Supporting Information

Hydrocarbon Macrocycle Conformer Ensembles and $^{13}$C-NMR Spectra

F. Bohle, S. Grimme*
SUPPORTING INFORMATION

Contents

1 Linear regression scaling factors for obtaining computed chemical shifts S3
   1.1 $^1$H-NMR Shifts .......................................................... S5
   1.2 $^{13}$C-NMR Shifts .......................................................... S6

2 Cycloalkane chemical shifts S7
   2.1 $^{13}$C-chemical shifts .................................................. S7
   2.2 $^1$H-chemical shifts ................................................... S8

3 Number of conformers in each crenso step S10

4 Timings S11

5 Influence of symmetry on final chemical shifts S12

6 Technical settings applied in crenso S13

7 Ensemble strain energy S14

*Mulliken Center for Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstr. 4, 53115, Bonn, Germany

†Corresponding author: grimme@thch.uni-bonn.de (phone: +49-228/732544)
8 Cycloalkane discussion with respect to experimental or theoretical reference

8.1 Cyclobutane ........................................ S15
8.2 Cyclopentane ........................................ S15
8.3 Cyclohexane .......................................... S15
8.4 Cycloheptane ........................................ S15
8.5 Cyclooctane .......................................... S16
8.6 Cyclononane ......................................... S16
8.7 Cyclodecane .......................................... S16
8.8 Cycloundecane ...................................... S17
8.9 Cyclododecane ...................................... S17
8.10 Cyclotridecane ..................................... S17
8.11 Cyclotetradecane .................................. S18
8.12 Cyclopentadecane ................................ S18
8.13 Cyclohexadecane .................................. S19
8.14 Cycloheptadecane ................................ S19
8.15 Cyclooctadecane .................................. S20
8.16 Cyclononadecane .................................. S20
8.17 Cycloicosane ...................................... S21

9 Application examples ................................ S23
9.1 Application example I: cyclic peptide PGLGF ........................................ S23
9.2 Application example II: lobophytolins A .............................................. S27
9.3 Application example III: (-)-dactylolide .............................................. S30

10 References ............................................ S34
1 Linear regression scaling factors for obtaining computed chemical shifts

Chemical shifts ($\delta_{DFT,corr}$) are obtained by subtracting computed DFT chemical shielding constants ($\sigma_{DFT,raw}$) from a fitted reference chemical shielding constant ($\sigma_{fitted,reference}$) (intercept) and applying a scaling factor (scaling_factor) (slope), see eq. 1. The scaling factor and reference chemical shielding constant are obtained from a linear regression fit between computed chemical shielding constants ($\sigma_{DFT,raw}$) and experimentally measured chemical shifts ($\delta_{exp}$). The linear regression correction is applied to reduce systematic errors between theory and the experiment which are, for example, the result of errors in the DFT treatment, solvation or rovibrational effects. The intercept from the linear regression represents the reference isotropic shielding value of a common standard, typically TMS.

$$\delta_{DFT,corr} = \text{scaling_factor} \times (\sigma_{fitted,reference} - \sigma_{DFT,raw}) \quad (1)$$

For the linear regression, a database of small organic molecules compiled with experimentally determined $^1$H and $^{13}$C data is taken from the supporting information in Ref. [1], which can also be found at the chemical shift repository[2]. The database is divided into the TESTSET and a smaller PROBESET for cross validation. The employed database containing small organic molecules and experimental references was first developed by Rablen[3] and augmented by Lodewyk, Siebert and Tantillo[1].

Following molecules were used in the the TESTSET:
- 1,1,1-trichlorethane,
- 1,1,1-trichloroacetone,
- 1,1-dichloroethylene,
- 1-chloro-2-methylpropene,
- 2-butyne,
- 2-chloro-2-methylpropane,
- 2-chloropropane,
- 2-chloropropionitrile,
- 2-cyclohexenone,
- 2-cyclopentenone,
- 2-methyl-2-butene,
- 2-methyl-2-nitropropane,
- 2-methylpropanenitrile,
- 2-methylpropene,
- 3,3-dimethyl-1-butene,
- 3,4-dichloro-2,5h-furanone,
- 3-butene-2-one,
- 3-chloropropene,
- 4-methylthiazole,
- acetaldehyde,
- acetone,
- acetonitrile,
- acetylenechloride,
- benzene,
- butyrolactone,
- chlorobenzene,
- chlorodimethylsilane,
- chloroethane,
- cyclobutane,
- cyclobutanone,
- cyclobutenone,
- cyclohexanone,
- cyclopentane,
- cyclopentanone,
- cyclopentene,
- cyclopropane,
- cyclopropanone,
- dimethylsulfoxide,
- fluorobenzene,
- furan,
- furfural,
- isoxazole,
- methanol,
- methylacetate,
- methylerter-butylether,
- nitrobenzene,
- nitroethane,
- nitromethane,
- N-methylpyridine,
- N-methylpyrrole,
- N-methylpyrrolidine,
- N,N-dimethylformamide,
- oxytane,
- oxirane,
- p-benzoquinone,
- propionitrile,
- pyrazine,
- pyridazine,
- pyridine,
- pyrimidin,
- tetrahydrofuran,
- tetrahydropropyran,
- thiirane,
- thiophene,
- trans-1,2-dichloroethylene,
- trichloroethylene,
- trimethylacetonitrile.

