RSC Advances

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *RSC Adv.*, 2025, **15**, 7257

Dihedral–torsion model potentials that include angle-damping factors†

Thomas A. Manz *

This groundbreaking study derives and tests several new dihedral torsion model potentials for constructing classical forcefields for atomistic simulations of materials. (1) The new angle-damped dihedral torsion (ADDT) model potential is preferred when neither contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $\neq$ 180°), at least one of the contained equilibrium bond angles is $\geq$ 130° (*i.e.*, (θ^{eq}_{ABC} or θ^{eq}_{BCD}) $\geq$ 130°), and the dihedral torsion potential contains some odd-function contributions (*i.e.*, $U[\phi] \neq U[-\phi]$). (2) The new angle-damped cosine only (ADCO) model potential is preferred when neither contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $\neq$180°), at least one of the contained equilibrium bond angles is $\geq$ 130° (*i.e.*, (θ^{eq}_{ABC} or θ^{eq}_{BCD}) $\geq$ 130°), and the dihedral torsion potential contains no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$). (3) The new constant amplitude dihedral torsion (CADT) model potential is preferred when neither contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $\neq$ 180°), both contained equilibrium bond angles are <130° (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) < 130°), and the dihedral torsion potential contains some odd-function contributions (*i.e.*, $U[\phi] \neq U[-\phi]$). (4) The constant amplitude cosine only (CACO) model potential is preferred when neither contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $\neq$180°), both contained equilibrium bond angles are <130° (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) <130°), and the dihedral torsion potential contains no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$). (5) The new angle-damped linear dihedral (ADLD) model potential is preferred when at least one contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} or θ^{eq}_{BCD}) = 180°). Most importantly, this article derives combined angle-dihedral coordinate branch equivalency conditions and angle-damping factors that ensure the angle-damped torsion model potentials (*e.g.*, ADDT, ADCO, and ADLD) are mathematically consistent and continuously differentiable even as at least one contained bond angle approaches linearity (*i.e.*, as (θ_{ABC} or θ_{BCD}) $\rightarrow$ 180°). This article introduces the torsion offset potential (TOP). I show the TOP gives rise in some materials to the unusual physical phenomenon of slip torsion. For various molecules, extensive quantitative comparisons to high-level quantum chemistry calculations (*e.g.*, CCSD) and experimental vibrational frequencies showed these new dihedral torsion model potentials perform superbly.

Received 23rd December 2024
Accepted 22nd January 2025

DOI: 10.1039/d4ra08960j

rsc.li/rsc-advances

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

1. Introduction

The directed dihedral ϕ_{ABCD} measures the directed angle between the plane containing atoms ABC and the plane containing atoms BCD.[1] The allowed range is:

$$-\pi < \phi_{ABCD} \leq \pi \quad (1)$$

The directed dihedral is defined for all situations except when atoms ABC or atoms BCD reside on a line, which corresponds to $\theta_{ABC} = \pi$, $\theta_{BCD} = \pi$, $\theta_{ABC} = 0$, $\theta_{BCD} = 0$, or at least two of the atoms have the same nuclear position. The latter condition of

Chemical & Materials Engineering, New Mexico State University, Las Cruces, NM 88001, USA. E-mail: tmanz@nmsu.edu

† Electronic supplementary information (ESI) available: A PDF file containing analytic formulas for gradients of the kangals and directed dihedral; analytic formulas for derivatives of the angle-damping factors; analytic formulas for the first derivatives and forces of the CADT, CACO, ADDT, ADCO, and ADLD model potentials; selectivity rule for single-linear dihedrals; supplementary tables and supplementary figure; summary of validation tests of analytic derivative and force formulas; and derivation of torsion scan *R*-squared formulae. A zip archive containing optimized geometries of molecules; spreadsheets containing torsion mode analysis for all molecules studied in this work; computed data for potential energy surface plots for HNCO, acetonitrile, boranecarbonitrile, isocyanoborane molecules; computed data for energy parity and contour plots for acetylene and HCNO molecules; training and validation datasets and forcefield parameterization codes and results for 2-amino-propanal, acetylene, HNCO, and HOOH molecules; computed data for bond-angle scan energy curves for 10 molecules; Matlab codes implementing the analytic derivatives and force formulas for the ADDT, ADCO, CADT, CACO, and ADLD model potentials including ten runs each comparing analytically-computed forces to forces computed numerically using the central finite-difference approximation; and other data workup spreadsheets. See DOI: https://doi.org/10.1039/d4ra08960j

