A set of Quantum–Mechanically Derived Force Fields for Natural and Synthetic Retinal Photoswitches

Razan E. Daoud,[a] Simone Veglianti,[a] Anna Piras,[b] Abderrahmane Semmeq,[b] Samuele Giannini,[b] Giacomo Prampolini,[b*] and Daniele Padula[a†]

[a] *Dipartimento di Biotecnologie, Chimica e Farmacia, Università di Siena, Via A. Moro 2, 53100 Siena, Italy*

[b] *Istituto di Chimica dei Composti OrganoMetallici (ICCOM-CNR), Area della Ricerca, Via G. Moruzzi 1, I-56124 Pisa, Italy*

Supporting Information

*To whom correspondence should be addressed. Email: giacomo.prampolini@pi.iccom.cnr.it

†To whom correspondence should be addressed. Email: daniele.padula@unisi.it

Contents

List of Figures

List of Tables

S1 Electronic Data

Electronic versions of QMD–FFs are available at a public Github repository:
`https://github.com/dpadula85/rets_qmd_ffs`.

S2 QMD–FFs parameterisation details

Following the JOYCE procedure, the internal coordinates (ICs) chosen to represent the target molecule dynamics can be classified into two different sets: stiff or flexible, depending on whether small displacements from the equilibrium geometry or large amplitude motions are expected. E^{intra} is therefore further partitioned in several contributions, namely

$$E^{intra} = E_s + E_b + E_{st} + E_{ft} + E^{intra}_{Nb} \tag{S1}$$

The first three terms refer to stiff IC, and can be approximated through harmonic potentials, *i.e.*

$$E_s = \frac{1}{2}\sum_{s}^{N_s} k_s(r - r^0)^2 \; ; \; E_b = \frac{1}{2}\sum_{b}^{N_b} k_b(\theta - \theta^0)^2 \; ; \; E_{st} = \frac{1}{2}\sum_{st}^{N_{st}} k_{st}(\phi - \phi^0)^2 \tag{S2}$$

where k_s, k_b, k_{st} are force constances associated to, respectively, the s-th bond length, the b-th bond angle, and the st-th stiff torsion, while r^0, θ^0, ϕ^0 are their equilibrium positions. While the first two terms, *i.e.* the stretching and bending contributions E_s and E_b, are widely used by popular FFs, the third expression (E_{st}) is seldom employed for describing torsions of stiff dihedral angles in standard approaches. However, it is better suited to account for the internal energy due to small and fast distortions of a dihedral ϕ from its equilibrium position ϕ^0. The JOYCE procedure routinely employs the E_{st} term, to mimic the behaviour of a stiff and fast oscillating torsion, as those ruling the planarity of aromatic rings or conjugated double bonds.[1–6] The last two terms of Eq. S1 are instead assigned to flexible ICs, which are expected to experience larger displacements

during simulation. In fact, the term

$$E_{ft} = \sum_{\mu}^{N_{ft}} \sum_{j}^{N_{cos_\mu}} k_{j\mu}^{ft} \left[1 + \cos\left(n_j^\mu \delta_\mu - \gamma_j^\mu\right)\right] \tag{S3}$$

(where δ_μ is the μ-th flexible torsion, $k_{j\mu}^{st}$ its force constant, n_j^μ the periodicity, and γ_j^μ a proper phase), is designed to describe the large amplitude displacements induced by a flexible dihedral, as those defining, for instance, the rotation around a σ bond. Finally, the last term of Eq. S1 contains the so-called nonbonded intramolecular interactions. In most popular transferable FFs, such contributions are counted among all possible pairs of interacting sites i-j, adopting the same set of Lennard–Jones (LJ) and charge parameters for both inter- and intra-molecular interactions. Conversely, the JOYCE protocol allows for a more specific parameterisation, where intramolecular nonbonded parameters may be included in the QMD–FF just for selected atom pairs and can take in principle different values than those adopted for the description of intermolecular interactions. Hence, following recent JOYCE parameterisations carried out for organic chromophores,[4,5,7,8] nonbonded intramolecular interactions have been modelled only through the LJ term,

$$E_{Nb}^{intra} = \sum_{i}^{N_{pairs}-1} \sum_{i<j}^{N_{pairs}} 4\varepsilon_{ij}^{intra} \left[\left(\frac{\sigma_{ij}^{intra}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}^{intra}}{r_{ij}} \right)^{6} \right] \tag{S4}$$

where i and j run over atoms belonging to the same molecule, ε_{ij} is the depth of the minimum of the LJ potential curve, σ_{ij} is the distance regulating the switch from attractive to repulsive in the LJ potential curve, r_{ij} is their distance, and N_{pairs} is the number of selected interacting pairs. Once the FF functional form and the underlying ICs have been selected, a complete set of QMD–FF parameters is retrieved through the JOYCE 3.0 code by minimizing the functional

$$I^{intra} = \sum_{K}^{3N-6} \sum_{K \leq L}^{3N-6} \frac{2W''_{KL}}{C} \left[H_{KL} - \left(\frac{\partial^2 E_{FF}^{intra}}{\partial Q_K \partial Q_L} \right) \right]_{g_0}^2 + \sum_{g}^{N_{geom}} W_g \left[\Delta U_g - E_g^{intra} \right]^2 \tag{S5}$$

In the first term on the right side of Eq. S5, the double sum runs over the $3N-6$ coordinates, C is a normalisation factor, Q_K is the Kth normal mode, whereas H_{KL} and $\frac{\partial^2 E^{intra}}{\partial Q_K \partial Q_L}$ are QM and MM Hessian matrix elements, both evaluated at the equilibrium geometry (g_0). The normalized diagonal elements of the weight matrix $\mathbf{W}''$ are set as in most previous applications,[6] at twice the value of those corresponding to the off diagonal terms. In the second term, ΔU_g is the QM computed energy difference between the the g and g_0 geometries and N_{geom} is the number of the different geometrical arrangements sampled for each target molecule. The weights W_g are all set to the same value.

While all detailed features of the JOYCE parameterization can be found in the original papers,[1,6,9] a brief summary of the protocol recently validated on the rPSB11 retinal[6] (see Fig. 1) and applied herein to all considered targets is outlined in the following:

1. Assign specific atomic types, based on symmetry and chemical equivalence considerations (as depicted in Fig. 1)

2. Select a suitable set of internal coordinates (ICs) which completely define the molecular geometry and select appropriate model functions for each IC. Following the JOYCE protocol,[1,6] torsional coordinates are further classified in stiff harmonic dihedrals (ϕ), such as the ones ruling the planarity of the C=C double bonds and their surrounding atoms, and flexible dihedrals (δ). These coordinates are modelled through the periodic function given in Eq. S3, appropriate to describe large amplitude rotations expected for flexible bonds along the conjugated chain. Rotation of methyl groups is also treated in the same way, albeit their torsional potential is not modelled *ad hoc* but taken from AMBER parameters.

3. Select non–bonded distances between atoms pairs not directly bonded to each other that could come close as a consequence of molecular flexibility. Interactions between separate regions of the same molecule (*e.g.* head–to–tail) can lead to atomic overlaps that are unphysical, making necessary the inclusion of intramolecular LJ terms entering in Eq. S4.

4. Obtain the optimal intramolecular parameters assigned to the chosen ICs solving the least

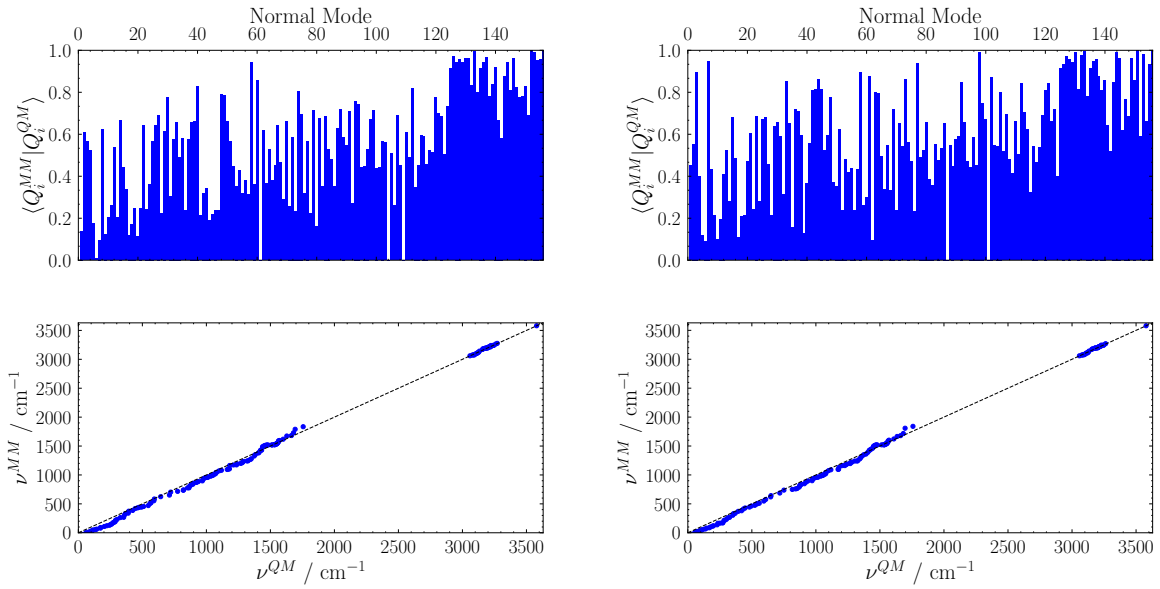
squares problem between the QM reference data and the MM result. The full set of best-fit parameters is achieved by the standard two–step procedure adopted in the JOYCE protocol,[6] where all stiff ICs are simultaneously parameterised in the first step, while flexible torsional coordinates are subsequently modelled, exploiting QM relaxed scans as reference data.

The procedure outlined above can be carried out in an automatic fashion to enhance reproducilibity and to reduce users' arbitrary choices, thanks to the latest developments implemented in the JOYCE 3.0 code and thoroughly described in recent work.[6] We highlight, also, that we did not focus on the parameterisation of torsional coordinates governing the conformation of the β–ionone ring, but rather we adopted Fourier series with periodicity $n = 3$ in combination with 1,4 non–bonded intramolecular terms, as done in AMBER. Transfer of torsional coordinates from GAFF for this ring is due to the fact that ring parameterisation is a complicated process (see Section 3.2), and that this region of the molecule is not involved in the conjugated system responsible for the electronic properties of common interest

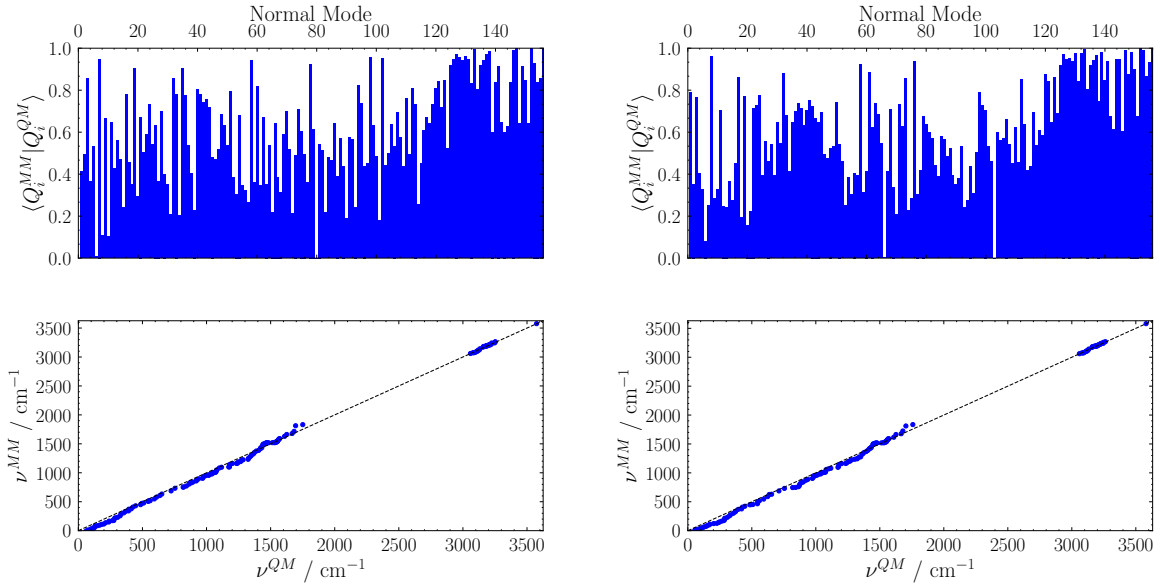
S3 Natural Retinal Protonated Schiff Bases (rPSBs)

S3.1 QMD–FFs

S3.1.1 Vibrational modes and frequencies



(a) rPSB11 **(b)** rPSB9



(c) rPSB13 **(d)** rPSBT

Fig. S1: Comparison between QM and MM vibrational frequencies and overlap between QM and MM normal modes for naturally occurring rPSBs.

S3.1.2 Torsional Profiles

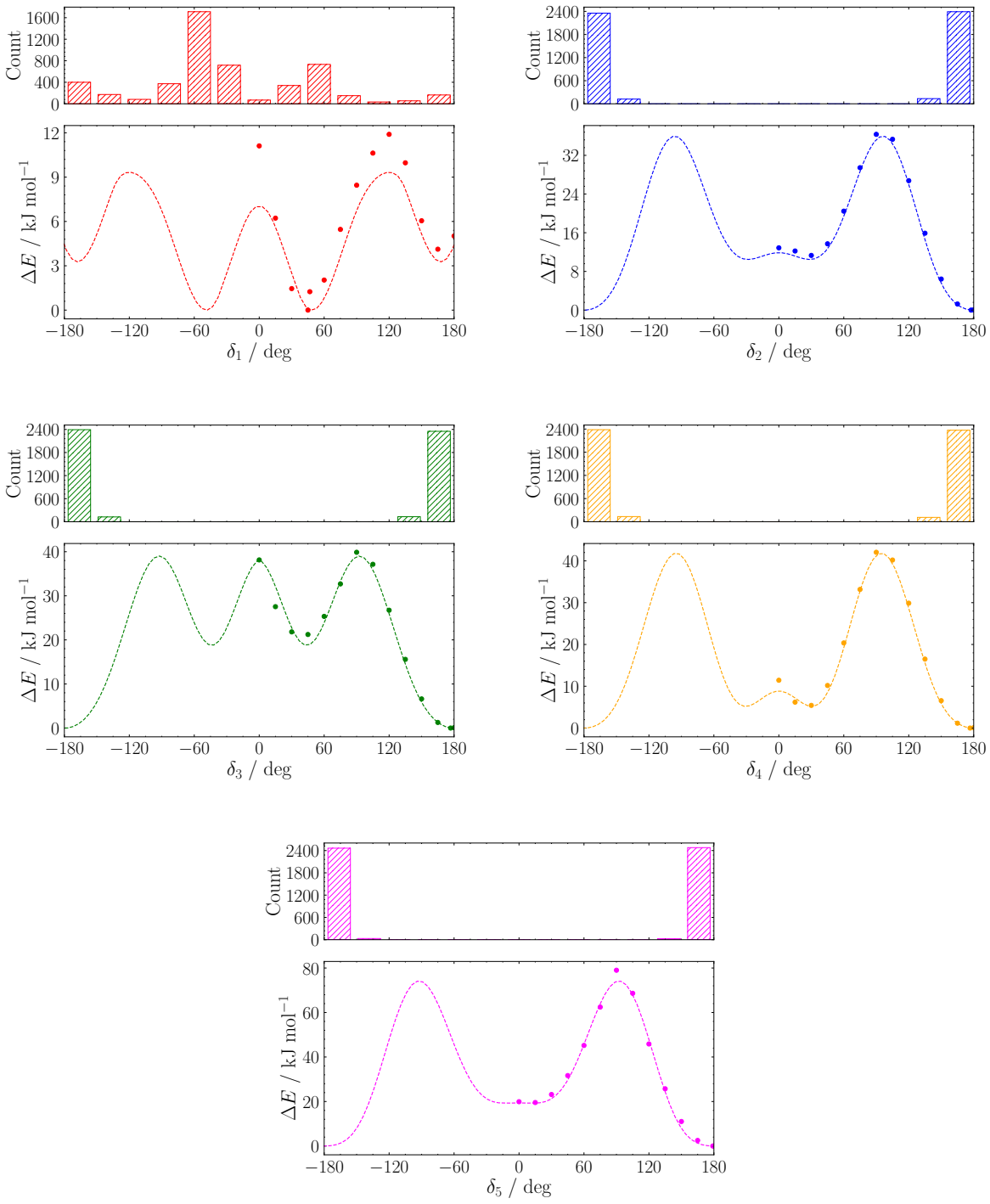


Fig. S2: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB11, and their distribution along a 500 ps MD run.

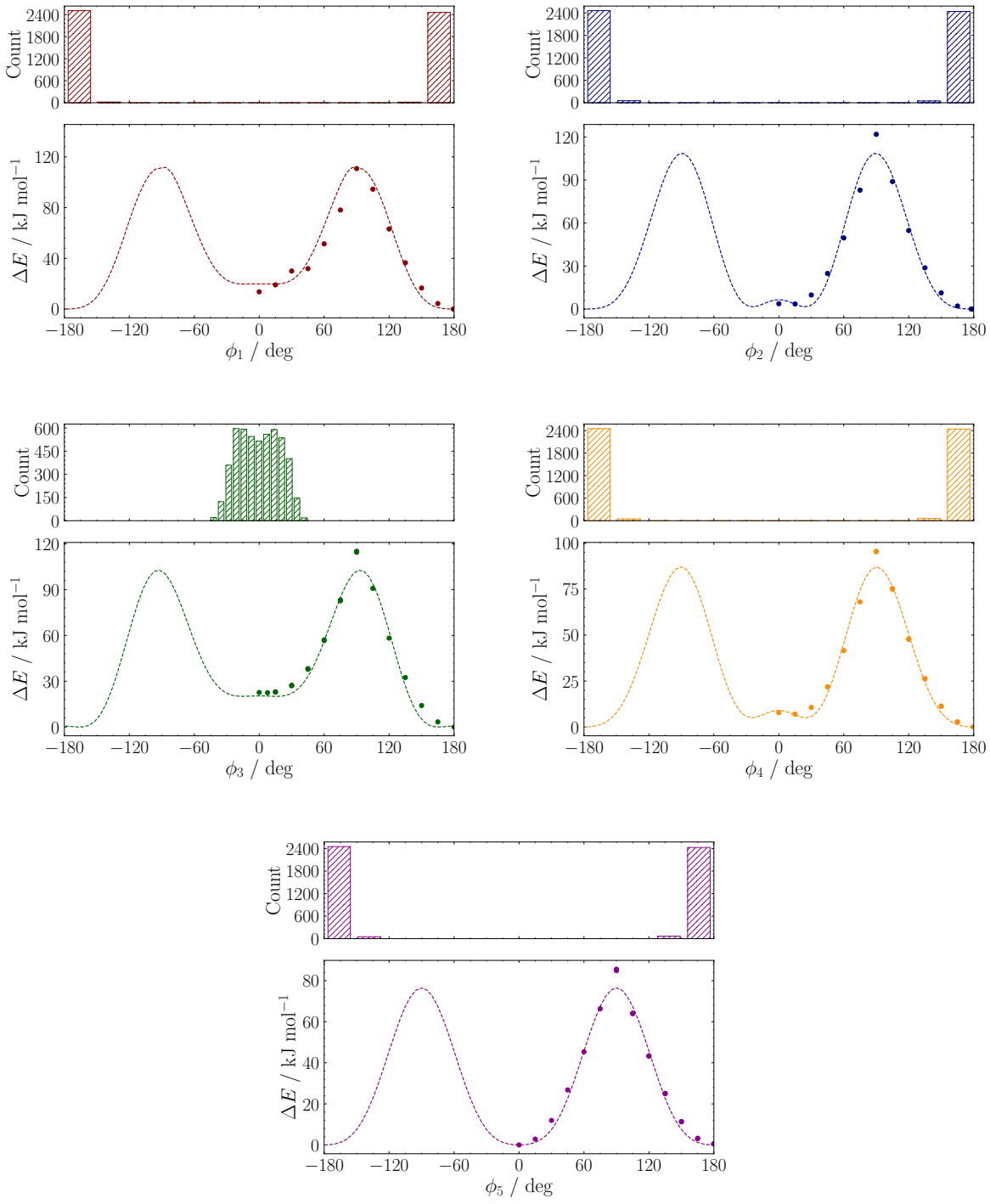


Fig. S3: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11, and their distribution along a 500 ps MD run.