The following molecules were used in the the PROBESET:
- 1,3,5-triazine,
- 1,3-dioxane,
- 2,5-dihydrofuran,
- 2-methylpyrazine,
- 2-methylthiophene,
- allylthiocyadate,
- crotonaldehyde,
- diketene,
- dimethylacetel,
- dimethylcyanamide,
- E-3,4-dimethyl-2-pentene,
- ethylycyanofomrate,
- ethylene carbonate,
- ethylsulfide,
- fumaronitrile,
- isoprene,
- maleicenydride,
- malononitrile,
- methylenecyclbutane,
- methylisothiocyanate,
- pivalaldehyde,
- pyridine,
- quadricyclane,
- tert-butylisocyanide.

The computed chemical shielding constants of the molecules in the TESTSET and PROBE-
SET were obtained in the following way: Conformers, rotamers, chemical- and magnetical equivalencies were determined in a crest run with implicit solvation (CHCl₃). The resulting conformer rotamer ensembles (CRE) are sorted by censo while improving both the description of free energy (G) and the molecular solution phase structure. Shielding constants are calculated for the dominantly populated conformers. Through consideration of the chemical equivalencies (identified during the conformer generation in CREST) and Boltzmann factors the shielding constants are properly thermally averaged over the CRE. The following method combinations were applied in the censo calculations:

Geometry: r²SCAN-3c[DCOSMO-RS(CHCl₃)]
Solvent: CHCl₃
T in K: 298.15
Electronic Energy ΔE: r²SCAN-3c
Thermal corrections to ΔG_{RRHO} : GFN2-xTB[GBSA(CHCl₃)]
Solvation free Energy (ΔδG_{solv}): COSMO-RS Version 19 (Param: ctd = BP_TZVP_19.ctd)
Chemical shielding constants (σ_{DFT,raw}): PBE0/def2-TZVP[DCOSMO-RS(CHCl₃)]

The linear regression parameters are calculated from the TESTSET and analyzed using the PROBESET. Linear regression parameters for $^1$H and $^{13}$C for the method combination of PBE0/def2-TZVP[DCOSMO-RS(CHCl₃)//r²SCAN-3c[DCOSMO-RS(CHCl₃)] are evaluated below.
## 1.1 $^1$H-NMR Shifts

### $^1$H-NMR Shifts

**TESTSET (68 molecules)**

- $R^2$: 0.9975
- Scaling factor: 0.933
- Fitted reference: 31.686
- Unscaled RMSD ($\delta$): 0.387
- Scaled RMSD ($\delta$): 0.129

**PROBESET (24 molecules)**

- $R^2$: 0.9963
- Scaling factor: 0.945
- Fitted reference: 31.690
- Unscaled RMSD ($\delta$): 0.327
- Scaled RMSD ($\delta$): 0.158
1.2 \(^{13}\text{C-} \text{NMR Shifts}\)

**\(^{13}\text{C-} \text{NMR Shifts}\)**

**TESTSET (68 molecules)**

- R\(^2\): 0.9427
- Scaling factor: 0.942
- Fitted reference: 186.156
- (using TESTSET-scaling factor): unscaled RMSD (\(\delta\)): 7.948
- (and fitted reference) scaled RMSD (\(\delta\)): 4.217

**PROBESET (24 molecules)**

- R\(^2\): 0.9386
- Scaling factor: 0.938
- Fitted reference: 187.106
- (using TESTSET-scaling factor): unscaled RMSD (\(\delta\)): 6.637
- (and fitted reference) scaled RMSD (\(\delta\)): 2.215

The data points clearly deviating from the fitted line stem from molecules with chlorine atoms.
2 Cycloalkane chemical shifts

2.1 $^{13}$C-chemical shifts

Table S 1: $^{13}$C chemical shifts of the investigated cycloalkanes in CHCl$_3$. a) cyclononanone chemical shift extrapolated from cyclooctane and cyclodecane, using references [4, 5]; b) cycloheptadecane chemical shift extrapolated from cyclohexadecane and cyclooctadecane using references [4]; c) unscaled theoretical chemical shifts obtained with tetramethylsilane (TMS) as standard; d) theoretical values by applying linear regression correction ($\delta_{DFT,corr}$) (used in the main text). All computed chemical shifts given in ppm are evaluated at 298.15 K.

| ringsize (# of -(CH$_2$)-) | experiment | theory unscaled$^c$ | theory scaled$^d$ |
|---------------------------|------------|---------------------|-------------------|
| 4                         | 22.4 [6]   | 25.9                | 22.3              |
| 5                         | 25.8 [6]   | 30.0                | 26.3              |
| 6                         | 27.0 [4]   | 31.1                | 27.2              |
| 7                         | 28.5 [5, 7]| 32.3                | 28.3              |
| 8                         | 26.8 [4]   | 30.6                | 26.7              |
| 9                         | 25.9 $^a$  | 29.2                | 25.4              |
| 10                        | 25.1 [4]   | 28.6                | 24.9              |
| 11                        | 26.3 [4]   | 30.2                | 26.4              |
| 12                        | 23.8 [4]   | 27.1                | 23.5              |
| 13                        | 26.2 [4]   | 30.0                | 26.2              |
| 14                        | 25.2 [4]   | 28.7                | 25.0              |
| 15                        | 27.0 [4]   | 30.5                | 26.6              |
| 16                        | 26.9 [4]   | 30.3                | 26.5              |
| 17                        | 27.2 $^b$  | 31.3                | 27.5              |
| 18                        | 27.5 [4]   | 31.3                | 27.5              |
| 19                        | 28.1 [4]   | 32.2                | 28.3              |
| 20                        | 28.0 [4]   | 32.0                | 28.1              |
### 2.2 $^1H$-chemical shifts