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

two or more atoms having the same nuclear position is physically prevented by the Pauli exclusion principle.[2]

In a material, a proper dihedral A–B–C–D corresponds to the situation in which bonds A–B, B–C, and C–D all exist.[3] An improper dihedral corresponds to the situation in which three atoms are directly bonded to a common atom; for example, three hydrogen atoms directly bonded to a carbon atom in a methyl group.[4] The remainder of this article refers to proper dihedrals. For simplicity, we subsequently refer to a 'proper dihedral' as a 'dihedral' and an 'improper dihedral' as an 'improper-dihedral'.

Let $\vec{R}_{AB}$ be the vector from the position of atom A to the position of atom B:

$$\vec{R}_{AB} = \vec{R}_B - \vec{R}_A \tag{2}$$

The directed dihedral ϕ_{ABCD} is computed as follows:[1,5]

$$\vec{t} = -\left(\vec{R}_{AB} \times \vec{R}_{BC}\right) \tag{3}$$

$$\vec{u} = -\left(\vec{R}_{BC} \times \vec{R}_{CD}\right) \tag{4}$$

$$\phi_{ABCD} = \text{sign} \times \text{acos}\left[\min\left[\max\left[-1, \left(\frac{\vec{t}\cdot\vec{u}}{|\vec{t}||\vec{u}|}\right)\right], 1\right]\right] \tag{5}$$

The sign is specified using the following method:

$$\text{sign} = \begin{cases} 1 & \vec{R}_{BC}\cdot(\vec{t}\times\vec{u}) = 0 \\ \dfrac{\vec{R}_{BC}\cdot(\vec{t}\times\vec{u})}{\text{abs}\left[\vec{R}_{BC}\cdot(\vec{t}\times\vec{u})\right]} & \text{otherwise} \end{cases} \tag{6}$$

To cancel the effect of roundoff error, we used min and max functions to guarantee the argument of acos is between −1 and 1. When ϕ_{ABCD} is 0 or 180°, all four atoms are in the same plane. When $\phi_{ABCD} = 0$, atoms A and D are located on the same side of the BC bond in the ABCD plane. When $\phi_{ABCD} = 180°$, atoms A and D are located on opposite sides of the BC bond in the ABCD plane. When ϕ_{ABCD} is −90° or 90°, the ABC and BCD planes are perpendicular.

Fig. 1 shows an achiral isomer of $(CClFH)_2$ and the *S* enantiomer of the chiral molecule C(OH)ClFH with different values of the HCCH or FCOH dihedral labeled, respectively. For these two molecules, Fig. 1 shows the particular isomers studied in this article.

The torsion potential for dihedral ABCD describes the change in potential energy that results from rotating bonds AB and CD around bond BC. Because this potential energy returns to its initial value when a complete full rotation is performed, the torsion potential is periodic in the dihedral value ϕ_{ABCD}. A whole number of periods must be traversed when the dihedral changes from $\phi_{ABCD} = \pi$ to $\phi_{ABCD} \rightarrow -\pi$.

Please note that $U[\phi_{ABCD}] = \text{constant}$ is not considered to be any kind of torsion potential. Rather, it is considered to be the absence of a torsion potential, because it has no dependence on the dihedral's value.

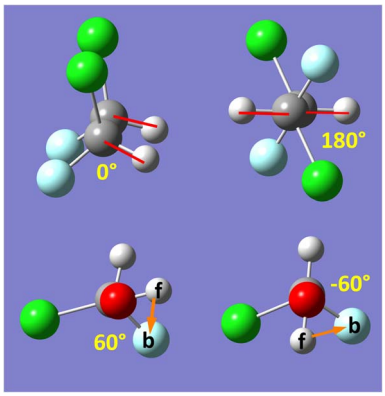

Fig. 1 The achiral molecule $(CClFH)_2$ (upper panels) and the *S* enantiomer of the chiral molecule C(OH)ClFH (lower panels) with different values of the HCCH or FCOH dihedral angle labeled, respectively. For $(CClFH)_2$, the minimum energy geometry corresponds to a HCCH dihedral value of 180°. For C(OH)ClFH, the minimum energy geometry corresponds to a FCOH dihedral value of −64.7°. Atom colors: grey (C), green (Cl), cyan (F), white (H), red (O). The dihedral value is positive (*cf.* negative) if the back substituent is located clockwise (*cf.* counter-clockwise) relative to the front substituent. For clarity, the front substituent is labeled 'f', and the back substituent is labeled 'b' in the lower panels.

