAN	N,N-DIMETHYLANILINE	CN(C)c1ccccc1
DMBZQ	2,5-DIMETHYL-1,4-BENZOQUINONE	Cc1cc(=O)c(cc1=O)C
DXLA	1,3-DIOXOLANE	O1CCOC1
FURA	FURAN	c1ccoc1
IMID	IMIDAZOLE	c1cnc[nH]1
IOBZ	IODOBENZENE	c1ccc(cc1)I
IOPHE	2-IODOPHENOL	c1ccc(c(c1)O)I
MBZOA	METHYL-BENZOATE	COC(=O)c1ccccc1
MCYPE	METHYLCYCLOPENTANE	CC1CCCC1
MESI	MESITYLENE	Cc1cc(cc(c1)C)C
MXYLE	m-XYLENE	Cc1cccc(c1)C
NBAPH	N-BOC-2-AMINOPHENOL	Oc1ccccc1NC(=O)OC(C)(C)(C)
NIBZ	NITROBENZENE	c1ccc(cc1)[N+](=O)[O-]
OMA	ortho-METHYLANISOLE	COc1ccccc1C
OXYLE	o-XYLENE	Cc1c(C)cccc1
PBZOA	PROPYL-BENZOATE	CCCOC(=O)c1ccccc1
PCRE	para-CRESOL	Cc1ccc(O)cc1
PCYM	p-CYMENE	c1cc(ccc1C(C)C)C
PHEN	PHENOL	c1ccc(cc1)O
PIPER	PIPERIDINE	C1CCNCC1
PXYLE	p-XYLENE	Cc1ccc(C)cc1
PYAZ	PYRIDAZINE	c1ccnnc1
PYLI	PYRROLIDINE	C1CCNC1
PYMI	PYRIMIDINE	c1cncnc1

PYRI	PYRIDINE	c1ccncc1
PYRR	PYRROLE	[nH]1cccc1
STYR	STYRENE	c1ccccc1C=C
TCLBZ	1,2,4-TRICHLOROBENZENE	Clc1ccc(Cl)c(Cl)c1
THAZ	THIAZOLE	n1ccsc1
THF	TETRAHYDROFURAN	C1CCOC1
THIO	THIOPHENE	c1ccsc1
THP	TETRAHYDROPYRAN	O1CCCCC1
THPH	THIOPHENOL	Sc1ccccc1
THT	TETRAHYDROTHIOPHENE	S1CCCC1
TOLU	TOLUENE	Cc1ccccc1
XBZ	METHOXYBENZENE	COc1ccccc1
2T	2,2'-BITHIOPHENE	c1cc(sc1)c2cccs2
ANTH	ANTHRACENE	c1ccc2cc3ccccc3cc2c1
BZIM	BENZIMIDAZOLE	c1ccc2c(c1)[nH]cn2
BZTA	BENZOTHIAZOLE	n1c2ccccc2sc1
BZTH	BENZOTHIOPHENE	c1c2ccccc2sc1
CAFF	CAFFEINE	Cn1cnc2n(C)c(=O)n(C)c(=O)c12
CNAP	1-CHLORO-NAPHTHALENE	Clc2cccc1ccccc12
ENAPH	1-ETHYLNAPHTHALENE	CCc1cccc2ccccc21
INDA	INDAZOLE	c2ccc1[nH]ncc1c2
INDO	INDOLE	c12c(ccn2)cccc1
MIND	3-METHYL-1H-INDOLE	c1cccc2c1c(c[nH]2)C
MINDA	1-METHYL-INDAZOLE	Cn1c2ccccc2cn1
MNAP	1-METHYL-NAPHTHALENE	Cc2cccc1ccccc12
NAPH	NAPHTHALENE	c1ccc2ccccc2c1
NAPY	1,5-NAPHTHYRIDINE	c1ccc2ncccc2n1
QUIN	QUINOLINE	n1cccc2ccccc12

TECE	TETRACENE	c34cc2cc1ccccc1cc2cc3cccc4
THPY	THIENO[2,3-c]PYRIDINE	c1cncc2c1ccs2
XNAPH	1-METHOXYNAPHTHALENE	COc1cccc2ccccc21

5. **Table S2 - List of molecules from the flexible drug-like molecule dataset**

Name	SMILES
Molecule 1	CCCC/C=C/C=C/CCCCC/C=C/C
Molecule 2	CCCC/C=C/CC/C=C/CCCC(=O)O
Molecule 3	CCCC/C=C/CCC/C=C/CCCC1CCCCC1
Molecule 4	CCCC/C=C/CCC/C=C/CCCC1CCC(=O)CC1
Molecule 5	C(C/C=C/Cc1ccccc1)C/C=C/c1ccccc1
PET (plastic)	OCCOC(=O)c1ccc(cc1)C(=O)OCCOC(=O)c1ccc(cc1)C(=O)O
PEF (plastic)	OCCOC(=O)c1ccc(o1)C(=O)OCCOC(=O)c1ccc(o1)C(=O)OCCO
Pitolisant	Clc1ccc(CCCOCCCN2CCCCC2)cc1
Pomalidomide	Nc1cccc2c1C(=O)N([C@@H]1CCC(=O)NC1=O)C2=O
Thyroxine	OC(=O)[C@H](N)Cc1cc(I)c(Oc2cc(I)c(O)c(I)c2)c(I)c1