Table S 2: $^1H$ chemical shifts of the investigated cycloalkanes in CHCl$_3$. a) chemical shift measured in neat; b) chemical shift measured in CHCl$_3$ at 38 °C [4, 5]; c) unscaled theoretical chemical shifts obtained with tetramethylsilane (TMS) as theoretical standard; d) scaled theoretical shifts by applying linear regression correction ($\delta_{DFT,corr}$). All computed chemical shifts given in ppm are evaluated at 298.15 K.

| ringsize (\# of -(CH$_2$)-) | experiment | theory unscaled$^c$ | theory scaled$^d$ |
|-----------------------------|------------|----------------------|-------------------|
| 4                           | 1.980$^a$  | 1.976                | 2.021             |
| 5                           | 1.506$^a$  | 1.586                | 1.632             |
| 6                           | 1.431$^b$  | 1.466                | 1.512             |
| 7                           | 1.526$^a$  | 1.566                | 1.612             |
| 8                           | 1.531$^a$  | 1.566                | 1.612             |
| 9                           | 1.527$^a$  | 1.536                | 1.582             |
| 10                          | 1.514$^b$  | 1.566                | 1.612             |
| 11                          | 1.441$^b$  | 1.476                | 1.522             |
| 12                          | 1.344$^b$  | 1.386                | 1.432             |
| 13                          | 1.348$^b$  | 1.356                | 1.403             |
| 14                          | 1.331$^b$  | 1.376                | 1.422             |
| 15                          | 1.333$^b$  | 1.356                | 1.403             |
| 16                          | 1.319$^b$  | 1.346                | 1.393             |
| 17                          | 1.316$^a$  | 1.356                | 1.403             |
| 18                          | 1.310$^b$  | 1.346                | 1.393             |
| 19                          | 1.307$^b$  | 1.346                | 1.393             |
| 20                          | 1.301$^b$  | 1.336                | 1.383             |
Figure S 1: Computed and experimental $^1$H chemical shifts of the investigated cycloalkanes in CHCl$_3$. Information on experimental data is provided in Table 2. Blue points represent the experimental measurements, orange crosses the theoretically derived proton chemical shifts with TMS as reference and computed chemical shifts corrected by linear scaling are shown in green.
3 Number of conformers in each crenso step

Figure S 2: Number of conformer candidates remaining after each of the 'crenso steps': crest_combi (conformer ensemble generation), CENSO:part0 (cheap pre-screening), CENSO:part1 (pre-screening), CENSO:part2 (optimization and free energy (G) evaluation).
4 Timings

Figure S 3: Timings of the sorting steps: crest_combi, CENSO:part0, CENSO:part1, CENSO:part2, CENSO:part4. Evaluated at Intel(R) Xeon(R) CPU E5-2660 v4 @ 2.00GHz using 28 cores (see Table 3 for employed threshold-settings).
5 Influence of symmetry on final chemical shifts

Figure S 4: Calculated chemical shifts when symmetry in the thermostastistical $G_{mRRHO}^{298.15K}$ treatment is considered or ignored, i.e., the influence of the Boltzmann factor on the ensemble averaged property.
## 6 Technical settings applied in crenso

Table S 3: Technical settings as applied in the crenso calculation using (`crenso -nmr CHCl₃ -l3`), but with slightly increased thresholds for evaluation purposes.

|                                    | CENSO general information:                                    |
|------------------------------------|----------------------------------------------------------------|
| CENSO version                      | 1.0.8                                                          |
| Temperature                        | 298.15 K                                                       |
| Solvent                            | CHCl₃                                                          |

### PART0 - cheap pre-screening

| Energy                             | b97-d/def2-SV(P) + GCP                                         |
|------------------------------------|----------------------------------------------------------------|
| Energy_settings                    | scfconv 5, grid 1 using D3(0) instead of D3(BJ)                |
| Gₘ RRHO                            | not included                                                   |
| Gₘ solv                            | ALPB[gn2]                                                      |
| Geometry                           | GFNₓ-xTB (input geometry)                                      |
| Threshold                          | 6.0 kcal/mol                                                   |
| main QM code                        | TM                                                             |

### PART1 - pre-screening

| Energy                             | r2scan-3c                                                      |
|------------------------------------|----------------------------------------------------------------|
| Energy_settings                    | grid m4 scfconv 6                                              |
| Gₘ RRHO                            | GFN2[alpb]-bhess SPH                                           |
| Gₘ solv                            | COSMO-RS-normal[r2scan-3c] param= 19-normal                     |
| Geometry                           | GFNₓ-xTB (input geometry)                                      |
| Threshold                          | 4.5 kcal/mol                                                   |
| main QM code                        | TM                                                             |

### PART2 - optimization

| Energy                             | r2scan-3c                                                      |
|------------------------------------|----------------------------------------------------------------|
| Energy_settings                    | grid m4 scfconv 6                                              |
| Gₘ RRHO                            | GFN2[alpb]-bhess SPH                                           |
| Gₘ solv                            | COSMO-RS-normal[r2scan-3c] param= 19-normal                     |
| Geometry                           | r2scan-3c[DCOSMORS] @optlevel: normal using grid m4 scfconv 6   |
| Threshold                          | Opt_limit: 2.5 kcal/mol, Boltzmann sum threshold: 95.0 %        |
| main QM code                        | TM                                                             |

### PART4 - NMR mode

| Energy                             | using Energy from part2 - optimization                         |
| Energy_settings                    | see part2 - optimization                                       |
| Gₘ RRHO                            | using Gₘ RRHO from part2 - optimization                        |
| Gₘ solv                            | using Gₘ solv from part2 - optimization                        |
| Geometry                           | r2scan-3c[DCOSMORS] @optlevel: normal                           |
| Threshold                          | Boltzmann sum threshold: 95.0 %                                 |
| main QM code                        | TM                                                             |
| NMR shielding constants            | pbe0/def2-TZVP[DCOSMO-RS] (TM)                                  |
|                                   | using grid m5 scfconv 7                                        |
| NMR coupling constants             | pbe0/def2-TZVP[DCOSMO-RS] (TM)                                  |
|                                   | using grid m5 scfconv 7                                        |
7 Ensemble strain energy

Figure S 5: Ensemble averaged strain free energy, calculated with implicit CHCl$_3$ at 298.15 K as a function of the cycloalkane ring size.