For convenience, we categorize torsion potentials into five classes such that each torsion potential is a member of exactly one of these five classes. (A) 'Dihedral-only' torsion potentials depend exclusively on the dihedral value (*e.g.*, ϕ_{ABCD}) with no explicit dependence on the bond lengths or bond angles. (B) 'Angle-damped' torsion potentials depend exclusively on the dihedral value (*e.g.*, ϕ_{ABCD}) and the two contained bond angle values (*i.e.*, θ_{ABC} and θ_{BCD}) with no explicit dependence on the bond lengths. (C) 'Distance-damped' torsion potentials depend exclusively on the dihedral value (*e.g.*, ϕ_{ABCD}) and the three contained bond lengths (*i.e.*, R_{AB}, R_{BC}, and R_{CD}) with no explicit dependence on the bond angles. (D) 'Fully-damped' torsion potentials depend exclusively on the dihedral value (*e.g.*, ϕ_{ABCD}), the two contained bond angle values (*i.e.*, θ_{ABC} and θ_{BCD}), and the three contained bond lengths (*i.e.*, R_{AB}, R_{BC}, and R_{CD}). (E) The final class contains all of the miscellaneous torsion potentials that do not fit into any of the first four classes.

Class A torsion potentials that depend only on the dihedral's value are abundant in the prior literature.[1,3,5–10] However, it is straightforward to prove that every Class A torsion potential

$$U_{\text{torsion}}^{\text{dihedral_only}}[\phi_{ABCD}] = \text{function}[\phi_{ABCD}] \neq \text{constant} \tag{7}$$

is mathematically and physically inconsistent if one of the contained bond angles approaches linearity.[11–13] Proof: suppose that π minus the contained bond angle θ_{ABC} becomes an infinitesimal positive value

$$(\pi - \theta_{ABC}) = \varepsilon > 0 \tag{8}$$

A dihedral scan can be performed by holding the atoms B, C, and D fixed and moving atom A in a circle whose radius is $d_{AB}\sin[\varepsilon]$, where d_{AB} is the distance between atoms A and B. The

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

circumference of this circle is $2\pi d_{AB}\sin[\varepsilon]$. Since $U_{\text{torsion}}^{\text{dihedral_only}}[\phi_{ABCD}] \neq$ constant and $U_{\text{torsion}}^{\text{dihedral_only}}[\phi_{ABCD}]$ satisfies the periodic boundary condition

$$U_{\text{torsion}}^{\text{dihedral_only}}[\phi_{ABCD}] = U_{\text{torsion}}^{\text{dihedral_only}}[\phi_{ABCD} + 2\pi] \quad (9)$$

it follows that we can find some dihedral values $\phi_{ABCD} = \varphi_2$ and $\phi_{ABCD} = \varphi_1$ such that

$$\frac{U_{\text{torsion}}^{\text{dihedral_only}}[\varphi_2] - U_{\text{torsion}}^{\text{dihedral_only}}[\varphi_1]}{(\varphi_2 - \varphi_1)} \gg \Delta \quad (10)$$

for any infinitesimal positive value Δ. For analogous reasons, we can also find some dihedral values $\phi_{ABCD} = \varphi_3$ and $\phi_{ABCD} = \varphi_4$ such that

$$\frac{U_{\text{torsion}}^{\text{dihedral_only}}[\varphi_4] - U_{\text{torsion}}^{\text{dihedral_only}}[\varphi_3]}{(\varphi_4 - \varphi_3)} \ll -\Delta \quad (11)$$