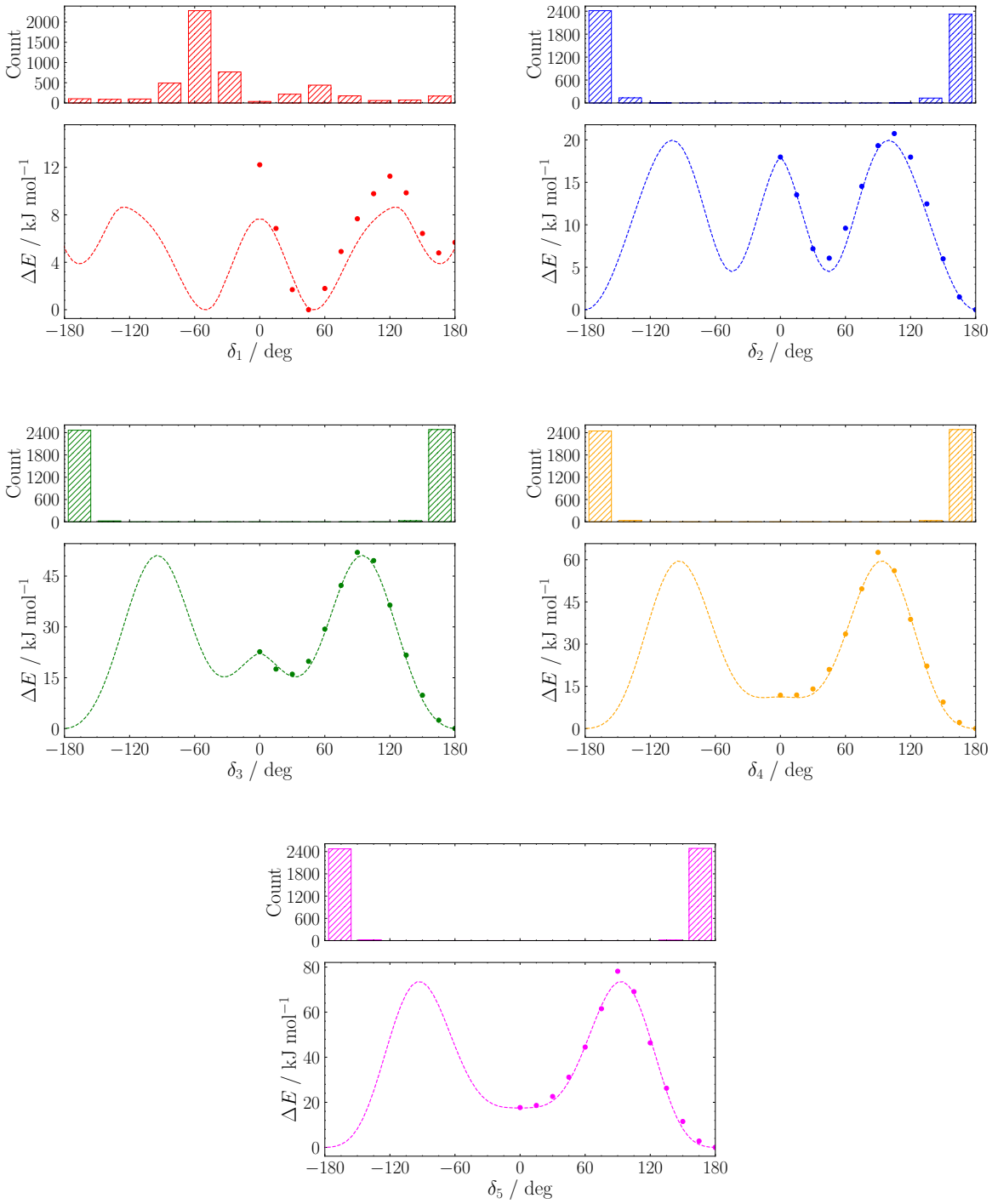


Fig. S4: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB9, and their distribution along a 500 ps MD run.

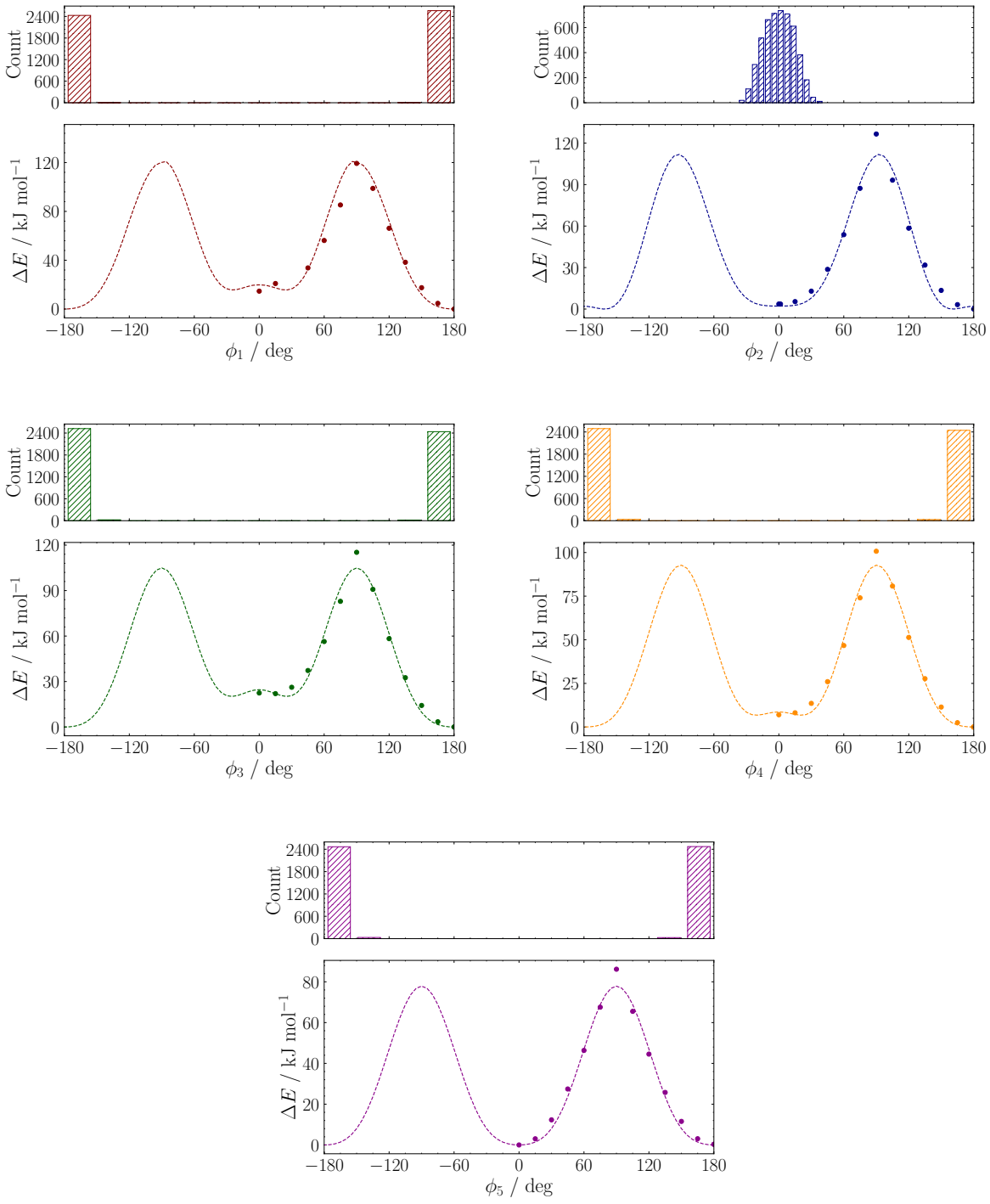


Fig. S5: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB9, and their distribution along a 500 ps MD run.

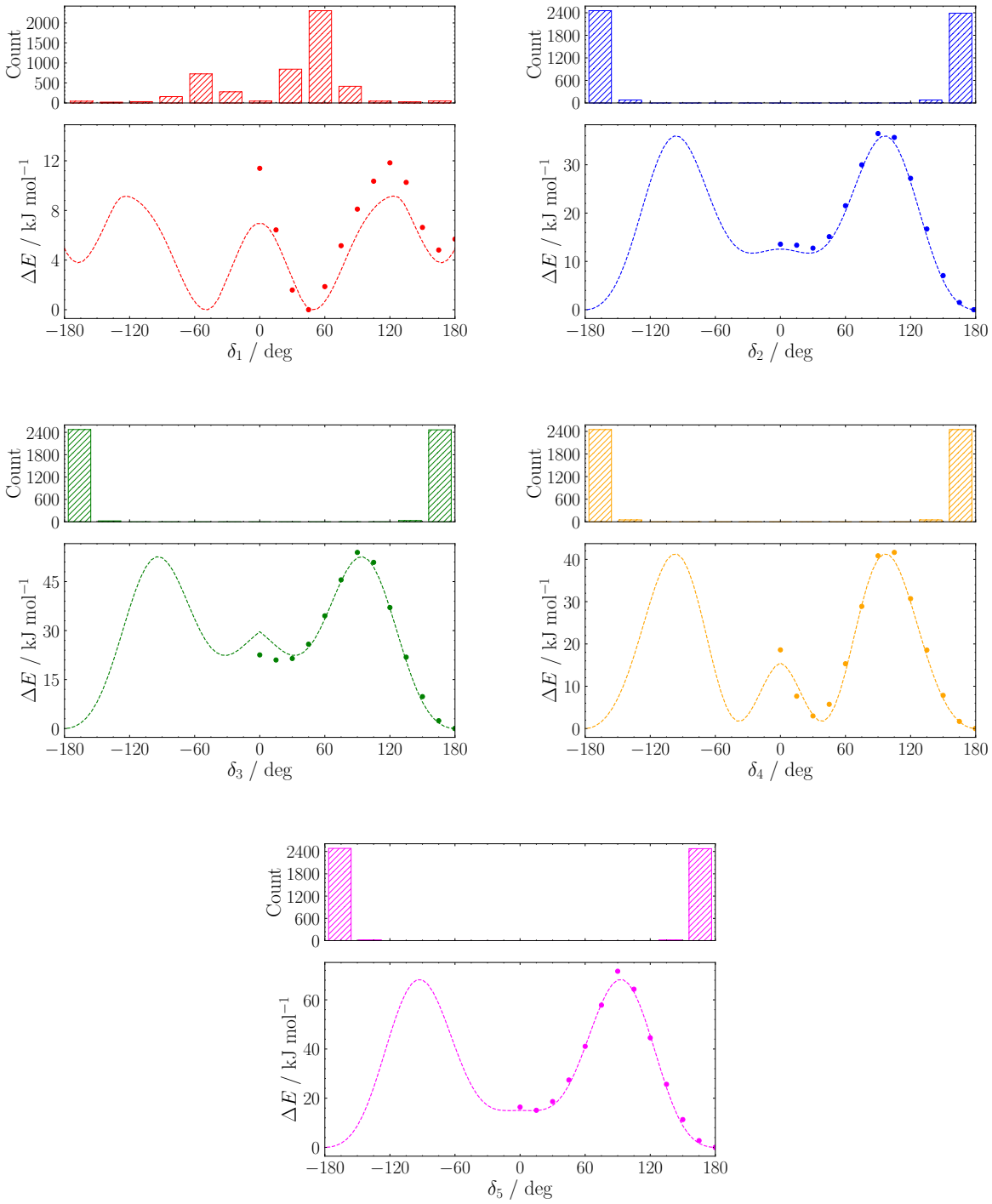


Fig. S6: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB13, and their distribution along a 500 ps MD run.

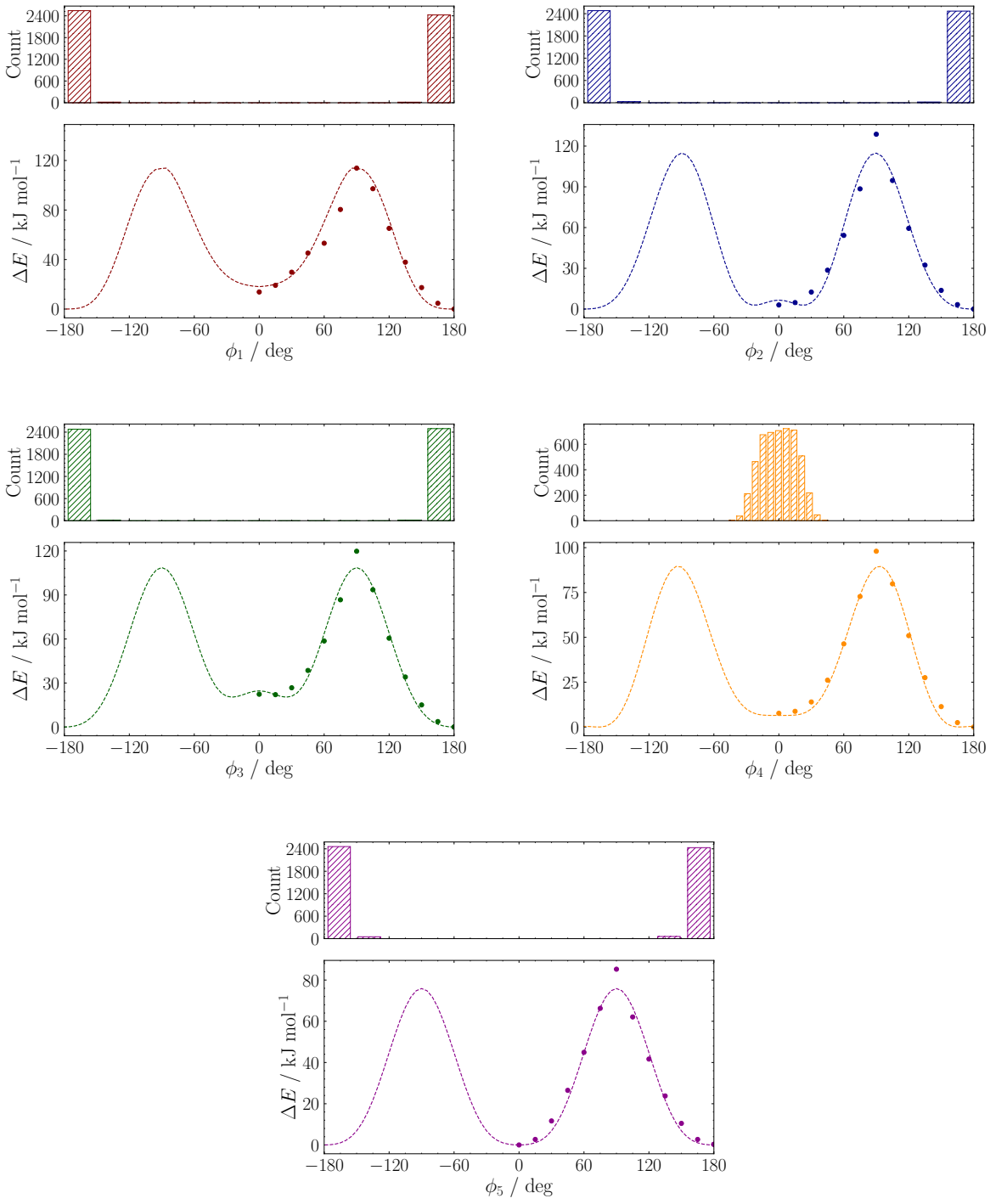


Fig. S7: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB13, and their distribution along a 500 ps MD run.

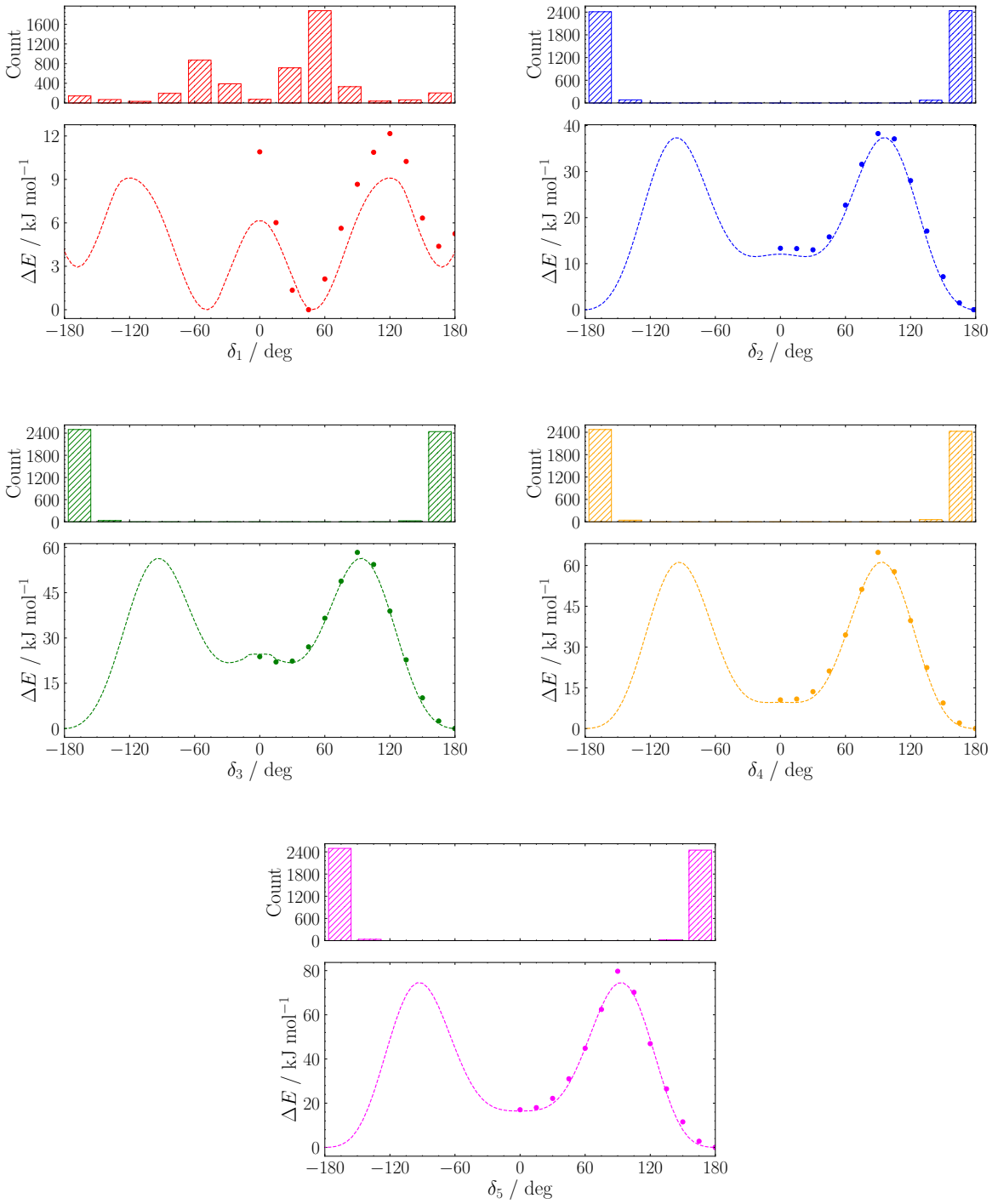


Fig. S8: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSBT, and their distribution along a 500 ps MD run.