The ensemble averaged Gibbs strain energy is calculated using the ensemble free energies of the cycloalkanes (c.f. Table S3(part2)). Methylyene is used as a computational group equivalent (pioneered by Wiberg[9] and Schleyer[10]) and the free energy of the methylene group equivalent is obtained from Eq. 2:

\[ G(CH_2) = \frac{(G_{ensemble}^{n-heptane} - G_{ensemble}^{n-propane})}{4} \]  

The ensemble averaged Gibbs strain energy is calculated from Eq. 3 and this approach is adapted from Ref. [11].

\[ \Delta G_{strain} = G_{ensemble}^{C_xH_{2x}} - x \cdot G(CH_2) \]  

The ensemble averaged Gibbs strain energies are shown in Fig. S5 and the values are discussed solely in terms of overall trend, with high values representing greater strain. Cyclohexane exhibits the lowest strain and cyclodecane the largest, if cyclobutane is disregarded. The overall strain decreases for larger cycloalkanes, approaching the limit of zero for an \( n \)-alkane.
8  Cycloalkane discussion with respect to experimental or theoretical reference

8.1 Cyclobutane

Figure S 6: Highest populated conformer of cyclobutane.

Cyclobutane is identified in its puckered (wing-shaped) ring form as a single conformer[12]. This puckered conformation is more stable compared to the higher energy planar conformation and reduces torsional strain (see also Figure 5). The puckered conformation has been experimentally identified by Dunitz and Schomaker using electron diffraction.[13]

8.2 Cyclopentane

Figure S 7: Highest populated conformer of cyclopentane.

Like cyclobutane the dominant conformation of cyclopentane is non planar and is present in a single envelope conformation. The ring strain is reduced considerably compared to cyclobutane and the bond angles come closer to the ideal tetrahedral angle of 109.5°. The envelope conformation was confirmed by Aston, Schumann, Fink and Doty in 1941[14].
8.3 Cyclohexane

The lowest energy conformer of cyclohexane is the chair conformation which is attested by Raman- and IR spectroscopy as well as electron diffraction.\cite{15}. The cyclohexane ring is considered to be strain free and two type of carbon-hydrogen bonds can be identified (axial and equatorial).

8.4 Cycloheptane

Wiberg computed the twist-chair TC conformation to have the lowest free energy at room temperature, which is matched by our CRENOSO workflow.\cite{16}

8.5 Cyclooctane

Figure S 8: Highest populated conformer of cyclohexane.

Figure S 9: Highest populated conformer of cycloheptane.

Figure S 10: Highest populated conformer of cyclooctane.
Anet and Basus showed that cyclooctane’s dominant conformation is the boat-chair (BC) form at room temperature. The BC conformation in $C_8$ symmetry is correctly identified by our algorithm.

### 8.6 Cyclononane

![Figure S 11: Highest populated conformers of cyclononane.](image)

$^{13}$C NMR measurements at -162 °C from Anet et al. with two resolved signals confirmed Allinger’s strain-energy results that the twist boat chair (TBC) conformation in $D_3$ symmetry is the dominant conformation. As seen in Fig. 11 this is in perfect agreement with our calculations, with the twist-chair-boat (TCB) conformation ($P_2$) being higher in free energy.

### 8.7 Cyclodecane

![Figure S 12: Highest populated conformers of cyclodecane and overlay of all shown structures.](image)

The lowest conformation is the boat-chair-boat ($P_1$) conformation as identified from NMR measurements and confirmed by theoretical investigation. The second conformer has been identified as the (TBCC) conformation in $C_2$ symmetry, followed by the (TBC) conformation.
8.8 Cycloundecane

Pawar et al. showed cycloundecane to exist as a mixture of two conformations [12323] (59%) and [335] (41%) at -183.1 °C. In our calculations in CHCl₃ at room temperature the ordering inverses and we identify [335] (P₁) as highest populated followed by [12323] (P₂).

8.9 Cyclododecane

Crystal structure and low temperature carbon NMR measurements reveal a square conformation [3333] in Dale’s nomenclature with D₄ symmetry, consistent with the lowest free energy conformation we have identified. Anet identified the [2334] conformation (C₁) through strain-energy calculations as next viable candidate, which is our P₂ second conformer.
8.10 Cyclotridecane

Figure S 15: Highest populated conformers of cyclotridecane and overlay of all shown structures.

Conformer identification for cyclotridecane, an odd-membered macrocycle, by NMR measurements is nontrivial as even at -135 °C the pseudo-rotation leads to a single signal. For this reason comparisons were made based on crystal structures of similar compounds, which match our lowest free-energy conformer very well. The experimentally (X-ray data) derived population estimate on the dominant conformation is with 80 % very close to our evaluation in CHCl$_3$ $P_1$ = 88.5 %. [21, 25]

8.11 Cyclotetradecane

Figure S 16: Highest populated conformers of cyclotetradecane and overlay of all shown structures.