The only way we could not find such values is if $U_{\text{torsion}}^{\text{dihedral_only}}[\phi_{ABCD}] =$ constant, which represents the absence of a torsion potential. The force exerted on atom A during each of these two parts of the dihedral scan is given by

$$\text{Force}_A = -\frac{\text{change_in_potential_energy}}{\text{distance_moved}} \quad (12)$$

$$Force_A^{1\to 2} = -\frac{U_{\text{torsion}}^{\text{dihedral_only}}[\varphi_2] - U_{\text{torsion}}^{\text{dihedral_only}}[\varphi_1]}{(\varphi_2 - \varphi_1)d_{AB}\sin[\varepsilon]} \quad (13)$$

$$\text{Force}_A^{3\to 4} = -\frac{U_{\text{torsion}}^{\text{dihedral_only}}[\varphi_4] - U_{\text{torsion}}^{\text{dihedral_only}}[\varphi_3]}{(\varphi_4 - \varphi_3)d_{AB}\sin[\varepsilon]} \quad (14)$$

Substituting eqn (10) into (13), substituting eqn (11) into (14), and taking the limit $\varepsilon \to 0$ gives

$$\lim_{\varepsilon\to 0}\text{Force}_A^{1\to 2} = -\infty \quad (15)$$

$$\lim_{\varepsilon\to 0}\text{Force}_A^{3\to 4} = \infty \quad (16)$$

As $\varepsilon \to 0$, the force on atom A fluctuates from infinitely positive to infinitely negative over an infinitesimally small distance. This extreme behavior is physically inconsistent. Practically, this means that every Class A torsion potential is inapplicable if one of the contained bond angles (*i.e.*, ∢ABC and/or ∢BCD) is statistically likely to approach linearity either (i) during the course of a molecular dynamics simulation or (ii) in thermally accessible conformations during a Monte Carlo simulation.[11,13,14]

Examples of previously published works on damped torsion potentials include the following. Grimme's quantum-mechanically-derived force field (QMDFF) includes a distance-damped torsion potential (*i.e.*, Class C torsion potential).[15] There have been a few prior studies of angle-damped (*i.e.*, Class B) torsion potentials, but these are not comprehensive or fundamental in nature.[11,12,14,16,17] Tuzun *et al.* describe a Class D torsion potential that includes both angle damping and distance damping.[13]

Herein, I introduce a comprehensive and fundamental theory of angle-damped dihedral torsion model potentials that are mathematically well-defined for all bond angle values. The remainder of this article is organized as follows. Section 2 introduces the combined angle-dihedral coordinate branch equivalency conditions, mathematical constraints on the angle-damping factors, a specific model function for the angle-damping factors, and principles used to derive the angle-damped torsion model potentials. Section 3 introduces the mirror image parameter S_{instance} that allows both mirror images to be classified within the same dihedral type and to be described by the same torsion force constant values. Section 4 derives the angle-damped dihedral torsion (ADDT) model potential as a Class B torsion potential that applies when neither equilibrium bond angle (*i.e.*, θ_{ABC}^{eq} and θ_{BCD}^{eq} for dihedral ABCD) is linear. Key components of the ADDT model potential include the torsion offset potential (TOP) and a set of orthonormal rotatable dihedral torsion modes. This ADDT model potential has continuous derivatives as either or both of the contained bond angles (*i.e.*, θ_{ABC} and θ_{BCD} for dihedral ABCD) approaches linearity. Section 5 introduces the constant-amplitude dihedral torsion (CADT) model potential that is a Class A torsion potential with orthonormal rotatable dihedral torsion modes. The CADT model potential emerges from the ADDT model potential when all of the angle-damping factors are set to a constant value. Section 6 introduces the angle-damped cosine only (ADCO) model potential as a Class B torsion potential that applies when neither equilibrium bond angle (*i.e.*, θ_{ABC}^{eq} and θ_{BCD}^{eq} for dihedral ABCD) is linear and the torsion potential contains no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$). Section 6 also introduces the constant amplitude cosine only (CACO) model potential as a Class A torsion potential that applies when the torsion potential contains no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$) and all of the angle-damping factors are set to a constant value. Section 7 describes rotatable dihedral mode smart selection using torsion scans. Section 8 discusses how the CADT and CACO model potentials compare to previously published Class A torsion potentials. Section 9 introduces the angle-damped linear dihedral (ADLD) model potential as a Class B torsion potential that applies when one or both of the equilibrium bond angles (*i.e.*, θ_{ABC}^{eq} and θ_{BCD}^{eq} for dihedral ABCD) is linear. Section 10 contains an extensive set of computed results for many molecules. These results show my torsion model potentials closely reproduce quantum-mechanically-computed potential energies surfaces for various torsion types. Section 10.8 compares vibrational frequencies computed from flexibility models (which include bond stretches, angle bends, and dihedral torsions) to prior experimentally-measured values for several small molecules. Section 11 concludes.