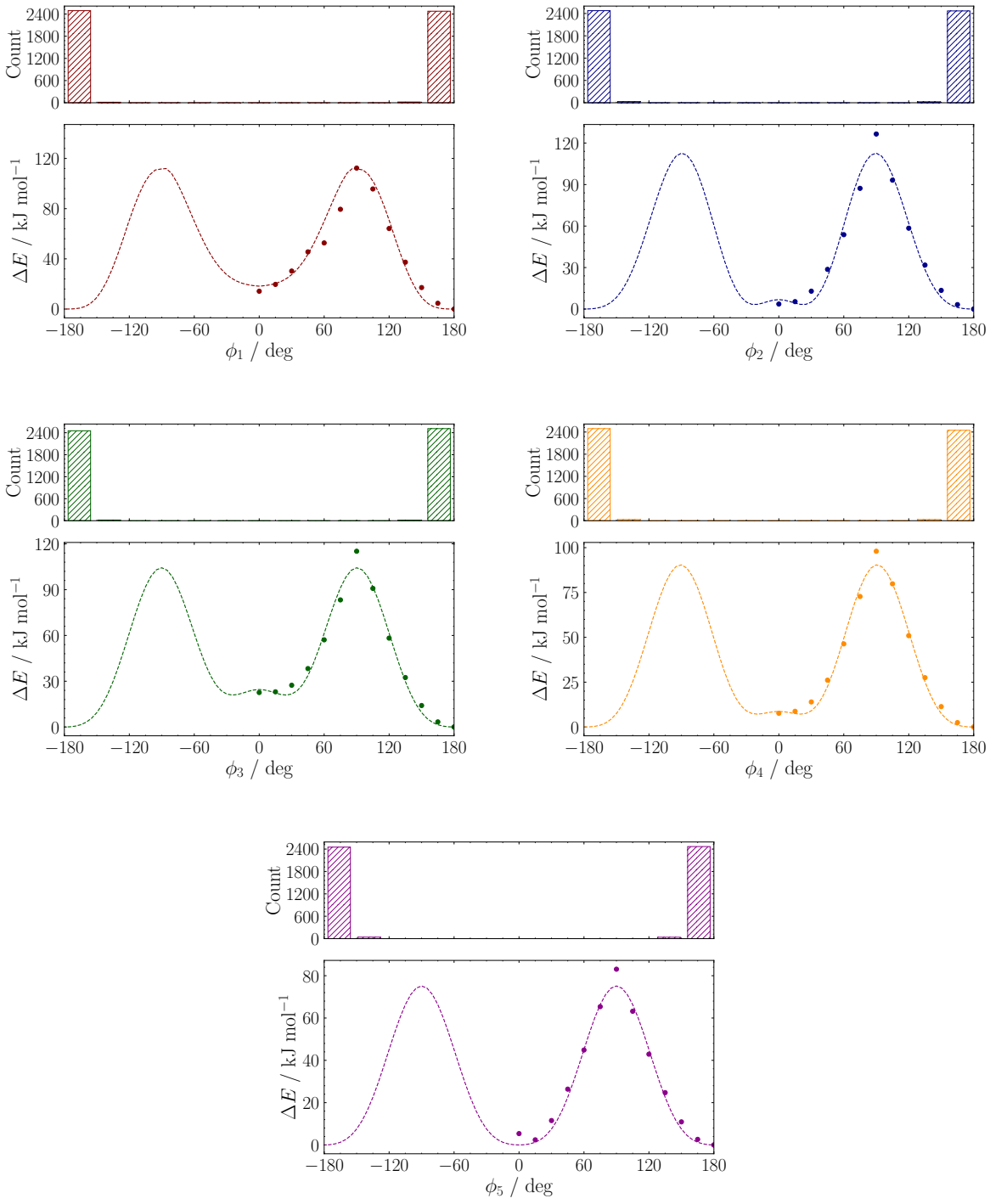


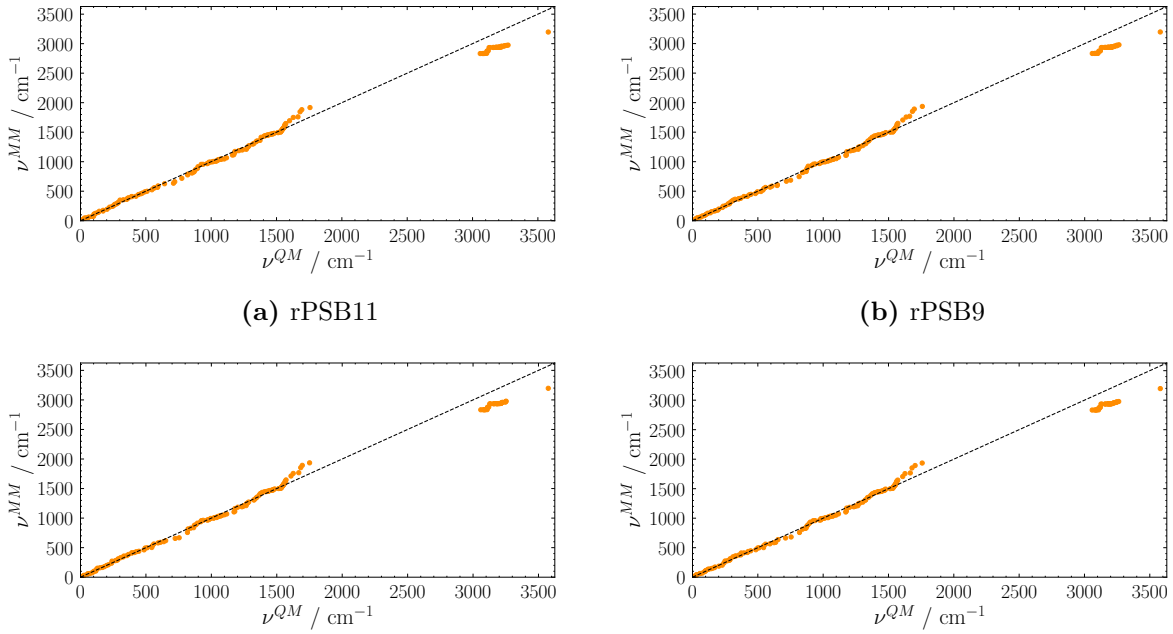
Fig. S9: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSBT, and their distribution along a 500 ps MD run.

S3.2 GAFF

Table S1: RMSD between QM and GAFF optimised geometries in terms of bond lengths, bending angles, dihedrals and molecular normal modes (total).

Molecule	Bond lengths / Å	Bending angles / °	Dihedrals / °	Total / Å
rPSB11	0.02	1.7	18.0	0.63
rPSB9	0.02	1.5	16.0	0.54
rPSB13	0.02	1.6	17.0	0.63
rPSBT	0.02	1.6	17.0	0.61

S3.2.1 Vibrational frequencies



(a) rPSB11 **(b)** rPSB9

(c) rPSB13 **(d)** rPSBT

Fig. S10: Comparison between QM and GAFF vibrational frequencies for rPSBs.

S3.2.2 Torsional profiles

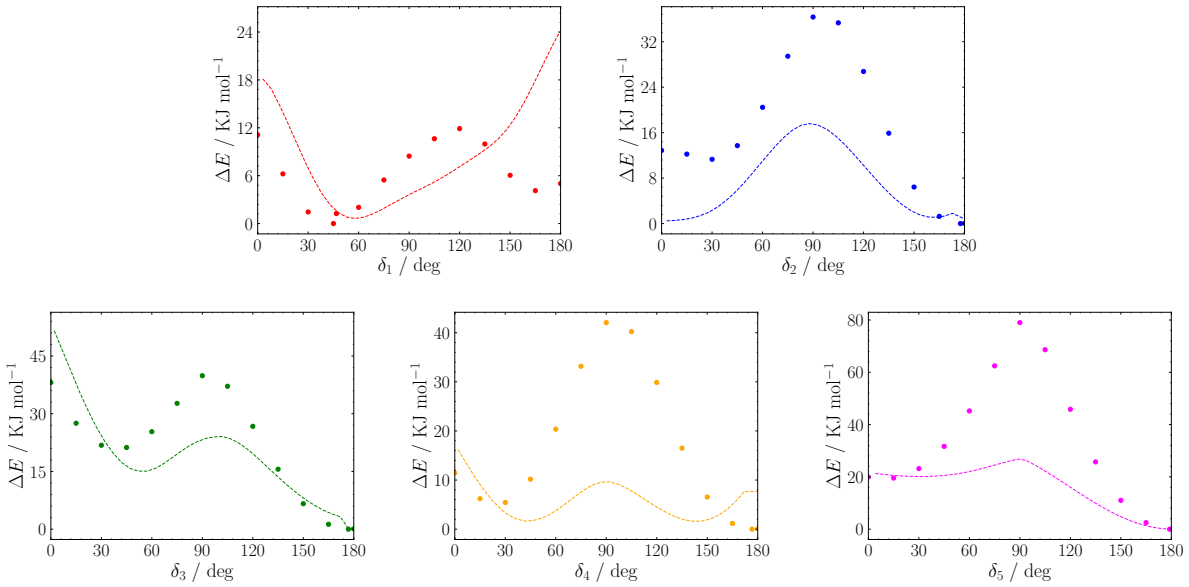


Fig. S11: QM (dots) and GAFF (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSBs.

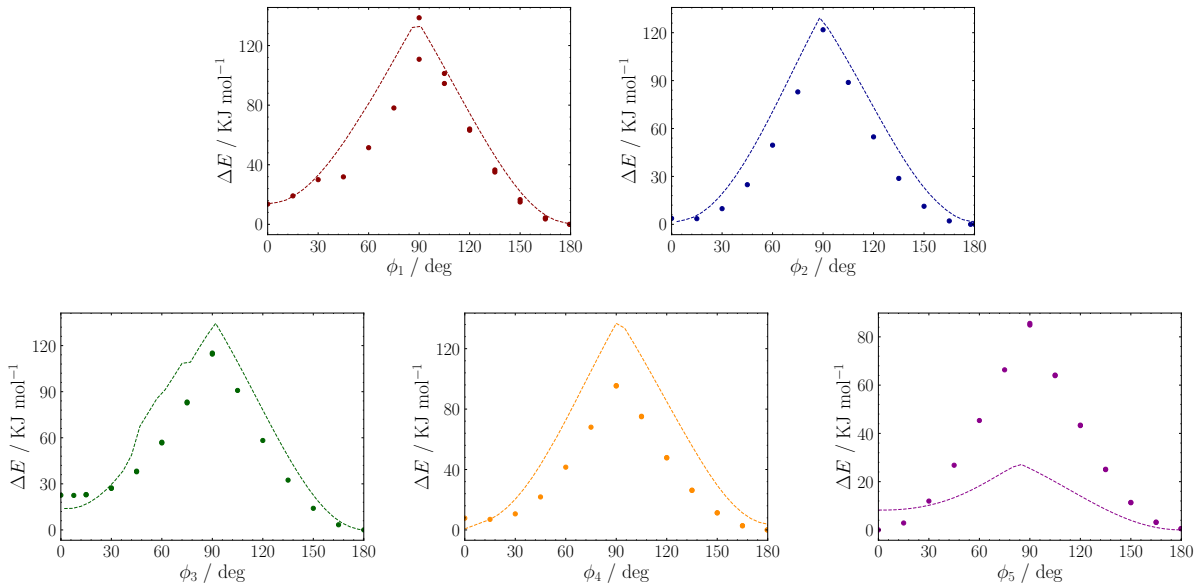


Fig. S12: QM (dots) and GAFF (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSBs.

S3.3 Optical spectra computed on MD trajectories with Berendsen thermostat

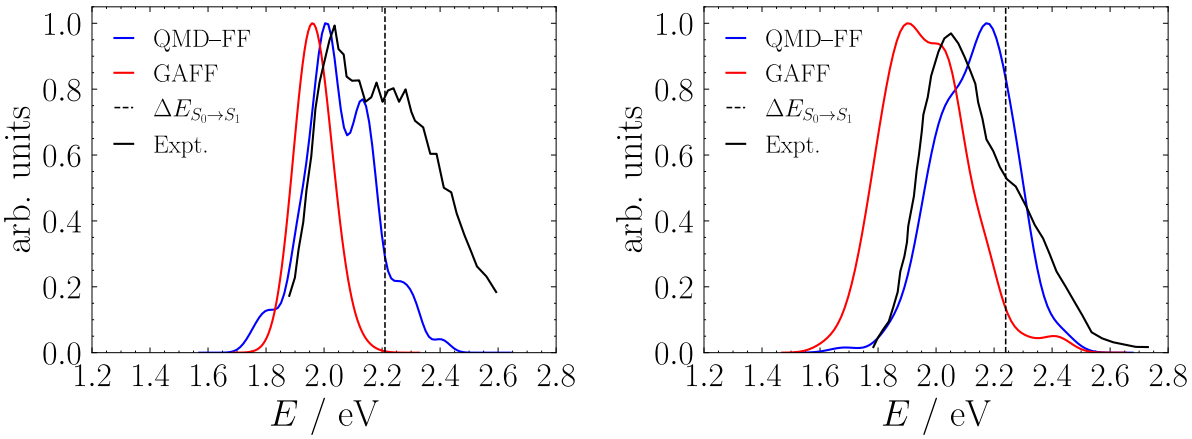


Fig. S13: Comparison between computed (RMS–CASPT2/cc–pVDZ) and experimental absorption spectra *in vacuo* for rPSB11 (left) and rPSBT (right) on trajectories using the Berenden thermostat. The vertical dashed lines represent RMS–CASPT2/cc–pVDZ vertical excitation energies computed on MP2/6–31G(d) equilibrium geometries. All spectra have been normalised with respect to their maximum intensity. Experimental spectra digitalised from refs. 10 and 11.

S3.4 Distributions of δ_n along 25 ns trajectories used to compute optical spectra

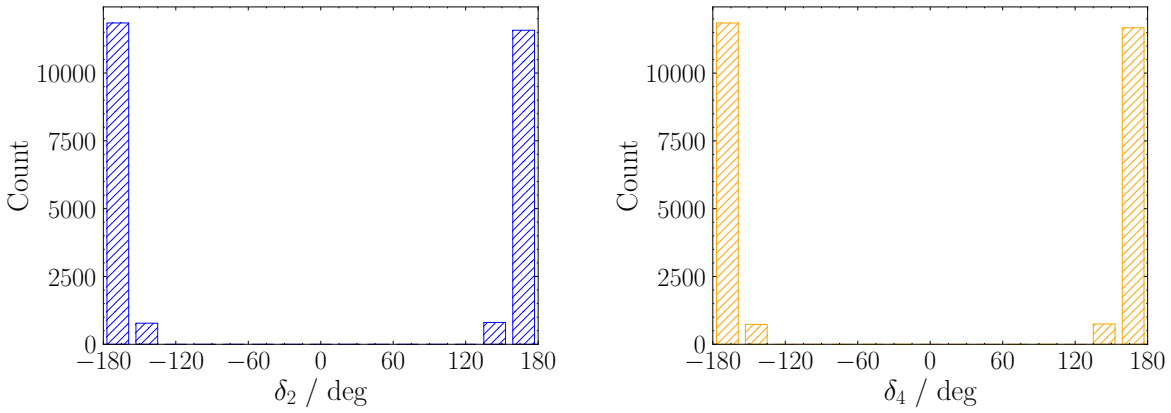


Fig. S14: Distribution of δ_2 and δ_4 flexible dihedrals along a 25 ns MD run for rPSB11 using the QMD–FF.

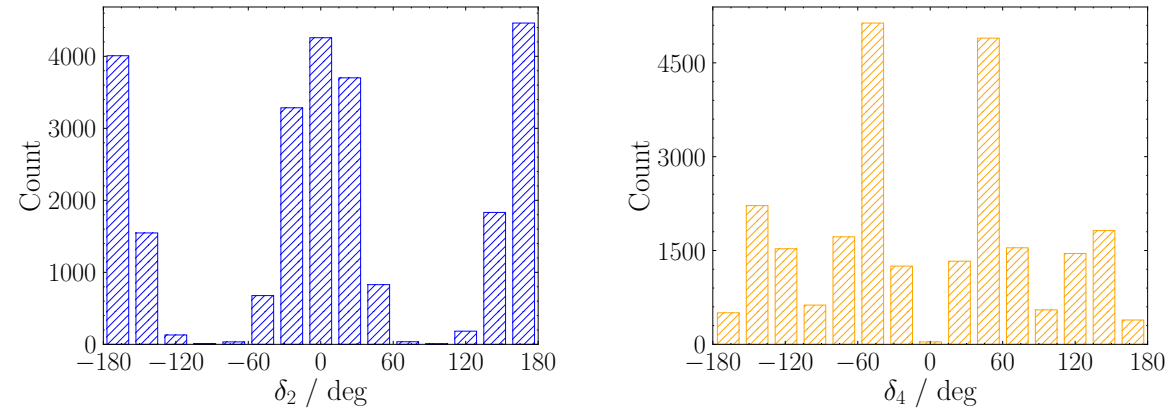


Fig. S15: Distribution of δ_2 and δ_4 flexible dihedrals along a 25 ns MD run for rPSB11 using GAFF.

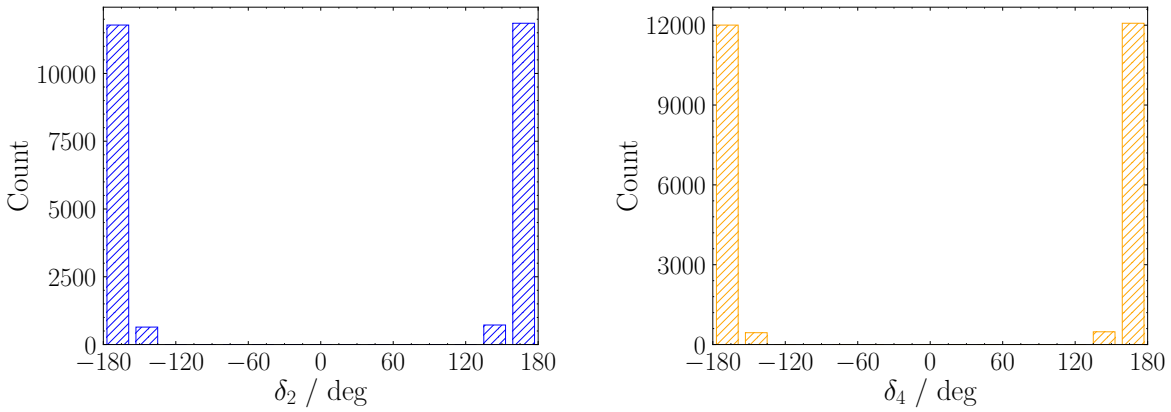


Fig. S16: Distribution of δ_2 and δ_4 flexible dihedrals along a 25 ns MD run for rPSBT using the QMD–FF.