Low temperature carbon NMR measurements[23] of cyclotetradecane (-132 °C) validate the rectangular diamond lattice conformation identified as our lowest lying conformer. Followed by the [3344] conformation, correctly identified as the second lowest lying conformer. [26]
8.12 Cyclopentadecane

Figure S 17: Highest populated conformers of cyclopentadecane and overlay of all shown structures.

Anet et al. predicted several stable cyclopentadecane conformations at room temperature, where the [13443] conformation agrees with our quinquangular lowest free energy confor-mer. The conformer $P_4$ closely resembles the predicted [13434] conformation.

8.13 Cyclohexadecane

Figure S 18: Highest populated conformers of cyclohexadecane and overlay of all shown structures.

In 1974 Anet et al. performed low temperature $^1$H and $^{13}$C NMR measurements as well as iterative strain-energy calculations and concluded that the predominant cyclohexadecane conformation is the [4444] or diamond lattice, square conformation. The rectangular [3535] conformation which Anet et al. also considered for the explanation of the carbon signal ratio (1:2:1 at -152 °C) is also found in our ensemble and is in accord with their strain-energy computations 1.65 kcal mol$^{-1}$ less favorable in free energy compared to the square conformation.
8.14 Cycloheptadecane

![Figure S 19: Highest populated conformers of cycloheptadecane.](image)

In a comparison of computational search methods Saunders et al. investigated the potential energy surface of cycloheptadecane[28]. Combining all search results they estimated a Boltzmann population of the most stable conformer to be 8% at 300 K which is in perfect agreement with our lowest lying conformer and Boltzmann population $P_1 = 8.5\%$ (in CHCl$_3$).

8.15 Cyclooctadecane

![Figure S 20: Highest populated conformers of cyclooctadecane.](image)

Investigations of larger cycloalkanes increase in complexity through their flexibility and the large amount of different accessible conformations. Using Monte Carlo computations Shah et al. investigated cyclooctadecane and identified the [3636] (not shown here, but identified in our ensemble), [4545] equal to $P_3$, and the 'nicked' conformation [234234] (our lowest lying conformer $P_1$) conformations to be the most stable.[29]
8.16 Cyclononadecane

The odd-membered cyclononadecane was also investigated by Shah et al. and all of their shown stable conformers can be found in our ensemble, but in a different order because of our Gibbs free energy sorting (in CHCl₃) compared to their gas phase MM2 energies.\[29\]

8.17 Cycloicosane

Cycloicosane as the largest here investigated cycloalkane has the lowest lying conformer in [7373] conformation (i.e. a rectangle of 4 sides with length of 7, 3, 7 and 3) which matches our P₁ conformer.\[29\] Like in the case of cyclononadecane other conformers identified and presented by Shah et al. are found in our ensemble but are reordered based on free energy.
9 Application examples

In all three following application examples, the same technical settings are used as in the cycloalkane studies described in Table 3.

9.1 Application example I: cyclic peptide PGLGF

The characterization of cyclic peptides is a nontrivial task because cyclization confines the chemical environment, resulting in different chemical shifts compared to the linear peptide analog. Zaretsky et al. proposed and evaluated a computational workflow for the QM prediction of chemical shifts in cyclic peptides [30], which was refined into the CANDLE workflow in 2018 [31]. As a cyclic peptide model they studied PGLGF (see Fig. 23), which we evaluate here with our CRENSO algorithm.

![Lewis structure of the cyclic peptide PGLGF with corresponding carbon atom numbering.](image)

In this and the following to examples we employ our default CRENSO workflow consisting of the following steps:

1. conformer ensemble (CRE) generation using meta-dynamics driven by force-field and semi-empirical quantum mechanical (SQM) methods of the GFN-xtb family.
2. clustering of the ensemble to group similar conformers and condense the often large ensembles to the essential candidate conformers
3. refinement and sorting of the CRE at DFT level using CENSO
4. calculation of NMR properties facilitated by CENSO

In all three application examples, the same technical settings are used as in the cycloalkane studies described in Table 3.
Figure S 24: Potential energy surface (PES) evaluation and conformer ensemble ranking. Hydrogen atoms in the shown geometries are omitted for clarity.

Figure 24 details the CRE generation and sorting of the 18-membered cycle PGLGF. Within an energy window of 8 kcal mol\(^{-1}\) crest.combi identifies 294 conformations. These are clustered using a special principle component analysis clustering algorithm whereby 47 conformers remain. In part0 an evaluation with a fast but improved energy description at DFT level, including a solvation contribution to the free energy, reorders the conformer ensemble compared to the SQM ranking. The sorting of conformers at the DFT level is obviously more distributed compared to the GFN-SQM methods, which is also evident in Part1 where free energies with thermostatistical correction are considered. In part2 conformers are optimized at r2scan-3c level with implicit solvation and the final ensemble is obtained. Here the picture is that a very stable conformer is identified with a second conformer 3.27 kcal mol\(^{-1}\) higher in Gibbs energy, i.e., not significantly populated anymore at room temperature. The search was started from Zaretsky’s CONF1, which was reoptimized using consistent CRENSO settings and shows to be substantially higher in free energy compared to our identified final conformer.
Figure S 25: Geometries created by the CRENSO algorithm and in comparison to conformers created by Zaretsky et. al.\cite{30,32}. Hydrogen atoms are omitted for clarity.