Note: this article adopts the convention that function arguments are enclosed in square brackets, while parentheses denote multiplication. For example, $y[x + 2]$ means 'y as a function of $(x + 2)$' while $y(x + 2)$ means 'y times $(x + 2)$'.

Throughout this article the R-squared statistical descriptor has the definition

$$R\text{-squared} = R^2 = 1 - \frac{\text{SSE}}{\text{SST}} \quad (17)$$

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Specific definitions for the sum of squared errors (SSE) and sum of squares total (SST) are presented in each context.

2. Mathematical principles governing the angle-dihedral coupling

2.1 Combined angle-dihedral coordinate branch equivalency conditions

The concept of combined angle-dihedral coordinate branch equivalency is illustrated in Fig. 2. As the BCD angle sweeps across the $\theta_{BCD} = \pi$ line (*e.g.*, from hydrogen atom position D1 to D2 in Fig. 2), the value of ϕ_{ABCD} discontinuously jumps from φ to $(\varphi \pm \pi)$. Accordingly, this means the coordinate pair $(\theta_{BCD},\phi_{ABCD}) = ((\pi - \Delta),\varphi)$ (aka 'coordinate branch 1') and the hypothetical coordinate pair $(\tilde{\theta}_{BCD},\tilde{\phi}_{ABCD}) = ((\pi + \Delta),(\varphi \pm \pi))$ (aka 'coordinate branch 2') describe the same ABCD conformation, where $0 < \Delta < \pi$. Consequently, the angle-damped dihedral torsion potential must satisfy the following coordinate branch equivalency relations:

$$\begin{aligned} U_{\text{torsion}}^{\text{angle-damped}}[\theta_{ABC},\theta_{BCD},\phi_{ABCD}] &= \\ U_{\text{torsion}}^{\text{angle-damped}}[\theta_{ABC},(2\pi-\theta_{BCD}),(\phi_{ABCD}\pm\pi)] &= \\ U_{\text{torsion}}^{\text{angle-damped}}[(2\pi-\theta_{ABC}),\theta_{BCD},(\phi_{ABCD}\pm\pi)] & \end{aligned} \quad (18)$$

As the BCD angle sweeps across the $\theta_{BCD} = \pi$ line (*e.g.*, from hydrogen atom position D1 to D2 in Fig. 2), this branch equivalency relationship allows us to choose values

$$(\bar{\theta}_{BCD},\bar{\phi}_{ABCD}) \in \{(\theta_{BCD},\phi_{ABCD}),(\tilde{\theta}_{BCD},\tilde{\phi}_{ABCD})\} \quad (19)$$

that change continuously. Analogously, we can find a set of coordinates

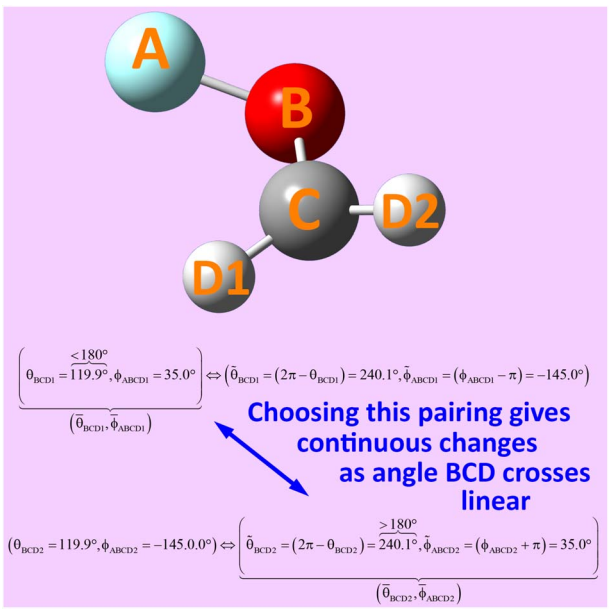

Fig. 2 Illustration of angle-dihedral coordinate branch equivalency. Coordinate branches can be selected to yield continuous changes as the adjusted bond angle $\bar{\theta}_{BCD}$ crosses linear (*i.e.*, crosses 180°).

$$(\bar{\theta}_{ABC},\bar{\phi}_{ABCD}) \in \{(\theta_{ABC},\phi_{ABCD}),(\tilde{\theta}_{ABC},\tilde{\phi}_{ABCD})\} \quad (20)$$

that change continuously as the angle ABC angle sweeps across the $\theta_{ABC} = \pi$ line.