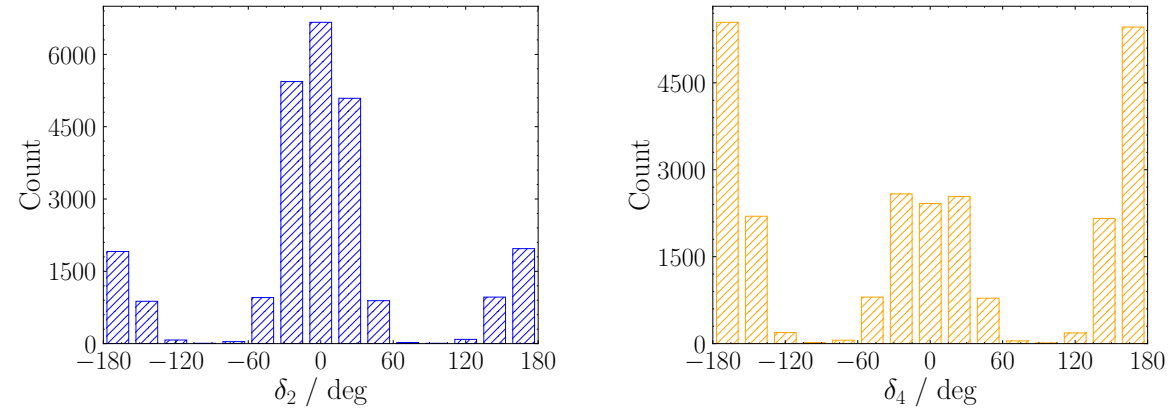
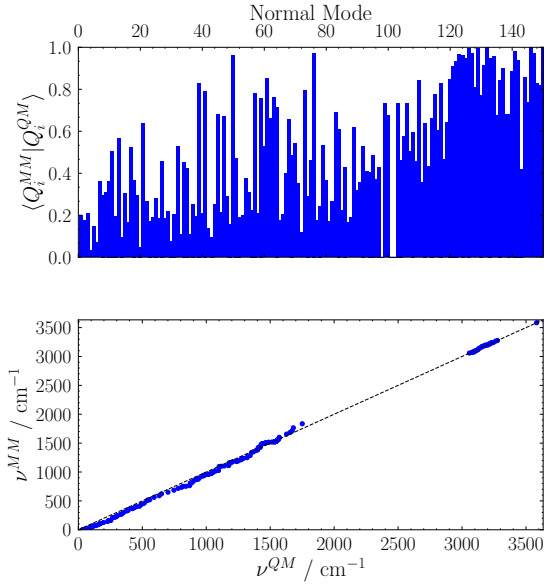


Fig. S17: Distribution of δ_2 and δ_4 flexible dihedrals along a 25 ns MD run for rPSBT using GAFF.

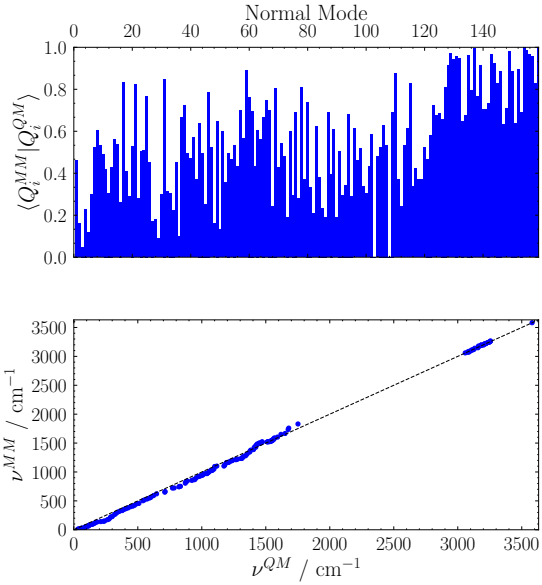
S4 Ring–locked retinals

S4.1 QMD–FFs

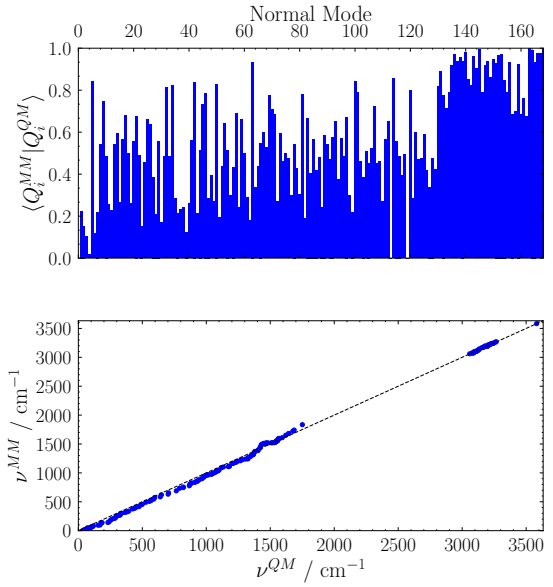
S4.1.1 Vibrational modes and frequencies



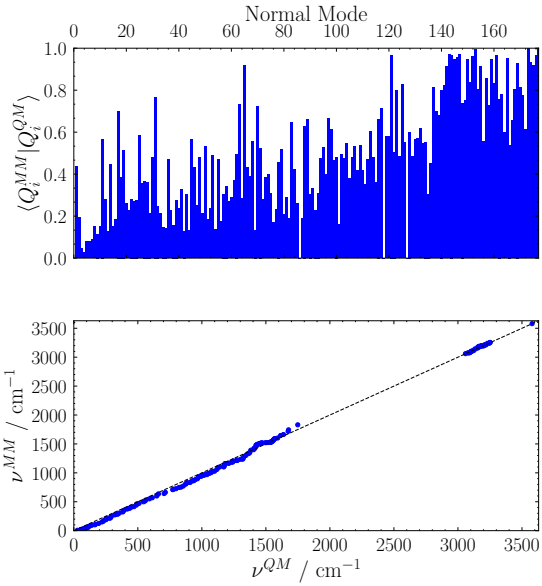
(a) rPSB11.5



(b) rPSB11.6



(c) rPSB11.7



(d) rPSB11.8

Fig. S18: Comparison between QM and MM vibrational frequencies and overlap between QM and MM normal modes for ring–locked rPSBs.

S4.1.2 Torsional Profiles

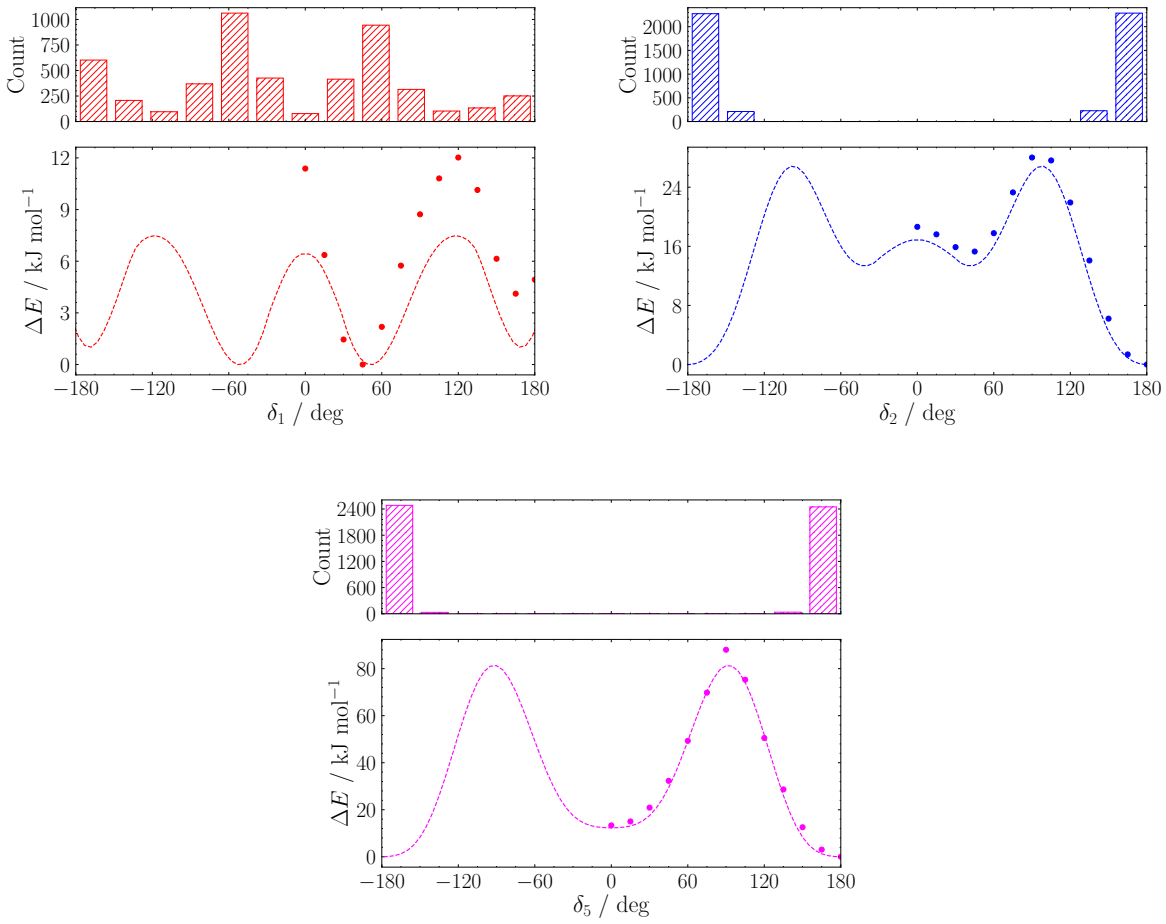


Fig. S19: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB11.5, and their distribution along a 500 ps MD run.

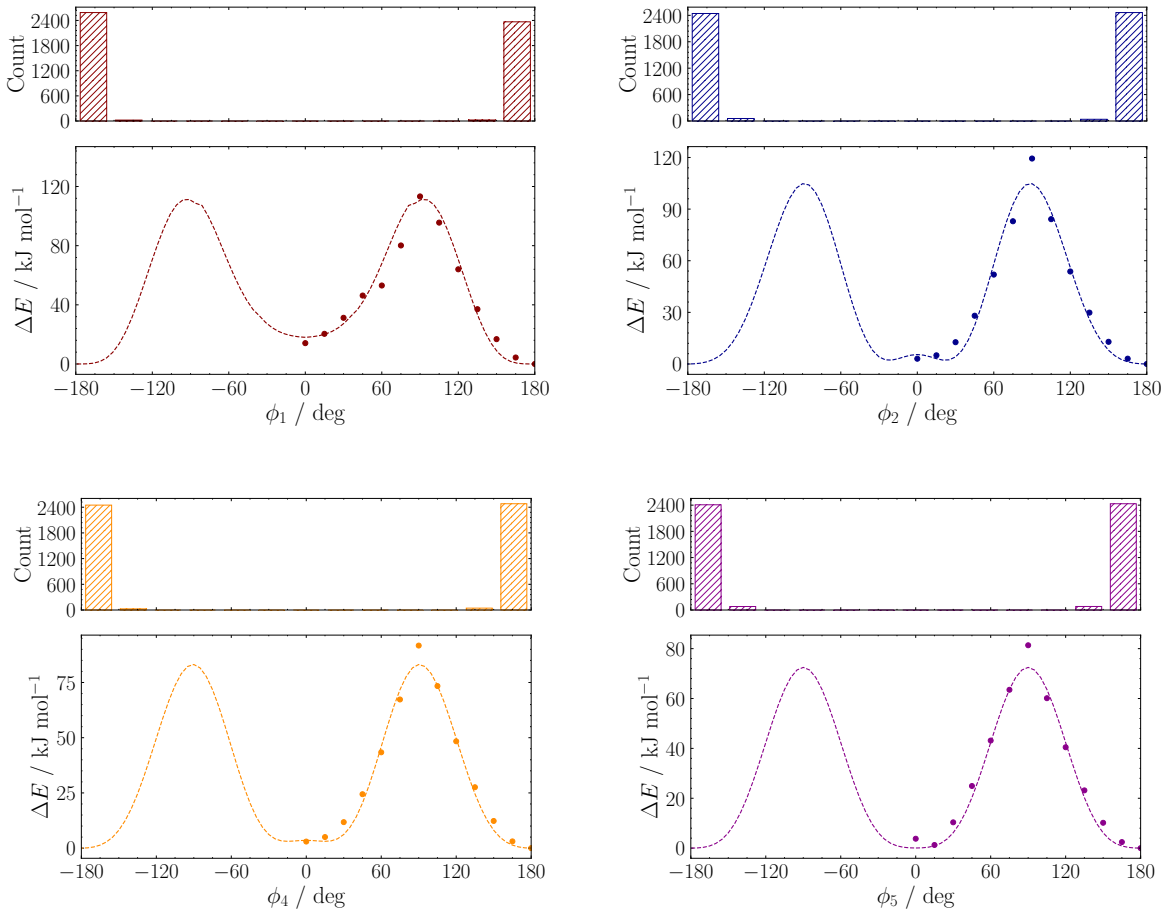


Fig. S20: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11.5, and their distribution along a 500 ps MD run.

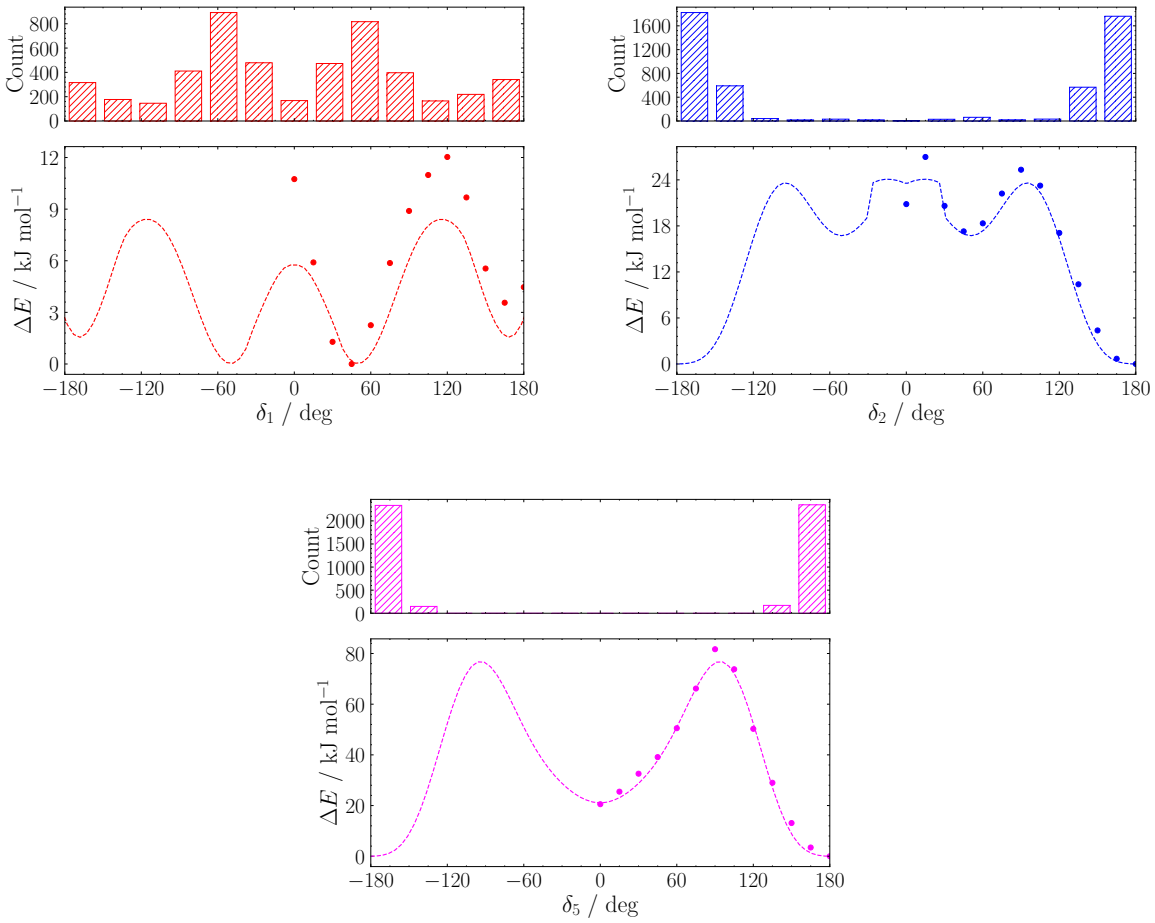


Fig. S21: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB11.6, and their distribution along a 500 ps MD run.

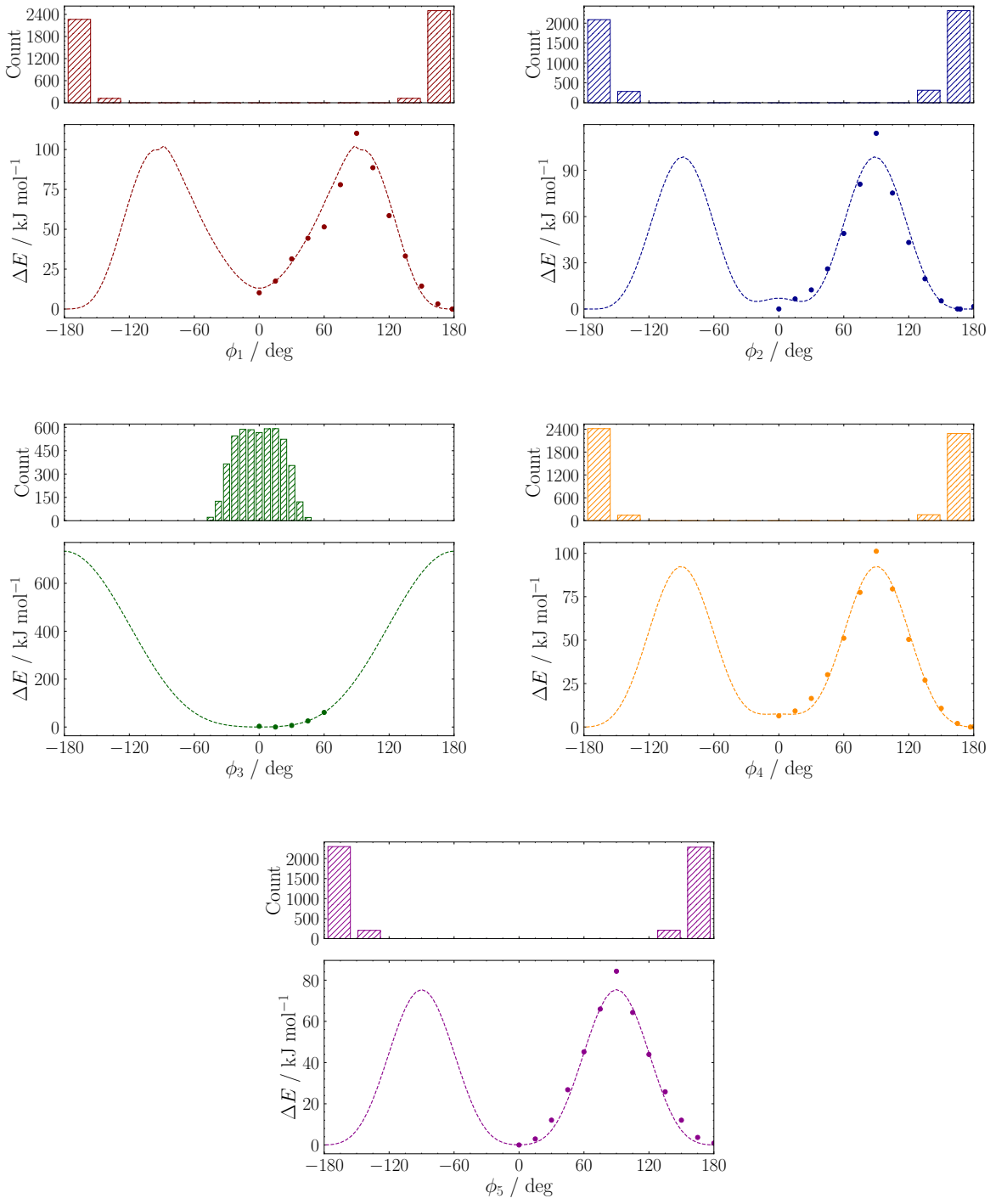


Fig. S22: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11.6, and their distribution along a 500 ps MD run.