6. Table S3 - Number of parameters used, modified and removed for the overlap analysis

Parameter type	Used directly	Adjusted before analysis	Removed from analysis
Harmonic Bonds	37	6	0
Constraints	44	1	0
Angles	30	2	10
Flexible Dihedrals	12	0	0
Improper Dihedrals	17	2	0

7. Example of the format for Bartender's input file

```
#Lines starting with "#" are comments. Only "full line" comments are allowed!
#note that you can assign half (or less) of an atom to a bead.
#Just divide it by the number you want. Note that all the
#atom and bead numbers are 1-based
BEADS
1 8,9,10,18,19,20,21
2 6/2,11,20
3 6/2,7,17
#you can add comment lines anywhere
4 16,5,22,1,12
5 15,4,3/2,14/2
6 13,2,3/2,14/2
#The virtual sites section is similar to the
#BEADS section, except for the last field
#which is the Gromacs function number. The way the
#virtual site is defined will depend on the function
```

```
#number, and the number of atoms. #This section HAS to go
#after the BEADS section
VSITES
7 1,2,3 1
#The next sections specify the bonded terms you
#want parametrized, in terms of the corresponding beads.
BONDS
1,2
1,3
1,4
4,5
4,6
ANGLES
1,4,5
1,4,6
1,2,4
2,4,5
5,4,2
4,2,1
DIHEDRALS
1,2,4,5
5,4,2,1
IMPROPERS
3,2,1,4
```

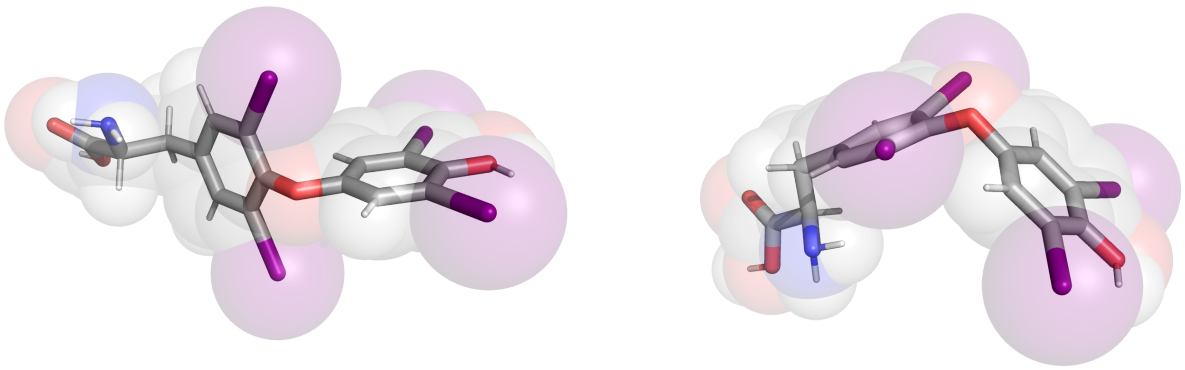
8. Methods and detailed results for the cisoid-transoid interconversion barrier in thyroxine.

Note: All references are given in the main text.

A candidate transition state structure was generated by rotating one of the ether-oxygen -- ring-carbon bonds with the goChem library (www.gochem.org) until the steric clash between the iodines of one ring and the hydrogens of the other was maximized (see Figure S3). The structure was optimized with TURBOMOLE at the r2SCAN-3c level, employing the RI approximation while keeping the rings in position. Numerical frequencies were obtained from the optimized structure, which resulted in one imaginary frequency. The corresponding eigenvector was followed for the transition state optimization, performed at the same level. Numerical, central frequencies were again obtained for the resulting transition state, verifying that it corresponds, not only to a transition state, but to the correct reaction coordinate. Similarly, the minimum structure was obtained by geometry optimization at the

same level of theory with frequency calculations to ensure the minimum character. For both TS and minimum, the Hessian matrices obtained in the frequency calculations were employed for the thermal corrections. On the optimized geometries, SCS-MP2/def2-QZVPP energies were obtained with ORCA. The RI and chain-of-spheres approximations were used for the self-consistent part of the calculations (the chain-of-spheres method was not used in the DFT case), while the RI approximation was also employed in the MP2 part. The Cosmotherm program was employed to obtain solvation free energies for both geometries with the COSMO-RS method at the BP-TZVPD-FINE level.