By constructing $U_{\text{torsion}}^{\text{angle-damped}}$ to be an infinitely differentiable function of the continuous coordinates $(\bar{\theta}_{ABC},\bar{\theta}_{BCD},\bar{\phi}_{ABCD})$, this ensures that the derivatives

$$\frac{\partial^{p+q+r}}{\partial\bar{\theta}_{ABC}{}^{p}\partial\bar{\theta}_{BCD}{}^{q}\partial\bar{\phi}_{ABCD}{}^{r}} U_{\text{torsion}}^{\text{angle-damped}}\left[\bar{\theta}_{ABC},\bar{\theta}_{BCD},\bar{\phi}_{ABCD}\right] \quad (21)$$

are continuous, where p, q, and r are any combination of non-negative integers. In summary, this provides a notion of continuous derivatives of all orders for $U_{\text{torsion}}^{\text{angle-damped}}$.

2.2 Mathematical constraints on the angle-damping factors

First, the angle-damping factor must go to zero as the corresponding bond angle approaches linearity. For simplicity, we seek a proxy variable that decreases monotonically from one when $\theta_{ABC} \to 0$ to zero when $\theta_{ABC} = \pi$. The 'kangal'

$$Ж_{ABC} = \cos[\theta_{ABC}/2] \quad (22)$$

provides an ideal proxy variable. (Kangal is a specific breed of guardian dog. The term 'kangal' as used here is a play on words that means 'keeper (or guardian) of the angle'.)

Consider a spherical coordinate system centered on atom C in Fig. 2, with radius $r = R_{CD}$, polar angle $\theta = \theta_{BCD}$, and azimuthal angle $\phi = \phi_{ABCD}$. For a point within this spherical coordinate system, the corresponding Cartesian coordinates are

$$x = r\sin[\theta]\cos[\phi] \quad (23)$$

$$y = r\sin[\theta]\sin[\phi] \quad (24)$$

$$z = r\cos[\theta] \quad (25)$$

This satisfies the coordinate branch equivalency condition that (r,θ,ϕ) has the same (x, y, z) coordinates as $(r,(2\pi - \theta),\phi \pm \pi)$.

Using the standard trigonometric formulas

$$\cos[\alpha + \beta] = \cos[\alpha]\cos[\beta] - \sin[\alpha]\sin[\beta] \quad (26)$$

$$\sin[\alpha + \beta] = \sin[\alpha]\cos[\beta] + \cos[\alpha]\sin[\beta] \quad (27)$$

angle-damped higher-order torsion multiplicities can be constructed as

$$x^2 - y^2 = r^2\sin^2[\theta]\cos[2\phi] \quad (28)$$

$$2xy = r^2\sin^2[\theta]\sin[2\phi] \quad (29)$$

$$x^3 - 3xy^2 = r^3\sin^3[\theta]\cos[3\phi] \quad (30)$$

$$3x^2y - y^3 = r^3\sin^3[\theta]\sin[3\phi] \quad (31)$$

and so forth. For $n = 1, 2, 3, 4, \ldots$ the left-hand side is a nth-order polynomial in the x and y coordinates, while the right-hand side is $r^n\sin^n[\theta]\cos[n\phi]$ or $r^n\sin^n[\theta]\sin[n\phi]$.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Combining the trigonometric identities

$$\sin[\theta_{BCD}] = 2\cos[\theta_{BCD}/2]\sin[\theta_{BCD}/2] \quad (32)$$

with the definition of the kangal gives

$$\sin[\theta_{BCD}] = 2\text{Ж}_{BCD}\sqrt{1-(\text{Ж}_{BCD})^2} \quad (33)$$

We can re-write eqn (28)–(31) in terms of the kangal as

$$\frac{x^2-y^2}{r^2} = 4\left(1-(\text{Ж}_{BCD})^2\right)\text{Ж}_{BCD}{}^2\cos[2\phi] \quad (34)$$

$$\frac{2xy}{r^2} = 4\left(1-(\text{Ж}_{BCD})^2\right)\text{Ж}_{BCD}{}^2\sin[2\phi] \quad (35)$$

$$\frac{x^3-3xy^2}{r^3} = 8\left(\sqrt{1-(\text{Ж}_{BCD})^2}\right)^3\text{Ж}_{BCD}{}^3\cos[3\phi] \quad (36)$$

$$\frac{3x^2y-y^3}{r^3} = 8\left(\sqrt{1-(\text{Ж}_{BCD})^2}\right)^3\text{Ж}_{BCD}{}^3\sin[3\phi] \quad (37)$$