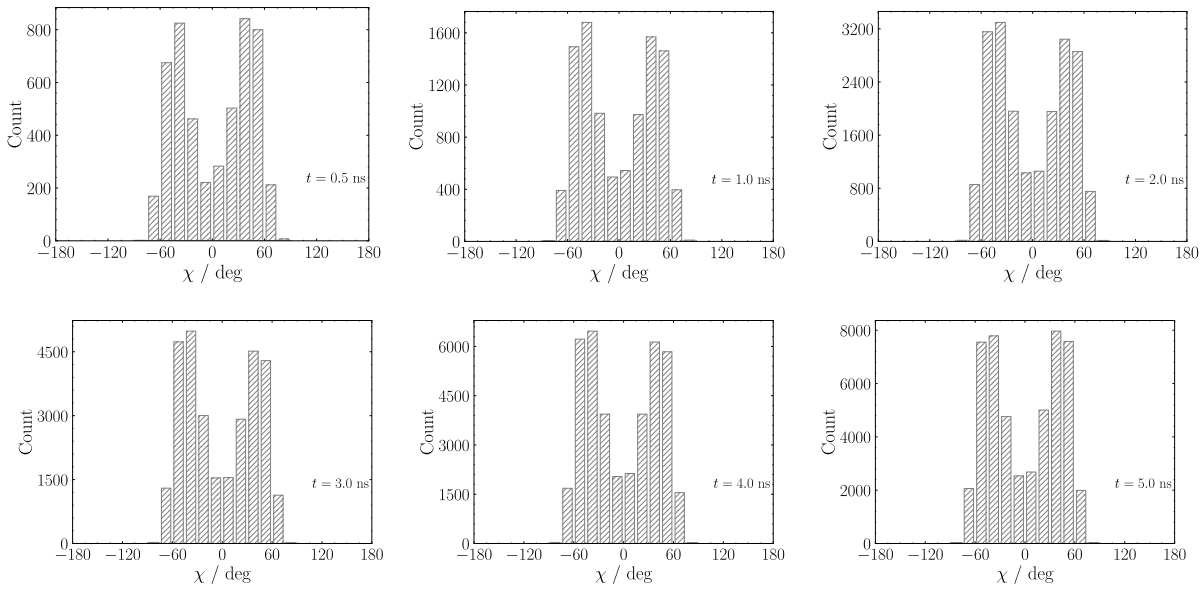


Fig. S23: Distribution of the dihedral χ along a 5 ns trajectory for rPSB11.6, to monitor convergence of the conformational ensemble.

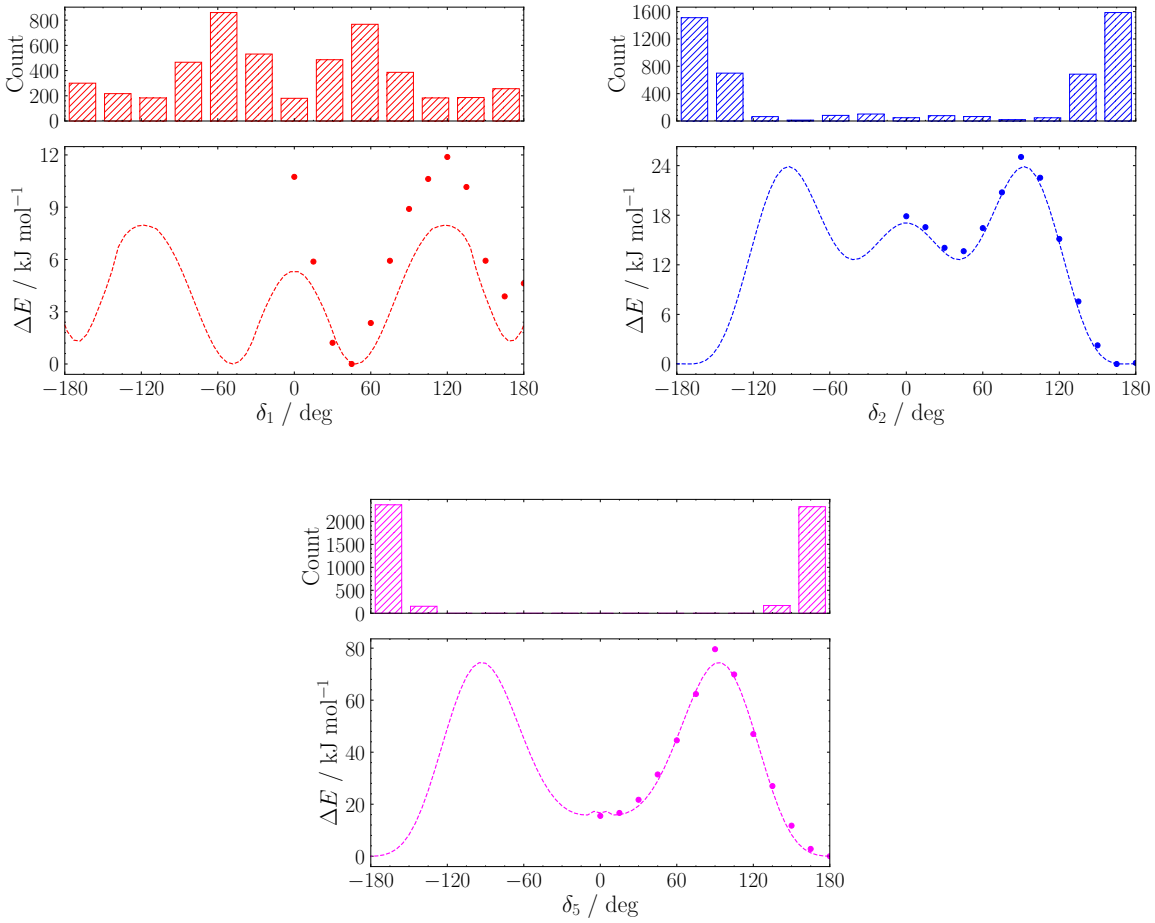


Fig. S24: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB11.7, and their distribution along a 500 ps MD run.

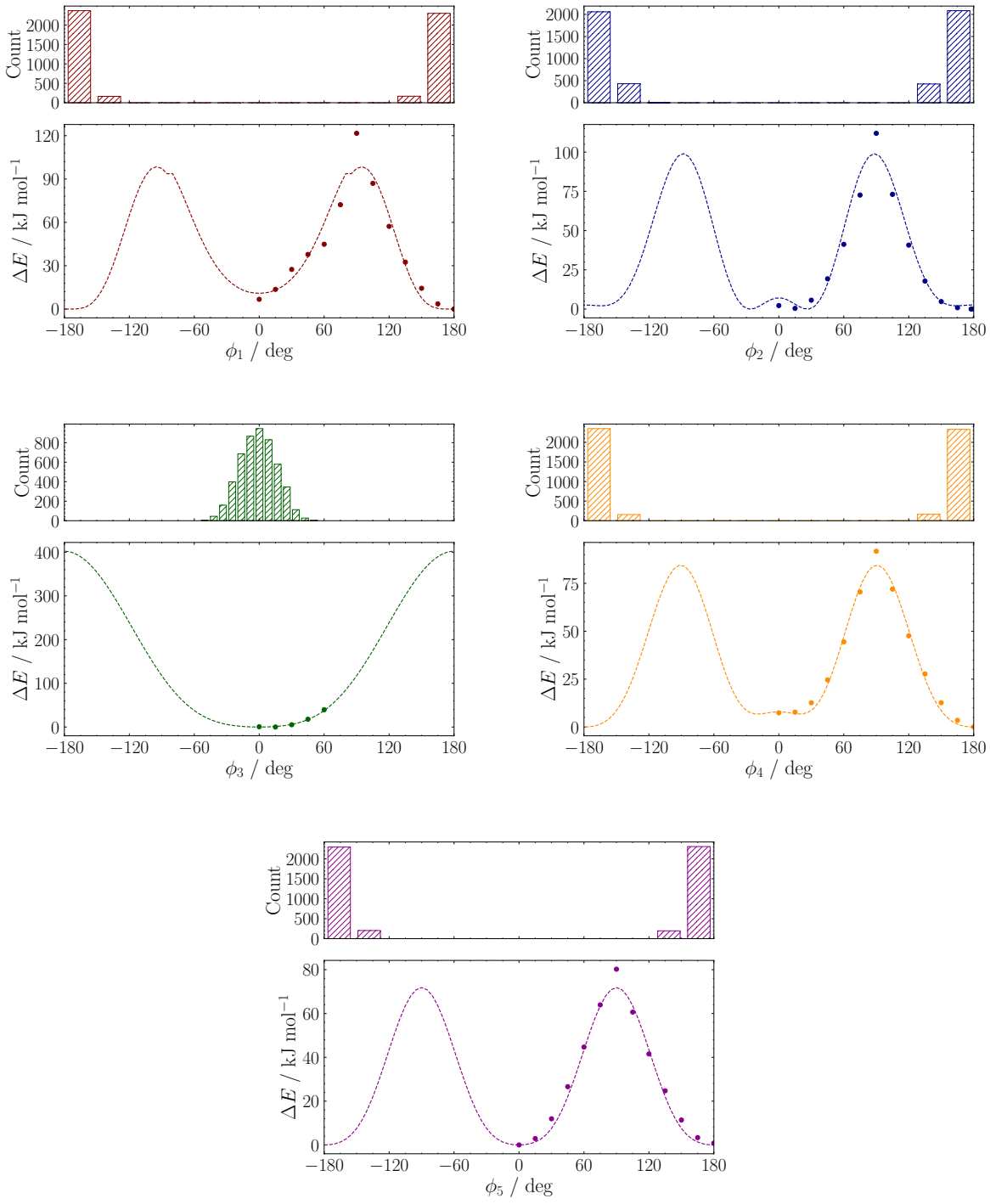


Fig. S25: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11.7, and their distribution along a 500 ps MD run.

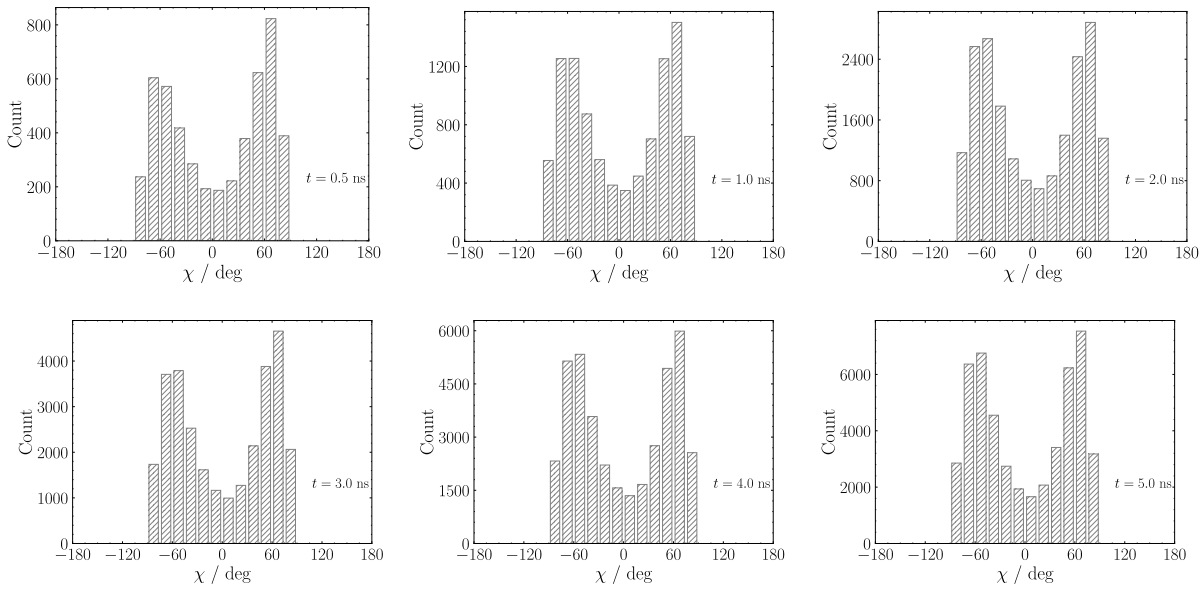


Fig. S26: Distribution of the dihedral χ along a 5 ns trajectory for rPSB11.7, to monitor convergence of the conformational ensemble.

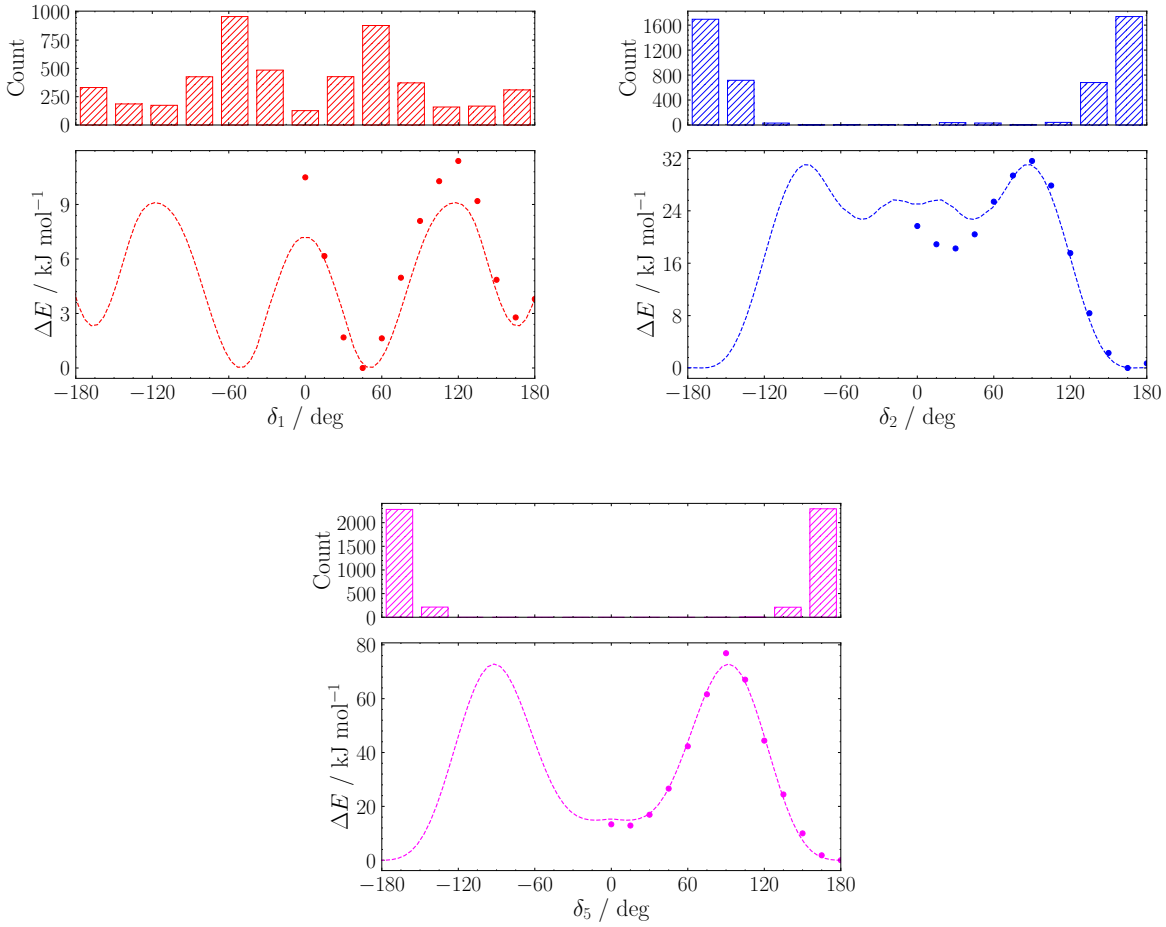


Fig. S27: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB11.8, and their distribution along a 500 ps MD run.