The generated geometries and structures of Zaretsky et al. are shown in Fig. 25. In A the overlay of the clustered crest_combi ensemble shows a large diversity in the orientation of the ring itself, as well as the phenyl-, isopropyl- and linker groups. B1 and B2 show the final conformation obtained by CRENSO in two orientations, where the B2 geometry is oriented like the PGLGF X-ray structure from Ref.\cite{32} and from visual inspection appears to be a very close match to the crystal structure, with all groups oriented on the same side of the macrocycle. In Fig. 25C the CRENSO lowest lying conformation is compared to the reoptimized CONF1 of Zartesky and a large difference in the orientation of the linker group is visible. Parts D and E show the low and high energy conformers as identified by Zaretsky et al.
| residue | position | CRENOSO | zaretsky | zaretsky_MSTD |
|---------|----------|---------|----------|---------------|
| 0       | Pro1     | C=O     | -0.47    | 1.84          | 1.50          |
| 1       | Gly2     | A       | 0.74     | 3.06          | 2.75          |
| 2       | Gly2     | C=O     | -2.82    | 4.44          | 7.78          |
| 3       | Leu3     | alpha   | -1.46    | 1.45          | 0.93          |
| 4       | Leu3     | beta    | -2.68    | 1.26          | -1.32         |
| 5       | Leu3     | gamma   | 0.21     | 3.40          | 3.69          |
| 6       | Leu3     | delta   | -0.37    | 1.13          | -0.62         |
| 7       | Leu3     | delta   | -4.05    | 1.13          | -0.62         |
| 8       | Leu3     | C=O     | -3.46    | 1.03          | 4.37          |
| 9       | Gly4     | alpha   | -0.11    | 4.05          | 3.70          |
| 10      | Gly4     | C=O     | -0.40    | 2.10          | 1.45          |
| 11      | Phe5     | alpha   | 2.71     | 4.35          | 3.66          |
| 12      | Phe5     | beta    | -0.17    | 2.50          | 2.43          |
| 13      | Phe5     | gamma   | -2.95    | 2.16          | -1.08         |
| 14      | Phe5     | delta   | -1.78    | 1.62          | -4.53         |
| 15      | Phe5     | epsilon | -3.09    | 1.93          | -4.81         |
| 16      | Phe5     | zeta    | -1.97    | 2.61          | -5.41         |
| 17      | Phe5     | epsilon | -3.09    | 1.93          | -4.81         |
| 18      | Phe5     | delta   | -1.78    | 1.62          | -4.53         |
| 19      | Phe5     | C=O     | -1.29    | 0.84          | 4.29          |
| 20      | Linker6  | alpha   | -0.63    | 1.48          | 1.17          |
| 21      | Linker6  | beta    | -0.47    | 1.83          | -1.17         |
| 22      | Linker6  | gamma   | -3.44    | 0.74          | -0.27         |
| 23      | Linker6  | Exo-C=O | -1.20    | 1.01          | 2.52          |
| 24      | Linker6  | tBuC    | 0.29     | 5.00          | 4.37          |
| 25      | Linker6  | tBuCH   | -2.10    | 2.89          | -2.52         |
| 26      | Linker6  | tBuCH   | -2.10    | 2.89          | -2.52         |
| 27      | Linker6  | tBuCH   | -2.10    | 2.89          | -2.52         |
| 28      | Pro1     | delta   | 3.57     | 3.99          | -4.28         |
| 29      | Pro1     | gamma   | 0.20     | 0.76          | -0.29         |
| 30      | Pro1     | beta    | 1.64     | 2.98          | -2.66         |
| 31      | Pro1     | alpha   | -3.33    | 2.45          | 1.52          |

**MAD** (mean absolute deviation):

\[
\frac{1}{n} \sum_{i=1}^{n} |\delta_{\text{calc}} - \delta_{\text{exp}}|
\]

Figure S 26: $^{13}$C chemical shift deviations compared to experiment.\[30\] \[32\]
The calculated isotropic shielding constants from the final CRENSEO conformation are converted to chemical shifts by the scaling factors and reference determined in section 1.2. The chemical shift deviations compared to the experiment are color coded in Fig. 26 and results from Zaretsky et al. who used a similar scaling and a multistandard (MSTD) approach for converting shielding constants to chemical shifts are listed for comparison. The mean absolute deviation (MAD) shows very good overall agreement with the experiment at 1.77 ppm for the CRENSEO workflow. The entire CRENSEO calculation took roughly 12 hours on 28 cores of an Intel(R) Xeon(R) CPU E5-2660 v4 @ 2.00GHz machine.

9.2 Application example II: lobophytolins A

In 2020 Li et al. isolated macrocyclic lobophytolins A from soft coral and assigned the structure using NMR and QM calculations. The lewis structure of lobophytolins A is shown in Fig. 29 and is investigated by the same procedure and settings as described in the previous example application section 9.1.

The evolution of CRE generation and sorting is shown in Fig. 27, starting with 626 conformers identified by crest_combi and reduced to 120 after clustering. Free-Energy evaluation at DFT level in part0 and part1 reorder the conformers and show a less flat PES compared to the SQM ranking. In the final populated CRENSEO ensemble remain 20 conformers, which are shown in Fig. 28.

![Conformer ranking during the CRENSEO procedure](image)

Figure S 27: PES evaluation and conformer ensemble sorting.
Figure S 28: Geometries created by the CRENOSO algorithm and in comparison to the crystal structure\cite{33}. Hydrogen atoms are omitted for clarity.