Next, we note that

$$\cos[n\phi_{ABCD}] = (-1)^n\cos[n(\phi_{ABCD}\pm\pi)] \quad (38)$$

$$\sin[n\phi_{ABCD}] = (-1)^n\sin[n(\phi_{ABCD}\pm\pi)] \quad (39)$$

$$(\text{Ж}_{BCD})^p = (\cos[\theta_{BCD}/2])^p = (-1)^p(\cos[(2\pi-\theta_{BCD})/2])^p \quad (40)$$

Therefore, to satisfy the coordinate branch equivalency condition that the torsion potential has the same numeric value for $(\theta_{ABC},\theta_{BCD},\phi_{ABCD})$ as for $(\theta_{ABC},(2\pi-\theta_{BCD}),\phi_{ABCD}\pm\pi)$, $\cos[2j\phi_{ABCD}]$ and $\sin[2j\phi_{ABCD}]$ must be multiplied by even-only powers of Ж_{BCD}, where j is a whole number. Similarly, $\cos[(2j+1)\phi_{ABCD}]$ and $\sin[(2j+1)\phi_{ABCD}]$ must be multiplied by odd-only powers of Ж_{BCD}, where j is a whole number.

As the BCD bond angle approaches linearity (*i.e.*, $\text{Ж}_{BCD}\to 0$), the angle-damping product on the right-hand side of eqn (34)–(37) have derivatives $(\partial^i/\partial\text{Ж}_{BCD}{}^i)$ equal to zero for $i = 1, 2, 3 \ldots (n-1)$, where $\cos[n\phi]$ or $\sin[n\phi]$ is the dihedral-dependent factor. To preserve this behavior, the ideal angle-damped $\cos[n\phi]$ term should have the form:

$$\left(\sum_{j=0}^{\infty}a_{n,j}(\text{Ж}_{ABC})^{n+2j}\right)\left(\sum_{j=0}^{\infty}a_{n,j}(\text{Ж}_{BCD})^{n+2j}\right)\cos[n\phi] \quad (41)$$

where $a_{n,0} > 0$. Note that the form of angle-damping factor for the ABC angle is analogous to that for the BCD angle. To preserve invariance in the choice of spherical coordinate system starting azimuthal angle, the angle-damping factor for $\sin[n\phi]$ is the same as the angle-damping factor for $\cos[n\phi]$.

2.3 A specific model for the elementary angle-damping functions

The zeroth-order angle-damping factor is a constant:

$$f_0^{ABC} = 1 \quad (42)$$

For $f_{n\geq 1}^{ABC}$, there are four regimes of bond angles:

• Linear bond angle: when $\theta_{ABC} = 180°$, then $\text{Ж}_{ABC} = 0$ and $f_{n\geq 1}^{ABC} = 0$.

• Wide-angle regime: when $130° \leq \theta_{ABC} < 180°$, it follows from eqn (41) that the leading-order term in $f_{n\geq 1}^{ABC}$ is proportional to $(\text{Ж}_{ABC})^n$ and forms the dominant contribution.

• Acute angle regime: this regime corresponds to $0° < \theta_{ABC} < 90°$ and $1 > \text{Ж}_{ABC} > \sqrt{\frac{1}{2}}$. Since we do not have any *a priori* information about how $f_{n\geq 1}^{ABC}$ should vary as a function of θ_{ABC} in this regime, a reasonable approximation is to mitigate this dependence by setting $f_{n\geq 1}^{ABC} \approx 1$ when $0 < \theta_{ABC} < \pi/2$. We also require that each $f_{n\geq 1}^{ABC}$ must be a monotonically increasing function of Ж_{ABC} and satisfy the limit

$$\lim_{\theta_{ABC}\to 0} f_n^{ABC} = 1 \quad (43)$$

• Intermediate-angle regime: this regime corresponds to $90° \leq \theta_{ABC} < 130°$ and defines the transition regime from perpendicular bond angle (*i.e.*, 90°) to wide bond angle (*i.e.*, 130°). Near the beginning of this transition regime when the bond angle is perpendicular or only slightly obtuse, we expect the different torsion modes to damp similarly as a constrained bond angle changes. In this case, the normalized shape of the angle-constrained torsion scan curve does not change substantially with small changes in the constrained θ_{ABC} value; however, the torsion barrier can change with constrained bond-angle changes. This can be accomplished by defining f_n^{ABC} *via* asymptotic matching such that $f_{n\geq 1}^{ABC} \approx f_1^{ABC}$ when $\theta_{ABC} \lesssim 90°$.