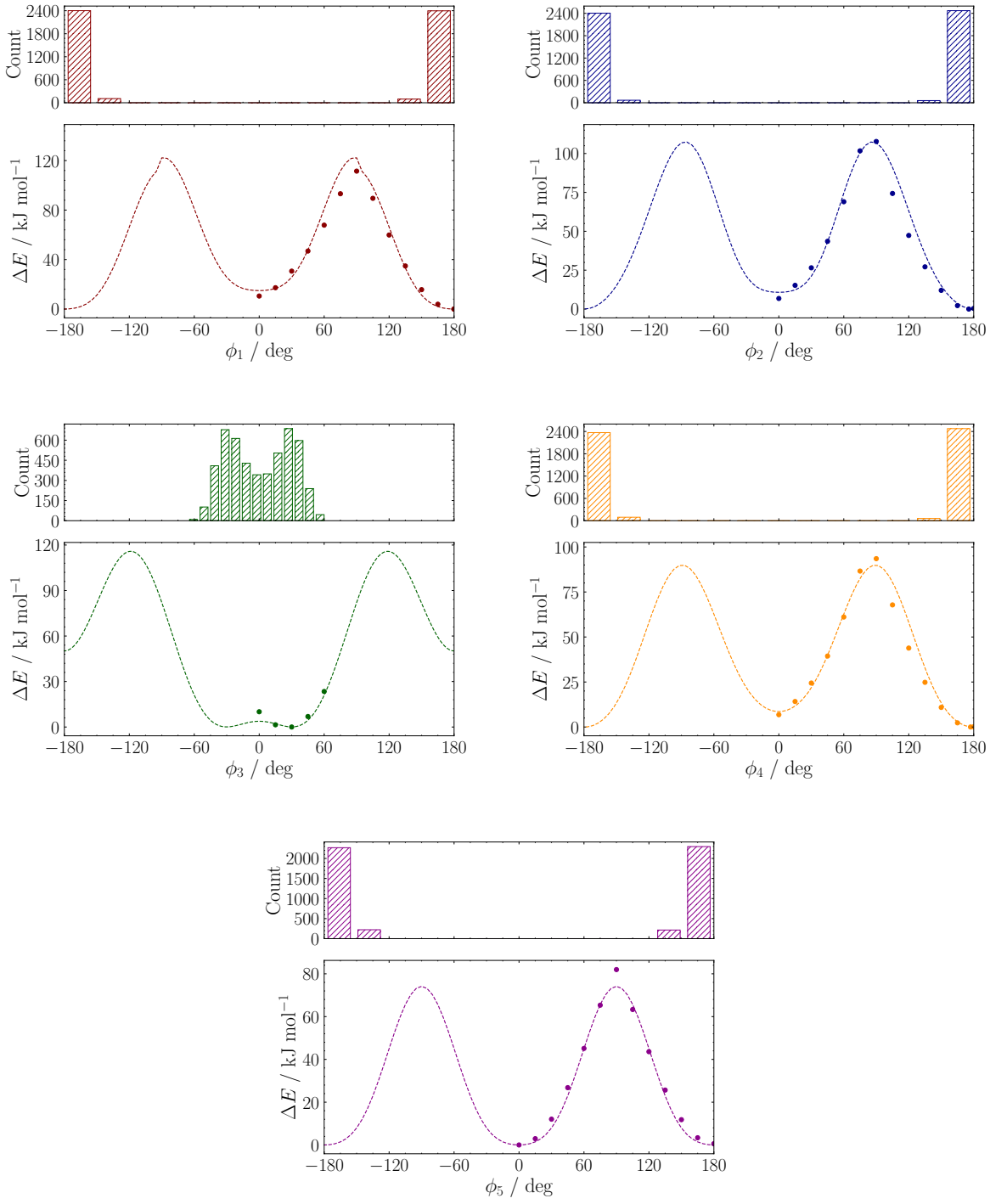
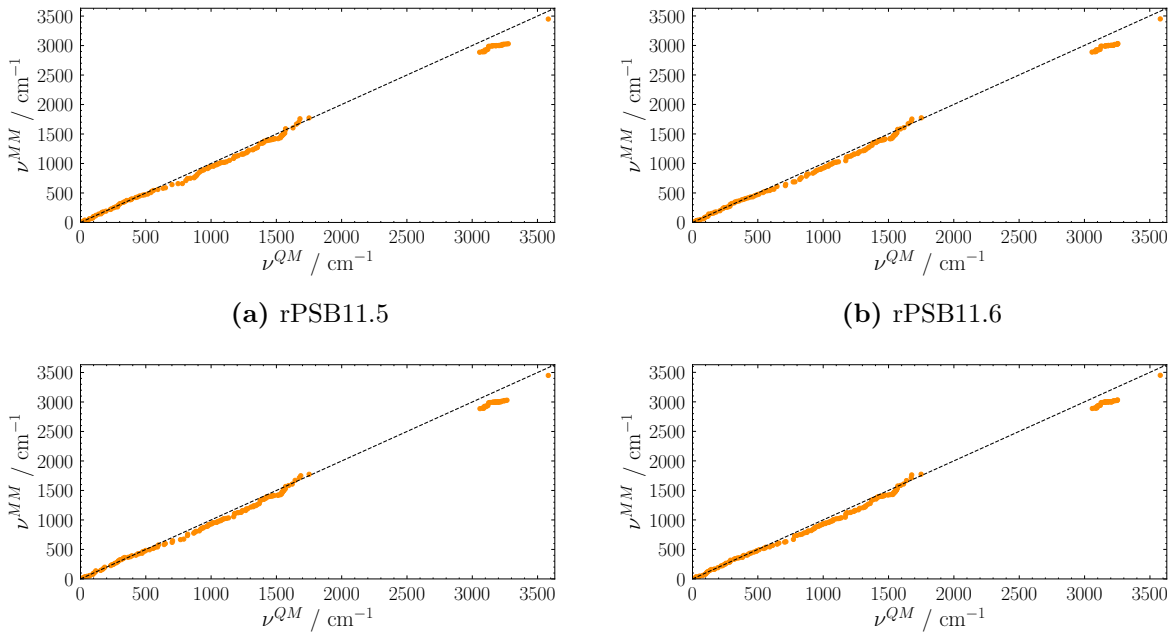


Fig. S28: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11.8, and their distribution along a 500 ps MD run.

S4.2 GAFF

S4.2.1 Vibrational frequencies



(a) rPSB11.5 **(b)** rPSB11.6

(c) rPSB11.7 **(d)** rPSB11.8

Fig. S29: Comparison between QM and GAFF vibrational frequencies for ring–locked rPSBs.

Table S2: RMSD between QM and MM (GAFF) optimised geometries in terms of bond lengths, bending angles, dihedrals and molecular normal modes (total).

Molecule	Bond lengths / Å	Bending angles / °	Dihedrals / °	Total / Å
rPSB11.5	0.02	1.9	13.8	0.65
rPSB11.6	0.02	1.4	15.8	0.83
rPSB11.7	0.02	1.6	15.9	1.07
rPSB11.8	0.02	1.5	14.2	0.75

S4.2.2 Torsional Profiles

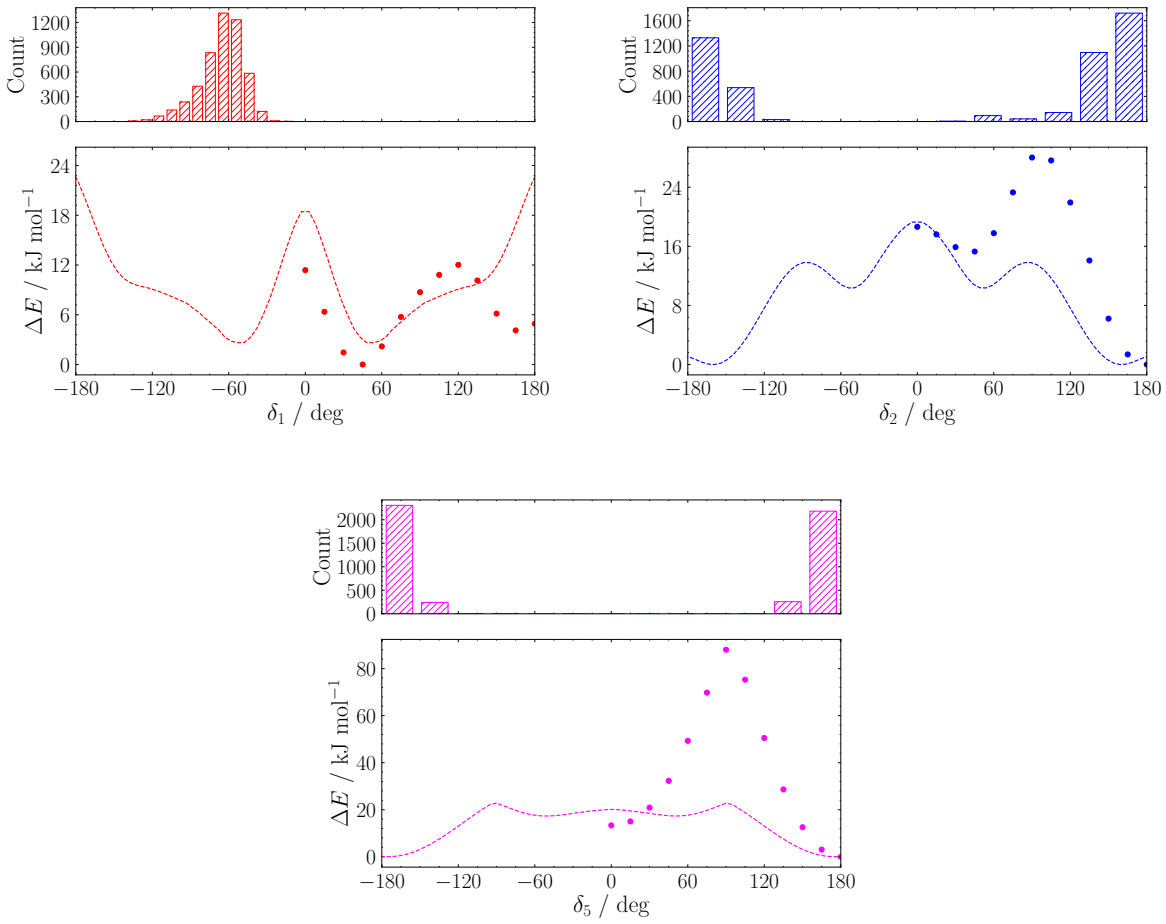


Fig. S30: QM (dots) and GAFF (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB11.5, and their distribution along a 500 ps MD run.

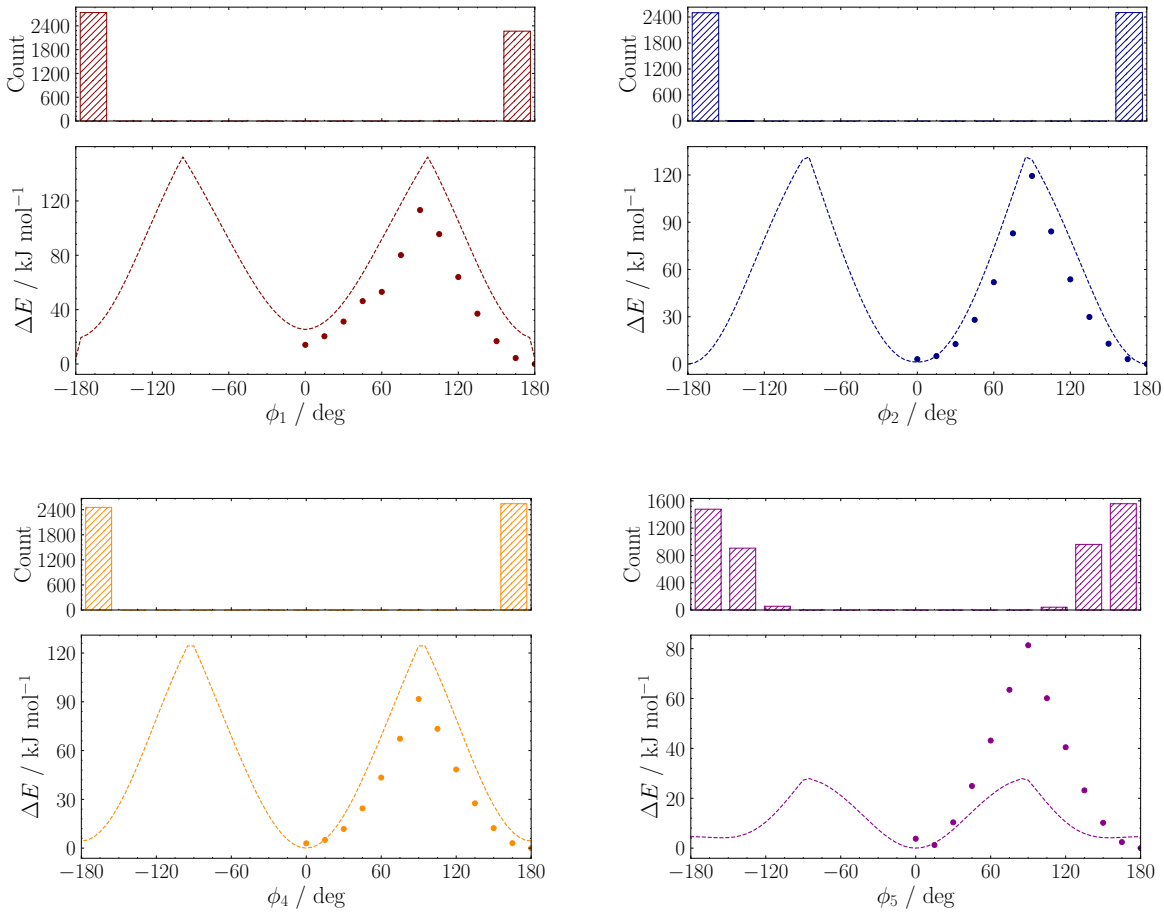


Fig. S31: QM (dots) and GAFF (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11.5, and their distribution along a 500 ps MD run.

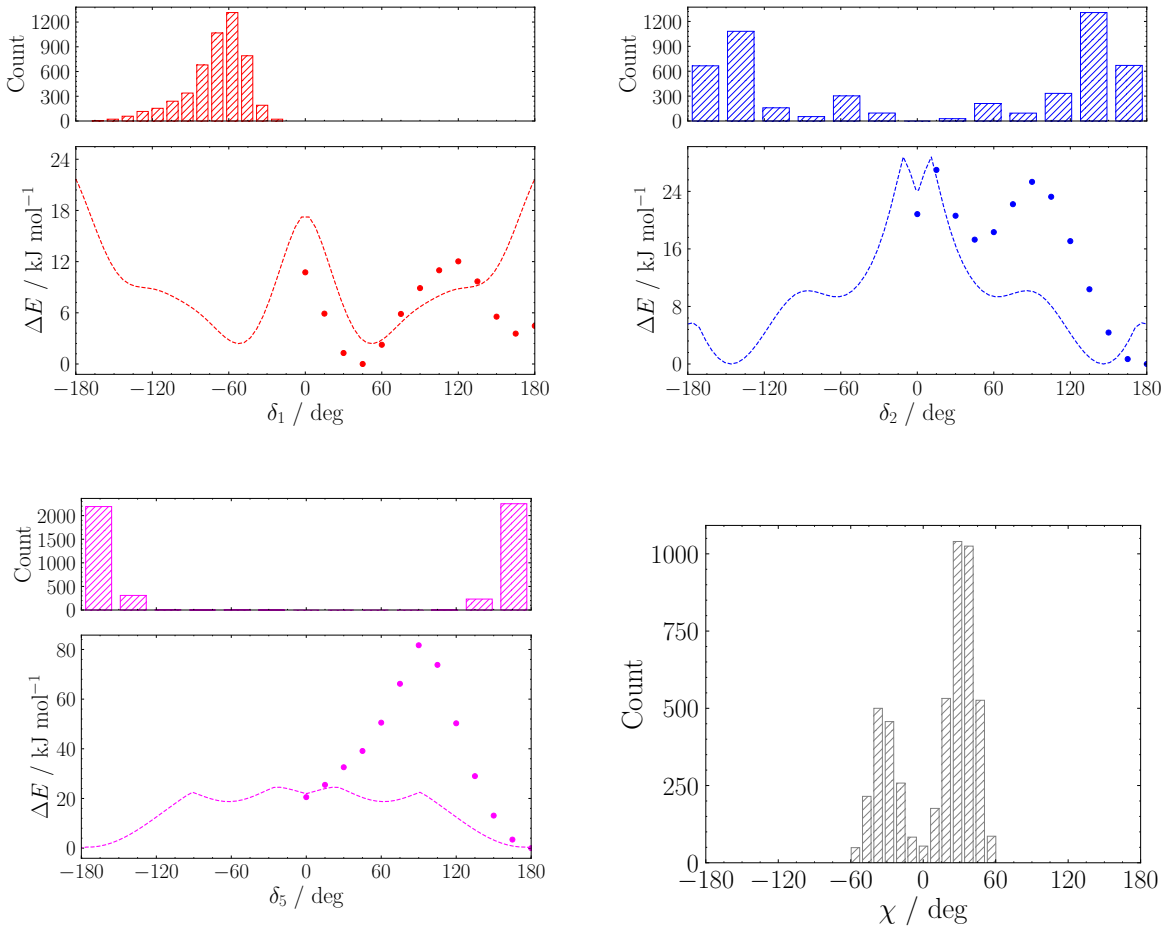


Fig. S32: QM (dots) and GAFF (dashed lines) potential energy profiles for δ_1–δ_5 and χ flexible dihedrals for rPSB11.6, and their distribution along a 500 ps MD run.

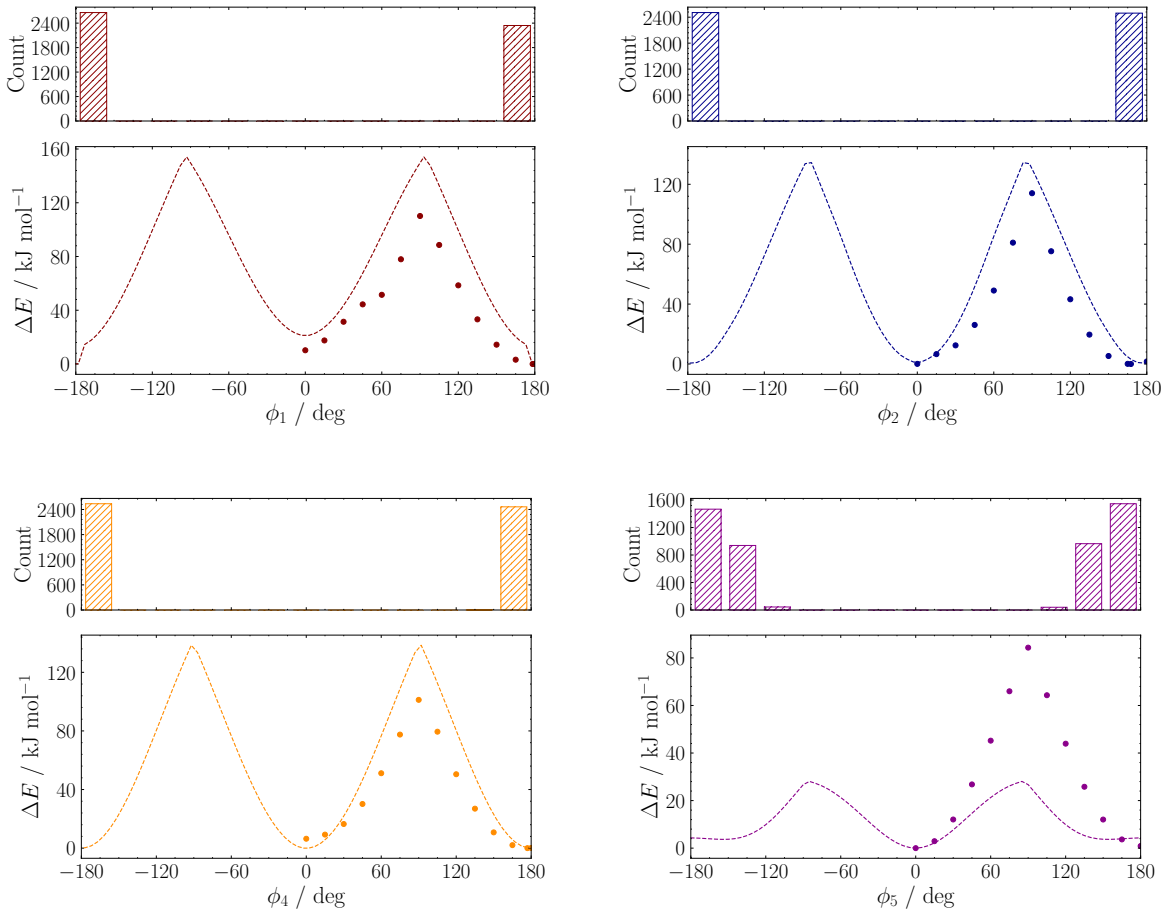


Fig. S33: QM (dots) and GAFF (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11.6, and their distribution along a 500 ps MD run.