Part A1 and A2 in Fig. 28 show the overlay of the clustered crest_combi ensemble. The overlay of the final CRENOSO ensemble is presented in B. Fig. 28C shows the overlay of the CRENOSO lowest lying CHCl$_3$ solution conformer ($P_1 = 30.07\%$) in green with the crystal structure in blue. The populations of the other conformers are listed in Table 4.
| #CONF | $\Delta G$ [kcal $\cdot$ mol$^{-1}$] | P [%] |
|------|-----------------|------|
| 1    | 0.00            | 30.07|
| 2    | 0.31            | 17.82|
| 3    | 0.81            | 7.71 |
| 4    | 1.06            | 5.07 |
| 5    | 1.10            | 4.76 |
| 6    | 1.22            | 3.88 |
| 7    | 1.27            | 3.58 |
| 8    | 1.30            | 3.40 |
| 9    | 1.30            | 3.40 |
| 10   | 1.37            | 3.04 |
| 11   | 1.39            | 2.93 |
| 12   | 1.40            | 2.87 |
| 13   | 1.57            | 2.17 |
| 14   | 1.62            | 2.00 |
| 15   | 1.69            | 1.77 |
| 16   | 1.71            | 1.71 |
| 17   | 2.01            | 1.04 |
| 18   | 2.04            | 0.99 |
| 19   | 2.06            | 0.95 |
| 20   | 2.13            | 0.85 |

Table S 4: Free energy differences and Boltzmann populations of the conformers in the final CRENSO ensemble of lobophytolins A.
The calculated isotropic shielding constants from the final CRENSO conformation are converted to chemical shifts by the scaling factors and reference determined in section 1.2. The agreement of the calculated chemical shifts with experiment are exceptionally good for the CRENSO solution ensemble, with a MAD of 1.06 ppm. The entire CRENSO calculation took roughly 15 hours on 28 cores of an Intel(R) Xeon(R) CPU E5-2660 v4 @ 2.00GHz machine.

9.3 Application example III: (-)-dactylolide

Larsen et al. studied the solution conformation behavior of (-)-dactylolide (see Fig 32), which acts as a tubulin binding agent and can be used as a chemotherapeutic agent. The CHCl₃ solution conformations of synthesised (-)-dactylolide was studied by NMR measurements.
and computational modeling, creating a solution ensemble using NMR data in the DISCON (Distribution of Solution Conformations) software.

We applied the default CRENSO workflow as already described in section 9.1 and the ensemble generation and evaluation is shown in Fig. 30.

Figure S 30: PES evaluation and conformer ensemble ranking. And lowest lying CRENSO solution conformer geometry.

The combined conformer search yields 483 conformers in an energy window of 8 kcal mol$^{-1}$ and 97 clustered conformations are evaluated in the CENSO algorithm. The sorting at DFT level using free energy thresholds drastically reduces the number of relevant conformer candidates and only 9 conformers remain in the final ensemble.
Table S 5: Free energy differences and Boltzmann populations of the conformers in the final CRENOSO ensemble of (-)-dactylolide.

| #CONF | ΔG [kcal · mol$^{-1}$] | P [%] |
|-------|------------------------|-------|
| 1     | 0                      | 90.31 |
| 2     | 1.72                   | 4.94  |
| 3     | 2.25                   | 2.01  |
| 4     | 2.87                   | 0.71  |
| 5     | 2.91                   | 0.66  |
| 6     | 3.02                   | 0.55  |
| 7     | 3.12                   | 0.47  |
| 8     | 3.33                   | 0.33  |
| 9     | 4.96                   | 0.02  |

Figure S 31: Geometries created by the CRENOSO algorithm and in comparison to zampanolide cocrystal structure bound to tubulin H229. \[35\]
The large population of 90.3% suggest a very stable conformation (see Fig. 31A2). The lowest lying conformation closely resembles the cocrystal structure of zampanolide (a similar polyketide natural product) bound to tubulin H229.\cite{35}

| position | type | δ (exp) | isotropic shielding constant | CRENNO |
|----------|------|----------|-----------------------------|----------|
| 0        | 1    | 166.50   | 10.46                       | -0.90    |
| 1        | 2    | 119.90   | 64.77                       | -5.50    |
| 2        | 3    | 140.60   | 34.96                       | 1.90     |
| 3        | 4    | 125.70   | 56.50                       | -3.50    |
| 4        | 5    | 144.20   | 22.76                       | 9.80     |
| 5        | 6    | 45.00    | 138.64                      | -0.20    |
| 6        | 7    | 197.60   | -23.98                      | 0.50     |
| 7        | 8    | 131.60   | 51.10                       | -4.30    |
| 8        | 9    | 146.20   | 22.73                       | 7.90     |
| 9        | 10   | 39.90    | 142.33                      | 1.40     |
| 10       | 11   | 76.60    | 105.20                      | -0.30    |
| 11       | 12   | 41.00    | 142.51                      | 0.10     |
| 12       | 13   | 143.60   | 30.54                       | 3.10     |
| 13       | 14   | 40.60    | 142.59                      | 0.50     |
| 14       | 15   | 75.90    | 105.46                      | 0.20     |
| 15       | 16   | 130.70   | 49.56                       | -1.90    |
| 16       | 17   | 131.10   | 44.03                       | 2.90     |
| 17       | 18   | 39.90    | 143.68                      | 0.10     |
| 18       | 19   | 75.50    | 106.32                      | -0.20    |
| 19       | 13a  | 109.50   | 73.74                       | -3.50    |
| 20       | 17a  | 16.20    | 170.27                      | -1.20    |
| 21       | 19a  | 199.30   | -30.26                      | 4.70     |
| 22       | 5a   | 24.30    | 160.78                      | -0.40    |

