Computational Workflow for Ring Strain and Cascade Ring-Opening Thermochemistry in Trimethyl-Lock Macrocycles

Abstract

We outline a reproducible quantum-chemistry workflow to (i) quantify the thermodynamic driving force for a specific ring-opening / ring-closing cascade reaction of trimethyl-lock (TML) macrocycles and (ii) extract a "pure" ring-strain measure tied to the scissile bond involved in the ring-opening step. The protocol is designed for SMILES-based structure generation, conformer sampling, DFT geometry optimization with dispersion and implicit solvation (COSMO), and harmonic frequency calculations using the def2-SVP basis.

1 Scope and target quantities

We consider macrocycles M that undergo an amine-triggered cascade consisting of: (1) nucleophilic attack and opening at an activated carbonate/ester in the TML motif and (2) rapid intramolecular cyclization to release a small lactone/benzolactone fragment C while regenerating a terminal amine on the opened chain B.

2 Molecular construction and mechanistic mapping

2.1 Input representation

Each species is represented by a SMILES string (or a set of SMILES enumerations for tautomers/protonation states). We require:

- Macrocycle M (intact ring).
- Open-chain product B (post-cascade, with regenerated terminal amine).
- Released lactone fragment C (TML-derived ring-closed leaving fragment).

2.2 Bond mapping and reaction templates

A bond-map is specified for the *scissile bond* (the bond broken upon amine attack) using either:

1. explicit atom indices in the M graph, or
2. a SMARTS pattern that matches the activated carbonate/ester linkage within the TML motif.

This mapping is used consistently for (a) generating B and C and (b) constructing ring-strain reference structures.

3 Conformer generation and pre-screening

3.1 RDKit sampling and force-field relaxation

For each species $S \in \{M, B, C\}$:

1. Generate N_{init} 3D conformers using ETKDG (macrocycles may require larger N_{init}).

2. Relax each conformer using MMFF94 (or UFF) to remove severe clashes.

3. Cluster conformers by RMSD (e.g., Butina clustering) and retain up to N_{keep} representatives.

Recommended starting points: $N_{\text{init}} \sim$ 500–5000 for M and B (macrocycle/open chain), and $N_{\text{init}} \sim$ 100–500 for C, followed by $N_{\text{keep}} \sim$ 25–200 after clustering.

4 DFT optimization, solvation, and frequencies

4.1 Electronic structure settings

For each retained conformer i of species S, compute:

1. Geometry optimization at DFT/def2-SVP with chosen dispersion correction.

2. Implicit solvation via COSMO with the target solvent dielectric.

3. Harmonic frequency calculation at the same level

4.2 Opening-with-caps relative to an acyclic control

If only M is readily available, define a ring-opening operation that breaks the scissile bond and caps both ends (e.g., with H/Me) to yield M_{open}. Apply the *same* opening-and-capping operation to an acyclic control molecule X that contains the same activated linkage but is not ring-closed, yielding X_{open}. Define

$$\Delta E^{\text{cyc}}_{\text{open}} = E(M_{\text{open}}) - E(M), \tag{1}$$

$$\Delta E^{\text{acyc}}_{\text{open}} = E(X_{\text{open}}) - E(X), \tag{2}$$

and the ring strain estimate

$$E_{\text{strain}} = \Delta E^{\text{cyc}}_{\text{open}} - \Delta E^{\text{acyc}}_{\text{open}}. \tag{3}$$