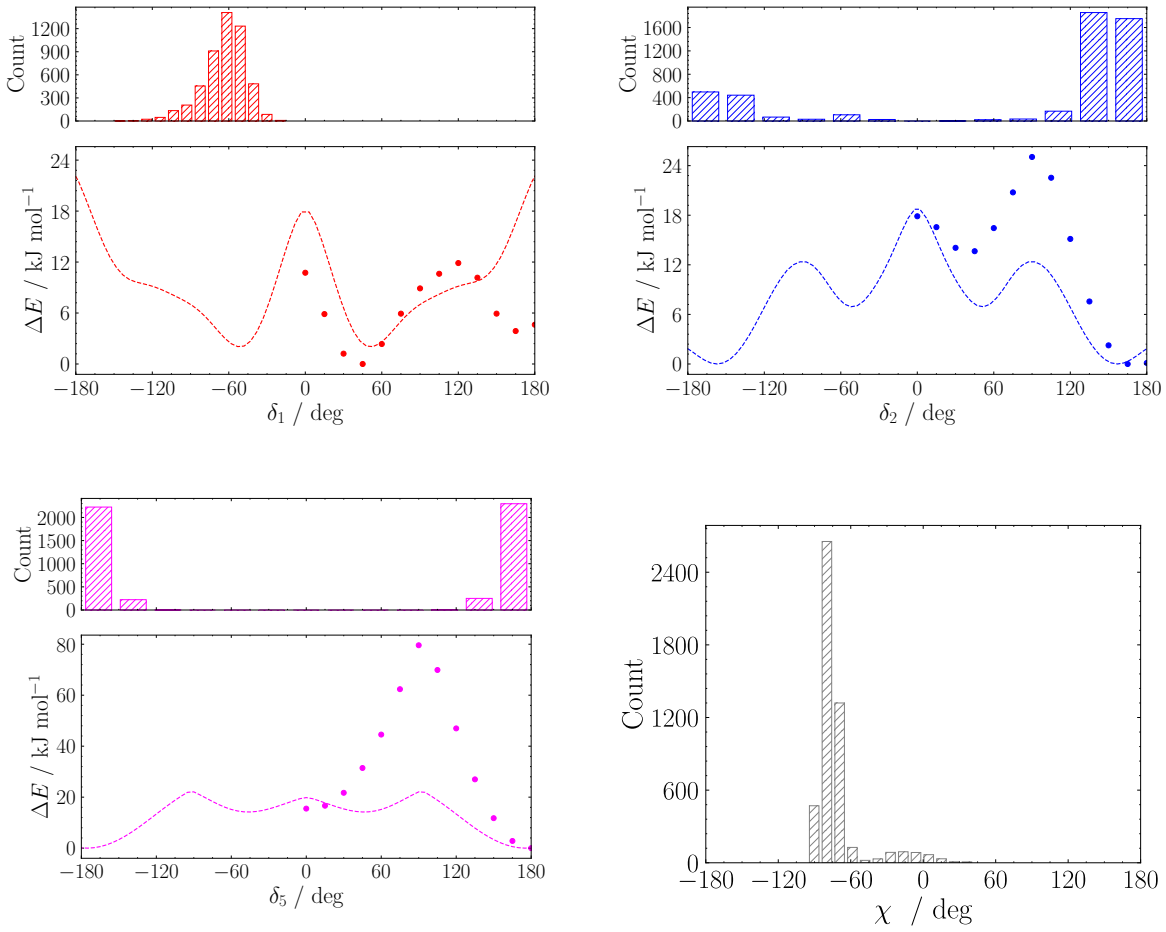


Fig. S34: QM (dots) and GAFF (dashed lines) potential energy profiles for δ_1–δ_5 flexible dihedrals for rPSB11.7, and their distribution along a 500 ps MD run.

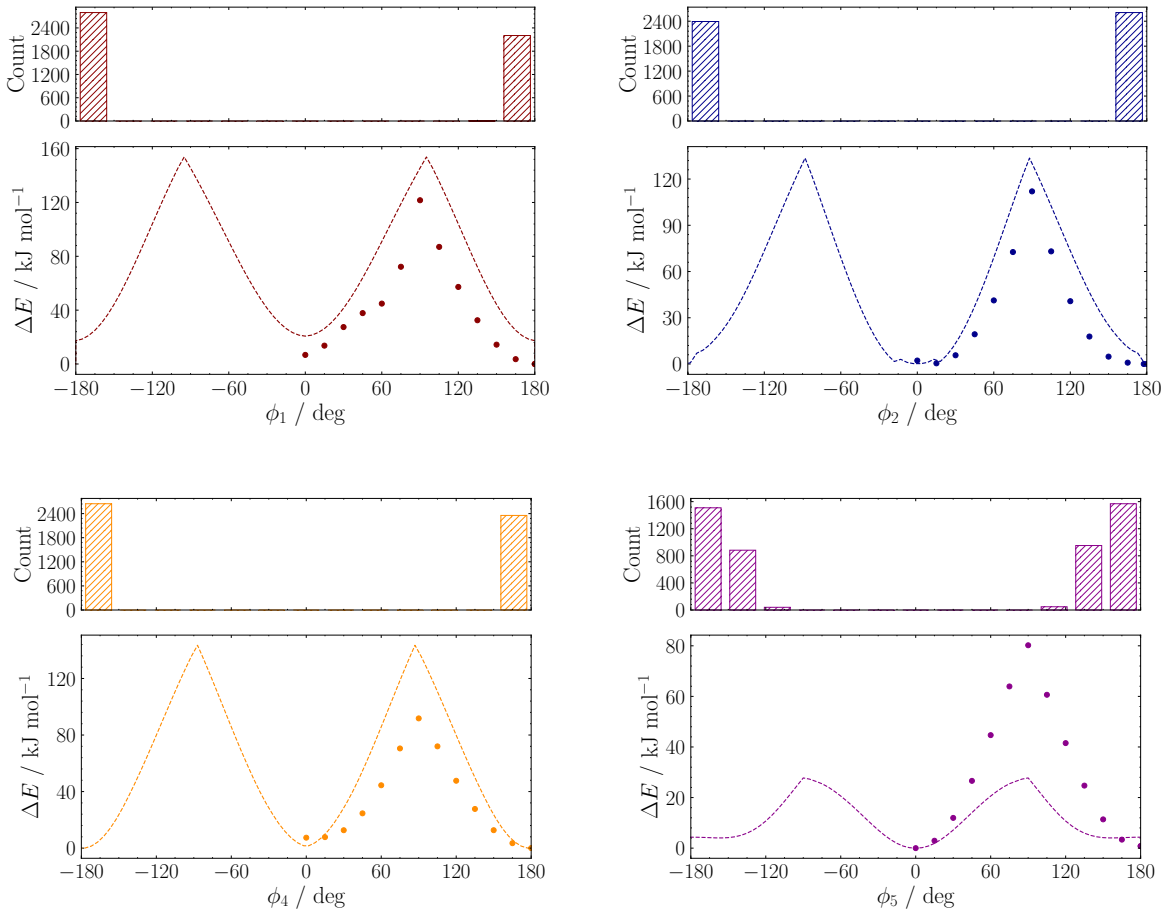


Fig. S35: QM (dots) and GAFF (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11.7, and their distribution along a 500 ps MD run.

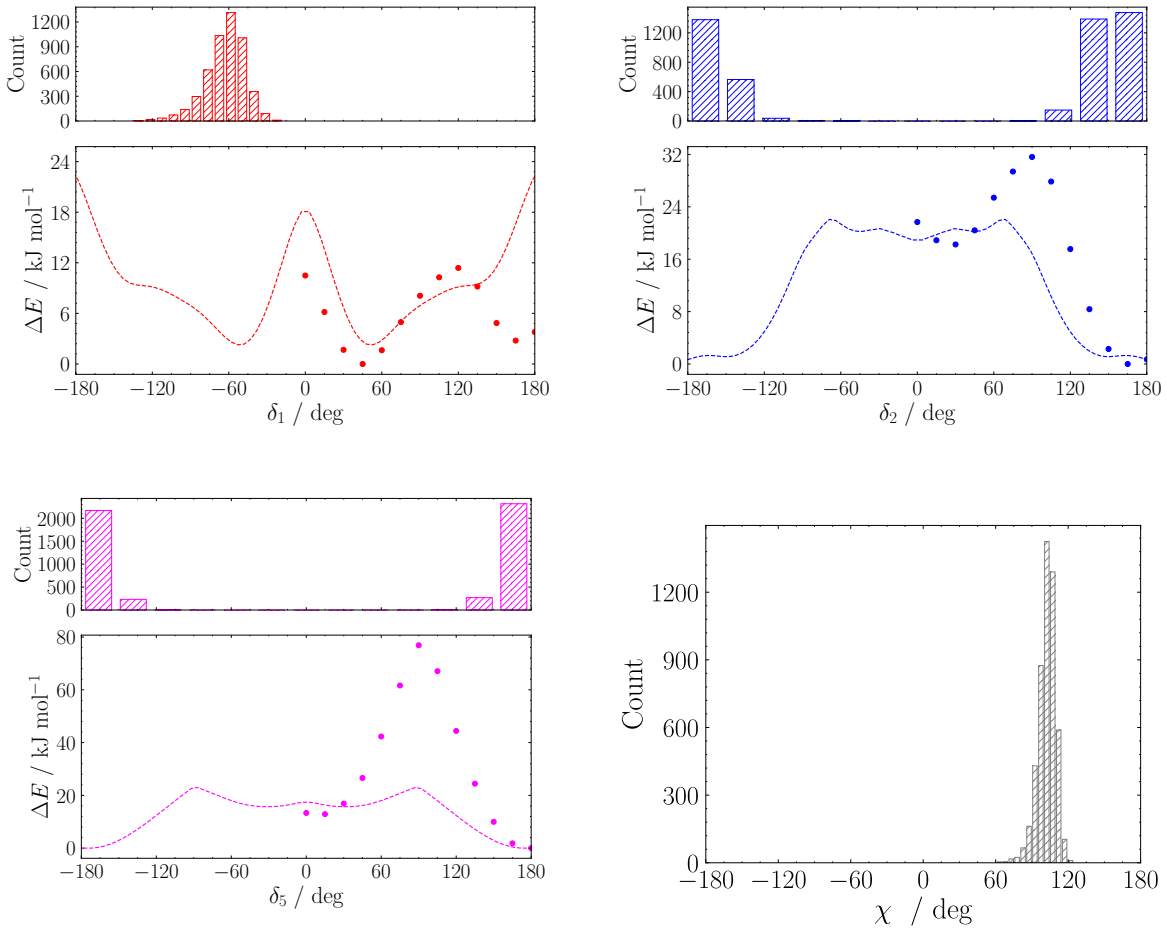


Fig. S36: QM (dots) and GAFF (dashed lines) potential energy profiles for δ_1–δ_5 and χ flexible dihedrals for rPSB11.8, and their distribution along a 500 ps MD run.

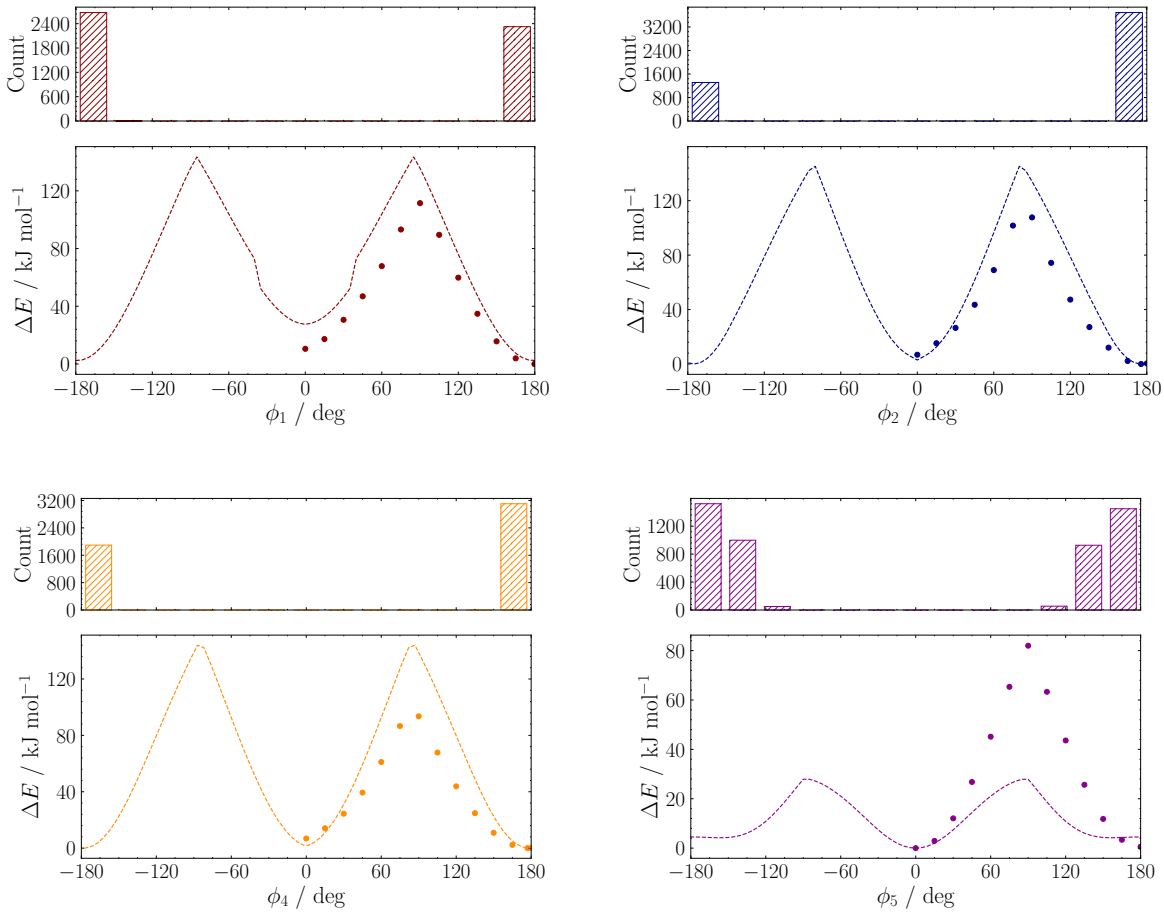
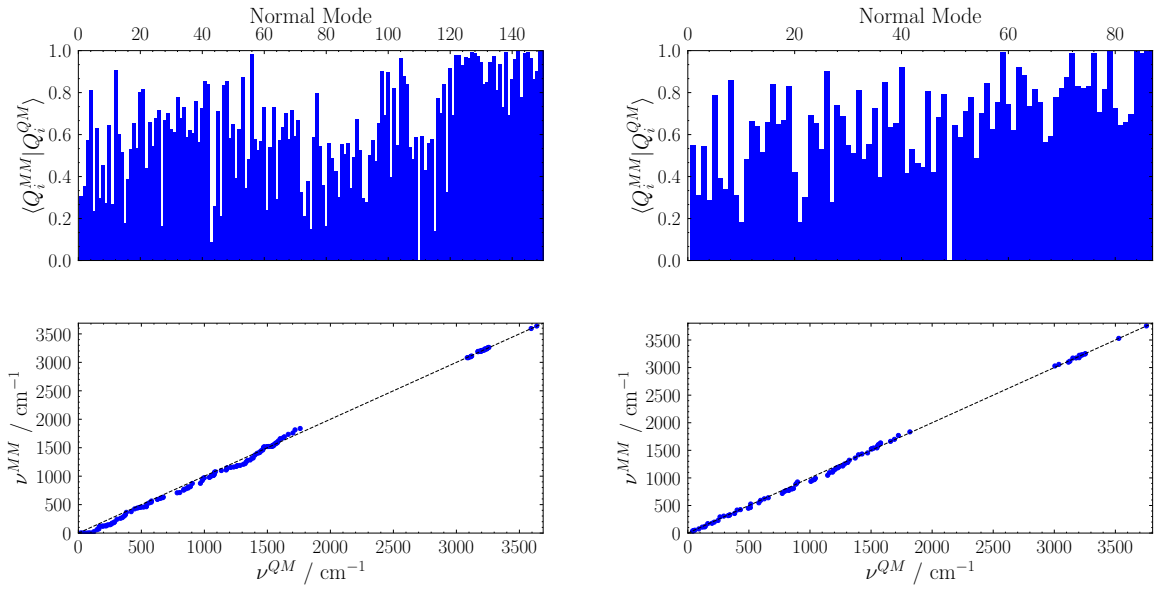


Fig. S37: QM (dots) and GAFF (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for rPSB11.8, and their distribution along a 500 ps MD run.

S5 Artificial retinal analogues

S5.1 QMD–FFs

S5.1.1 Vibrational modes and frequencies



(a) MMAR **(b)** MAMHPA

Fig. S38: Comparison between QM and MM vibrational frequencies and overlap between QM and MM normal modes for artificial retinal analogues.

S5.1.2 Torsional Profiles

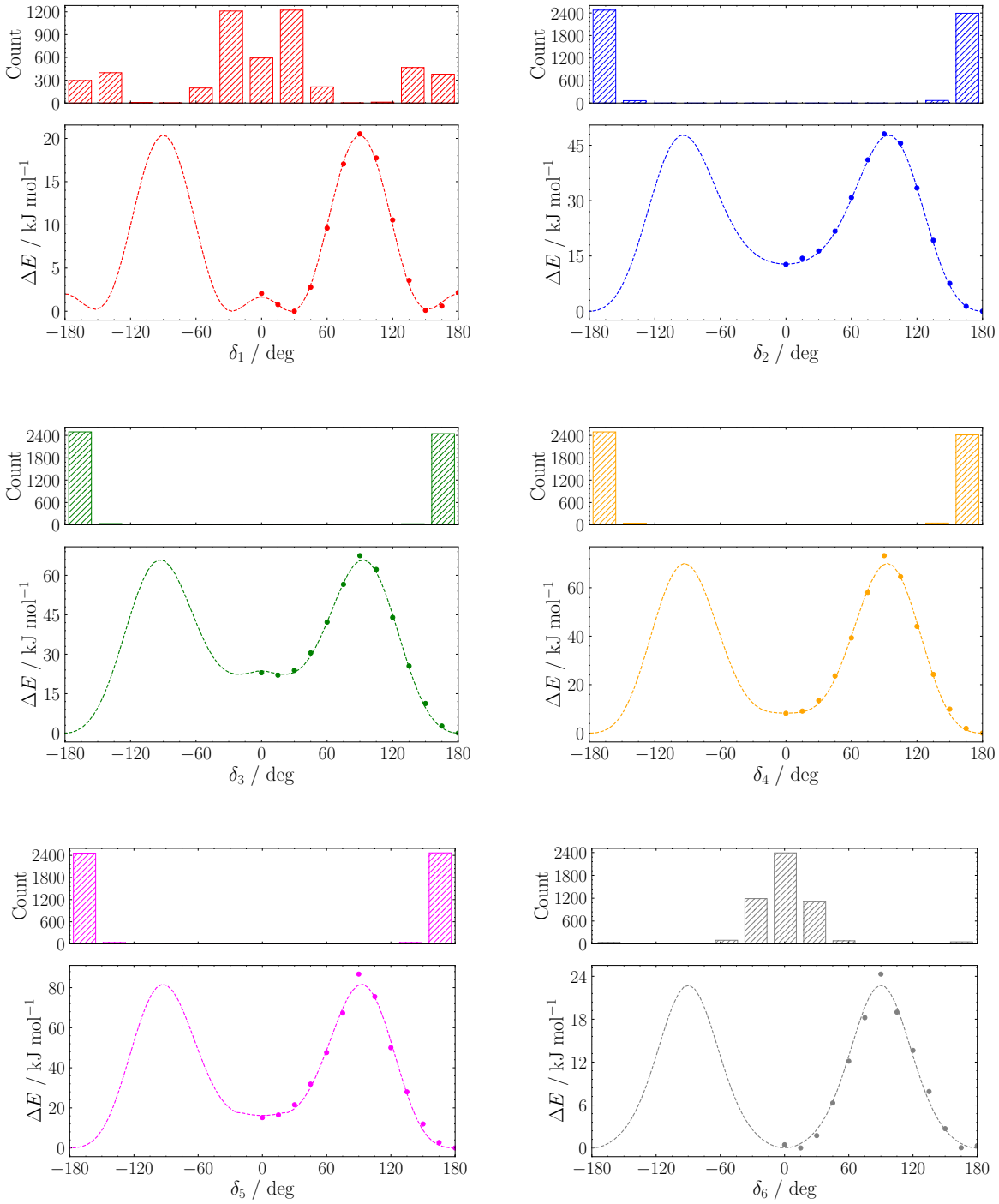


Fig. S39: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_6 flexible dihedrals for MMAR, and their distribution along a 500 ps MD run.