MAD: 2.39

Figure S 32: Comparison of $^{13}$C chemical shifts to experiment.\cite{34}

The calculated isotropic shielding constants from the final CRENNO conformation are converted to chemical shifts by the scaling factors and reference determined in section 1.2. The agreement of the calculated chemical shifts with the experiment is reasonable with an overall MAD of 2.39 ppm. The entire CRENNO calculation took roughly 9 hours on 28 cores of an Intel(R) Xeon(R) CPU E5-2660 v4 @ 2.00GHz machine.
10 References

[1] M. W. Lodewyk, M. R. Siebert, D. J. Tantillo, Chem. Rev. 2012, 112, 1839–1862.
[2] Chemical Shift Repository, http://cheshireNMR.info.
[3] P. R. Rablen, S. A. Pearlman, J. Finkbiner, J. Phys. Chem. A 1999, 103, 7357–7363.
[4] H. Fritz, E. Logemann, F. Schill, T. Winkler, Chem. Ber. 1976, 109, 1258–1268.
[5] J. J. Burke, P. C. Lauterbur, J. Am. Chem. Soc. 1964, 86, 1870–1871.
[6] K. B. Wiberg, W. E. Pratt, W. F. Bailey, J. Org. Chem. 1980, 45, 4936–4947.
[7] K. C. Rice, R. E. Wasylkisen, Org. Magn. Res. 1976, 8, 449–452.
[8] J. B. Stothers, Carbon-13 NMR spectroscopy, Academic Press, New York, 1972.
[9] K. B. Wiberg, J. Comput. Chem. 1984, 5, 197–199.
[10] M. R. Ibrahim, P. Von Ragué Schleyer, J. Comput. Chem. 1985, 6, 157–167.
[11] P. R. Rablen, Chemistry 2020, 2, 347–360.
[12] R. M. Moriarty in Topics in Stereochemistry, John Wiley & Sons, Ltd, 1974, pp. 271–421.
[13] J. D. Dunitz, V. Schomaker, J. Chem. Phys. 1952, 20, 1703–1707.
[14] J. G. Aston, S. C. Schumann, H. L. Fink, P. M. Doty, J. Am. Chem. Soc. 1941, 63, 2029–2030.
[15] D. H. Barton, R. C. Cookson, Q. Rev. Chem. Soc. 1956, 10, 44–82.
[16] K. B. Wiberg, J. Org. Chem. 2003, 68, 9322–9329.
[17] O. V. Dorofeeva, V. S. Mastryukov, N. L. Allinger, A. Almenningen, J. Phys. Chem. 1985, 89, 252–257.
[18] F. A. Anet, J. J. Wagner, J. Am. Chem. Soc. 1971, 93, 5266–5268.
[19] D. M. Pawar, S. V. Smith, H. L. Mark, R. M. Odom, E. A. Noe, J. Am. Chem. Soc. 1998, 120, 10715–10720.
[20] D. M. Pawar, J. Brown, K. H. Chen, N. L. Allinger, E. A. Noe, J. Org. Chem. 2006, 71, 6512–6515.
[21] H. F. Dos Santos, M. L. Franco, M. F. Venâncio, D. E. Ferreira, C. P. Anconi, W. R. Rocha, W. B. De Almeida, Int. J. Quantum Chem. 2012, 112, 3188–3197.
[22] J. D. Dunitz, H. M. M. Shearer, Helv. Chim. Acta 1960, 43, 18–35.
[23] F. A. Anet, A. K. Cheng, J. J. Wagner, J. Am. Chem. Soc. 1972, 94, 9250–9252.
[24] F. A. Anet, T. N. Rawdah, J. Am. Chem. Soc. 1978, 100, 7166–7171.
[25] B. H. Rubin, M. Williamson, M. Takeshita, F. M. Menger, F. A. L. Anet, B. Bacon, N. L. Allinger, J. Am. Chem. Soc. 1984, 106, 2088–2092.
[26] V. Dragojlovic, ChemTexts 2015, 1, 1–30.
[27] F. A. L. Anet, A. K. Cheng, J. Am. Chem. Soc. 1975, 97, 2420–2424.
[28] M. Saunders, K. N. Houk, Y. D. Wu, W. C. Still, M. Lipton, G. Chang, W. C. Guida, *J. Am. Chem. Soc.* **1990**, *112*, 1419–1427.

[29] A. V. Shah, D. P. Dolata, *J. Comput. Aided Mol. Des.* **1993**, *7*, 103–124.

[30] S. Zaretsky, J. L. Hickey, M. A. St. Denis, C. C. Scully, A. L. Roughton, D. J. Tantillo, M. W. Lodewyk, A. K. Yudin, *Tetrahedron* **2014**, *70*, 7655–7663.

[31] Q. N. N. Nguyen, J. Schwochert, D. J. Tantillo, R. S. Lokey, *Phys. Chem. Chem. Phys.* **2018**, *20*, 14003–14012.

[32] S. Zaretsky, C. C. G. Scully, A. J. Lough, A. K. Yudin, *Chem. Eur. J.* **2013**, *19*, 17668–17672.

[33] S.-W. Li, C. Cuadrado, L.-G. Yao, A. H. Daranas, Y.-W. Guo, *Org. Lett.* **2020**, *22*, 4093–4096.

[34] E. M. Larsen, M. R. Wilson, J. Zajicek, R. E. Taylor, *Org. Lett.* **2013**, *15*, 5246–5249.

[35] A. E. Prota, K. Bargsten, D. Zurwerra, J. J. Field, J. F. Díaz, K.-H. Altmann, M. O. Steinmetz, *Science* **2013**, *339*, 587–590.