As reported in the main text, the calculated barriers are 38.5 and 43.6 kJ/mol at 185 and 298 Kelvin, respectively. The result at 185 agrees very well with the reported value of 37 kJ/mol. The barriers calculated have a small entropic component of 0.04 kJ/K*mol, which causes a slightly larger barrier at 298 K. (a very small increase considering the large increase in temperature) The contributions to the barrier are given in table S4.



a **b**

Figure S3: **a.** Optimized geometry for the thyroxine molecule. **b.** Optimized transition state for the cisoid-transoid rotation.

8. Table S4 Components of the transition state barrier ($\Delta G^{\ddagger}$). All values in kJ/mol, except for ΔS, in kJ/molK.

	298 K	185 K
Δ Electronic energy	37.0	37.0
ΔS	-0.04	-0.04
ΔG_{therm}	10.3	5.4
ΔG_{solv}	-3.7	-3.9
$\Delta G^{\ddagger}$	38.5	43.6

9. Optimized geometries for the minimum and the transition state for the ring-ring rotation of Thyroxine, in XYZ format.

35

Minimum

O -1.6456854 -0.7581681 -1.3946582

C -1.0190031 -1.2333398 -0.2924119

O -1.5160142 -2.0497454 0.4452948

C 0.3791349 -0.6600930 -0.1400895

N 0.8838957 -0.9970803 1.1827046

C 0.4037275 0.8723741 -0.3563193

C 1.8115149 1.3943827 -0.2564151

C 2.6229925 1.4407356 -1.3913916

C 2.3441358 1.7681942 0.9788180

C 3.6718874 2.1638859 1.0737605

C 4.4925434 2.2047007 -0.0547772

C 3.9475018 1.8445175 -1.2879531

I 5.1588611 1.9165939 -3.0101028

I 4.4682130 2.6927915 2.9510509

O 5.8270127 2.4997014 0.0552589

C 6.2215950 3.8237487 -0.0089993

C 7.5782394 4.0524560 0.1792022

C 5.3432735 4.8723336 -0.2482090

C 5.8408282 6.1676221 -0.2945041

C 7.1984279 6.4478998 -0.1100582

C 8.0436117 5.3552489 0.1247391

O 7.6174679 7.7312753 -0.1664476

I 10.1100781 5.7189396 0.4122712

I 4.5088860 7.7592640 -0.6516768

H -2.5058507 -1.2052216 -1.4386918

H 0.9570713 -1.1157869 -0.9701497

H 1.8956802 -1.0391262 1.1775380

H 0.5268891 -1.9035409 1.4654401

H -0.2273258 1.3348852 0.4101770

H	-0.0150300	1.1075694	-1.3387315
H	2.2197420	1.1596479	-2.3597547
H	1.7226034	1.7257661	1.8667027
H	8.2416073	3.2161489	0.3652213
H	4.2850707	4.6878709	-0.3946630
H	8.5766120	7.7584940	-0.0179300

35

Transition state

O	0.5751359	-1.2141321	-4.0217114
C	0.2295020	-0.8103182	-2.7743249
O	-0.8314404	-1.0813310	-2.2682691
C	1.3042594	0.0650789	-2.1486591
N	2.6390166	-0.4584146	-2.4433673
C	1.0865261	0.1557770	-0.6307222
C	2.0381812	1.1401282	-0.0091054
C	1.7108958	2.4935838	0.0562220
C	3.2626272	0.7264061	0.5117623
C	4.1686499	1.6582331	1.0056976
C	3.9056782	3.0331797	0.9611417
C	2.6205156	3.4172381	0.5542588
I	1.9075361	5.3836391	0.8614924
I	5.8682285	0.8915072	2.0042830
O	4.8275748	3.9598447	1.4061753
C	5.6967748	4.4857313	0.4353040
C	6.9216561	3.8728481	0.2034456
C	5.3681812	5.6656999	-0.2221752
C	6.2542224	6.2093623	-1.1390819
C	7.4872011	5.6086358	-1.4192593
C	7.7908201	4.4367008	-0.7171046
O	8.3049317	6.1823522	-2.3236121

I	9.6596487	3.5015106	-1.0723684
I	5.7359246	7.9949751	-2.1298669
H	-0.1674970	-1.7379033	-4.3620901
H	1.1213763	1.0739691	-2.5776691
H	3.3353000	0.2104349	-2.1292951
H	2.7611401	-0.5628708	-3.4444441
H	1.2300714	-0.8417818	-0.2025932
H	0.0510203	0.4539770	-0.4441350
H	0.7279285	2.8274185	-0.2617436
H	3.5039273	-0.3312350	0.5422302
H	7.1934349	2.9778190	0.7456754
H	4.4310955	6.1614778	-0.0098543
H	9.1169761	5.6540863	-2.3984306