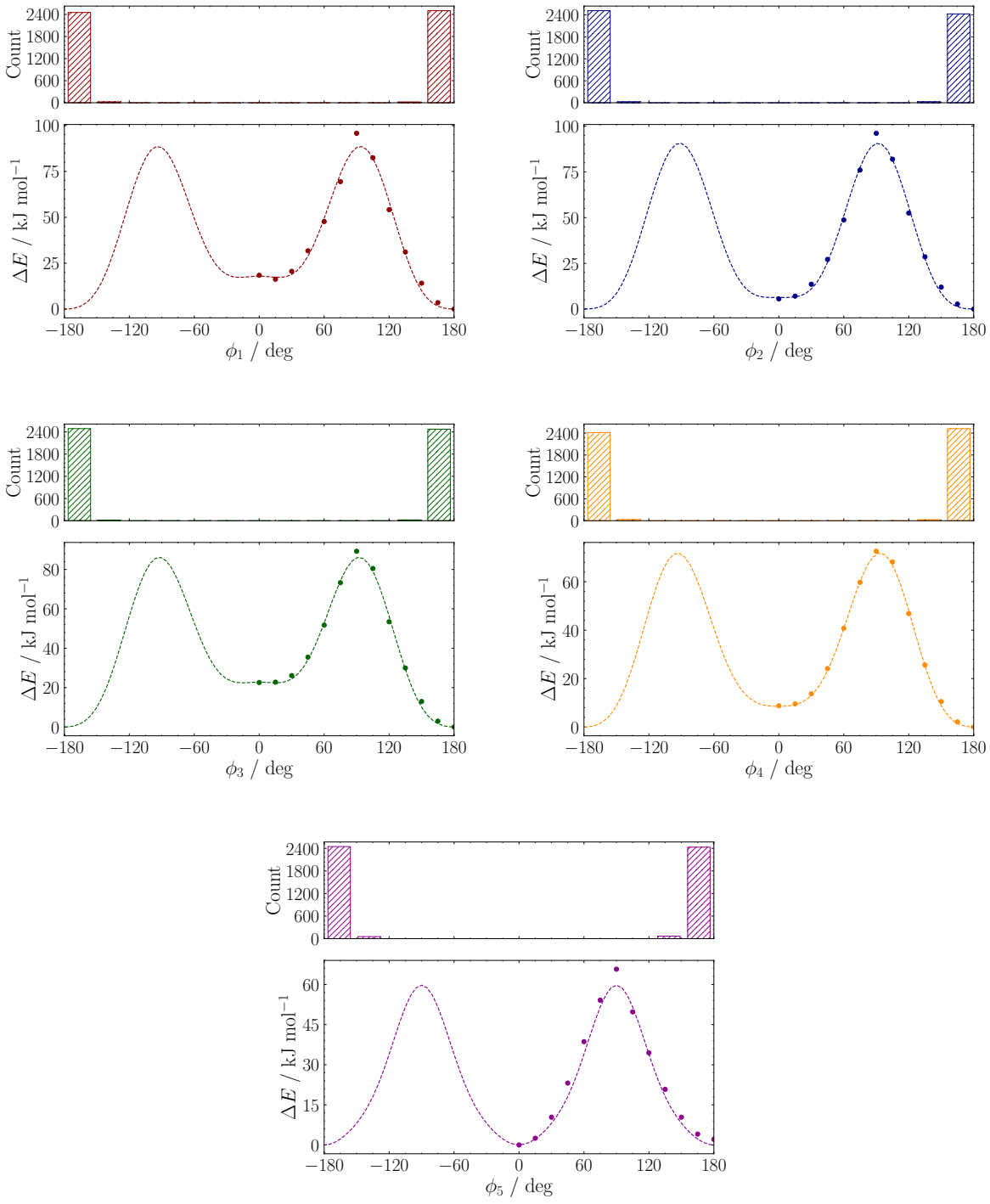


Fig. S40: QM (dots) and MM (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for MMAR, and their distribution along a 500 ps MD run.

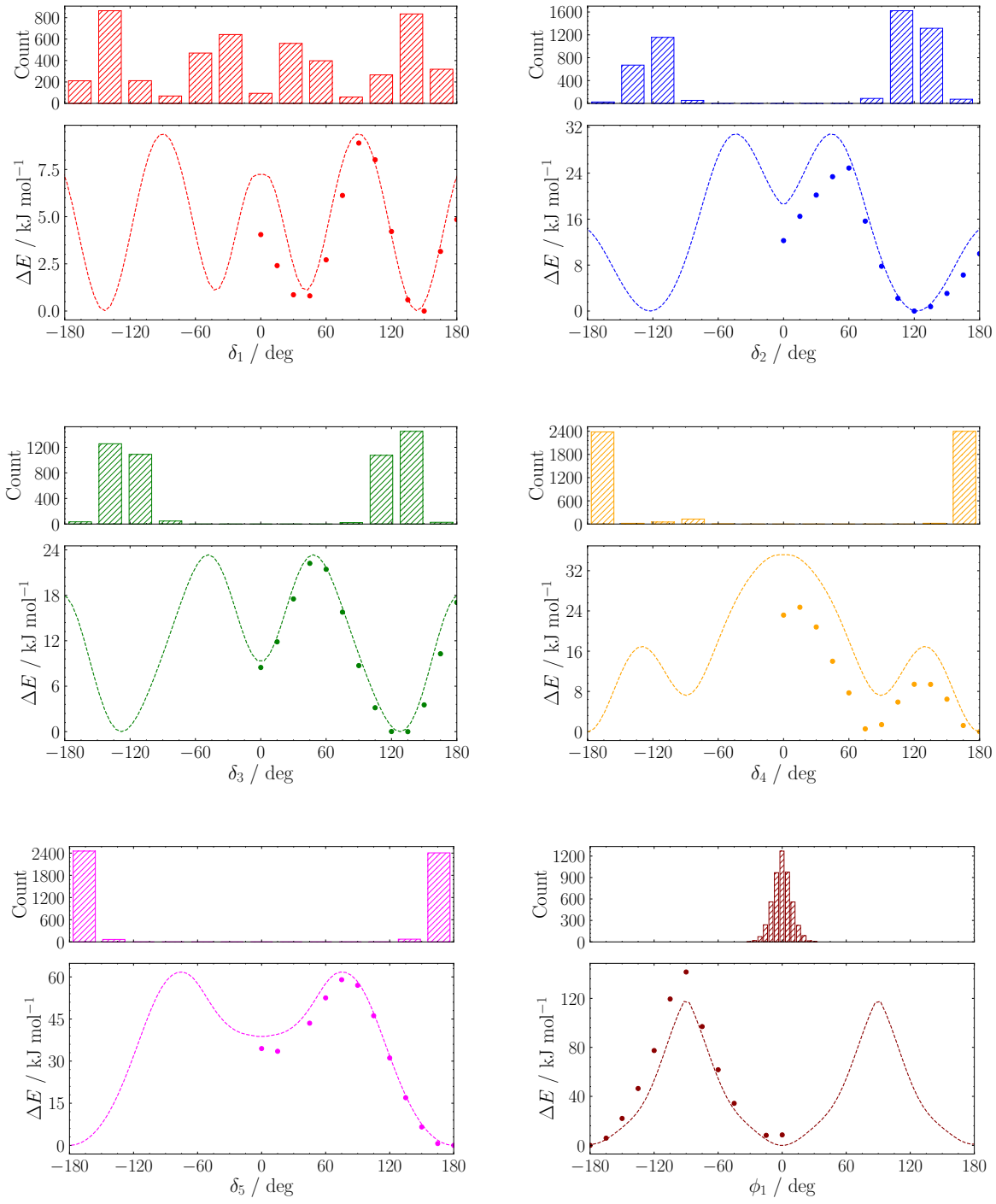


Fig. S41: QM (dots) and MM (dashed lines) potential energy profiles for δ_1–δ_5 and ϕ_1 flexible dihedrals for MAMHPA, and their distribution along a 500 ps MD run.

S5.2 GAFF

S5.2.1 Torsional Profiles

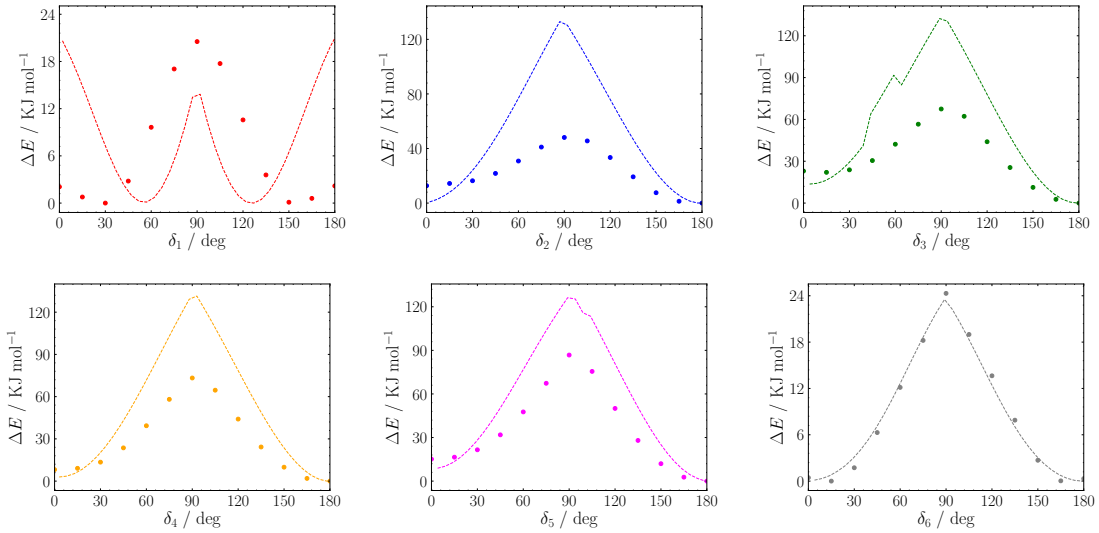


Fig. S42: QM (dots) and GAFF (dashed lines) potential energy profiles for δ_1–δ_6 flexible dihedrals for MMAR.

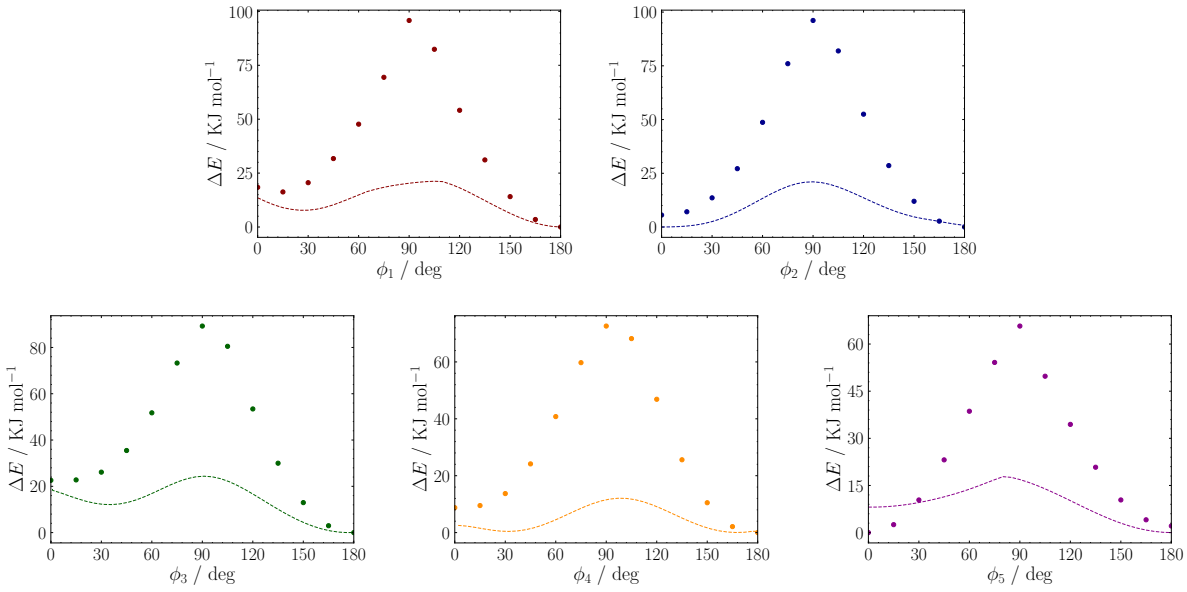


Fig. S43: QM (dots) and GAFF (dashed lines) potential energy profiles for ϕ_1–ϕ_5 flexible dihedrals for MMAR.

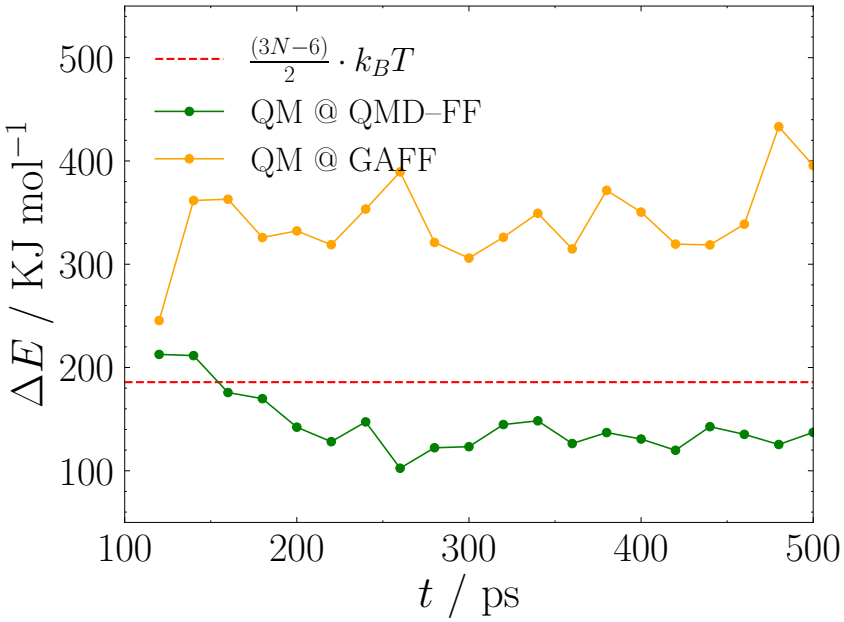


Fig. S44: Comparison of MP2/6–31G(d) internal energies on snapshots extracted from test MD runs at 298 K for MMAR carried out with either our QMD–FFs or GAFF with internal energy expected from the virial theorem. Energies are relative to the MP2/6–31G(d) optimised geometry.

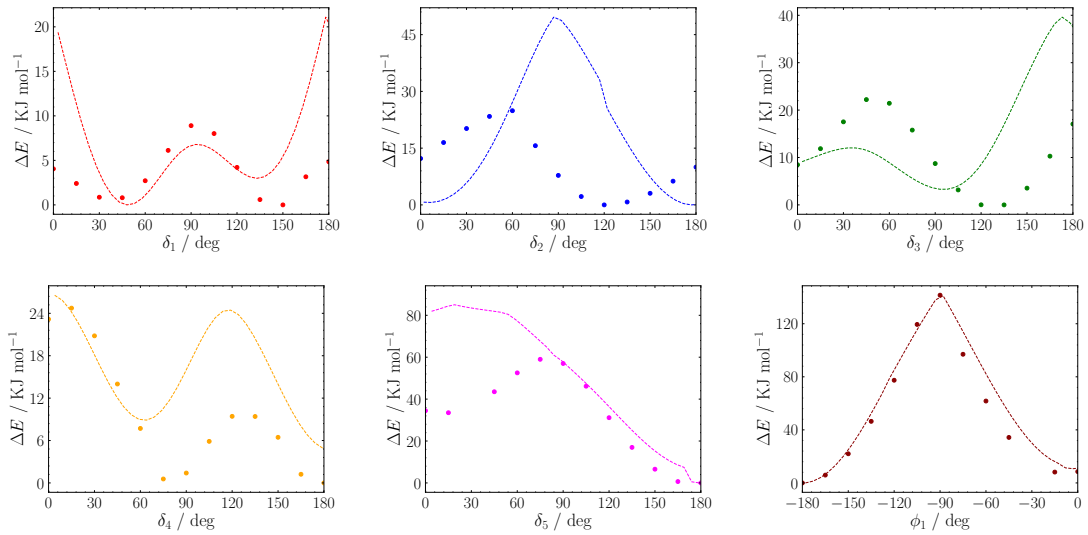


Fig. S45: QM (dots) and GAFF (dashed lines) potential energy profiles for δ_1–δ_5 and ϕ_1 flexible dihedrals for MAMHPA.

References

[1] I. Cacelli, G. Prampolini, *J. Chem. Theory Comput.* **2007**, *3*, 1803–1817.

[2] N. De Mitri, S. Monti, G. Prampolini, V. Barone, *J. Chem. Theory Comput.* **2013**, *9*, 4507–4516.

[3] O. Andreussi, I. G. Prandi, M. Campetella, G. Prampolini, B. Mennucci, *J. Chem. Theory Comput.* **2017**, *13*, 4636–4648.

[4] G. Prampolini, A. Andersen, B. I. Poulter, M. Khalil, N. Govind, E. Biasin, M. Pastore, *J. Chem. Theory Comput.* **2023**, *20*, 1306–1323.

[5] J. Cerezo, J. Gierschner, F. Santoro, G. Prampolini, *ChemPhysChem* **2024**, *25*, e202400307.

[6] S. Giannini, P. Martinez, A. Semmeq, J. Galvez, A. Piras, A. Landi, D. Padula, F. Santoro, J. Vilhena, J. Cerezo, G. Prampolini, *J. Chem. Theory Comput.* **2025**, *ASAP*, `https://doi.org/10.1021/acs.jctc.5c00010`.

[7] G. Prampolini, V. K. Porwal, A. Carof, F. Ingrosso, *J. Mol. Liq.* **2024**, *396*, 123898.

[8] A. Dellai, C. Naim, J. Cerezo, G. Prampolini, F. Castet, *Phys. Chem. Chem. Phys.* **2024**, *26*, 13639–13654.

[9] J. Cerezo, G. Prampolini, I. Cacelli, *Theor. Chem. Acc.* **2018**, *137*, 80.

[10] I. B. Nielsen, L. Lammich, L. H. Andersen, *Phys. Rev. Lett.* **2006**, *96*, 018304.

[11] L. H. Andersen, I. B. Nielsen, M. B. Kristensen, M. O. A. E. Ghazaly, S. Haacke, M. B. Nielsen, M. Å. Petersen, *J. Am. Chem. Soc.* **2005**, *127*, 12347–12350.