The following unique and parsimonious model achieves these goals. We construct polynomial functions $P_n[\text{Ж}_{ABC}]$ for $n = 1$, 2, 3, and 4 that are asymptotically matched to each other. Each $P_n[\text{Ж}_{ABC}]$ will be a monotonically increasing function of Ж_{ABC} over the range $0 \leq \text{Ж}_{ABC} \leq 1$ and satisfy the boundary conditions

$$P_n[\text{Ж}_{ABC} = 1] = 1 \quad (44)$$

$$P_{n\geq 1}[0] = 0 \quad (45)$$

$$P_{n\geq 1}[\varepsilon] \propto \varepsilon^n \quad (46)$$

where ε is positive infinitesimal.

First, we define two-term polynomials $P_1[\text{Ж}_{ABC}]$ and $P_2[\text{Ж}_{ABC}]$ that are asymptotically matched to each other in the $\text{Ж}_{ABC}\to 1$ limits. Define

$$P_1[\text{Ж}_{ABC}] = \gamma\text{Ж}_{ABC} + (1-\gamma)(\text{Ж}_{ABC})^3 \quad (47)$$

In eqn (47), the two coefficients (*i.e.*, γ and $(1-\gamma)$) sum to one to ensure eqn (44) is satisfied. To ensure that $P_1[\text{Ж}_{ABC}]$ is an odd function, there is no constant (power zero) and power two terms in eqn (47). Define

$$P_2[\text{Ж}_{ABC}] = \varpi(\text{Ж}_{ABC})^2 + (1-\varpi)(\text{Ж}_{ABC})^4 \quad (48)$$

In eqn (48), the two coefficients (*i.e.*, ϖ and $(1-\varpi)$) sum to one to ensure eqn (44) is satisfied. To ensures that $P_2[0] = 0$, there is

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

no constant (power zero) term in eqn (48). To ensure that $P_2[\text{Ж}_{ABC}]$ is an even function, there is no power one or power three terms in eqn (48). Substituting

$$\text{Ж}_{ABC} = 1 - \varsigma \tag{49}$$

into eqn (47) and (48) gives

$$P_1[1-\varsigma] = \gamma(1-\varsigma) + (1-\gamma)(1-\varsigma)^3 = 1 + \varsigma(2\gamma - 3) + \varsigma^2(3 - 3\gamma) + \ldots \tag{50}$$

$$P_2[1-\varsigma] = \varpi(1-\varsigma)^2 + (1-\varpi)(1-\varsigma)^4 = 1 + \varsigma(2\varpi - 4) + \varsigma^2(6 - 5\varpi) + \ldots \tag{51}$$

These two can be asymptotically matched by equating coefficients for their first- and second-order terms:

$$2\gamma - 3 = 2\varpi - 4 \tag{52}$$

$$3 - 3\gamma = 6 - 5\varpi \tag{53}$$

Eqn (52) and (53) comprise a linear equation system having the unique solution

$$\gamma = (1 - \varpi) = \tfrac{1}{4} \tag{54}$$

Substituting eqn (54) into (50) gives

$$P_1[1-\varsigma] = 1 - 2.5\varsigma + 2.25\varsigma^2 + \ldots \tag{55}$$

$P_3[\text{Ж}_{ABC}]$ and $P_4[\text{Ж}_{ABC}]$ can be asymptotically matched to $P_1[\text{Ж}_{ABC}]$ if we choose

$$P_3[\text{Ж}_{ABC}] = a^{(3)}(\text{Ж}_{ABC})^3 + b^{(3)}(\text{Ж}_{ABC})^5 + c^{(3)}(\text{Ж}_{ABC})^7 \tag{56}$$

$$P_4[\text{Ж}_{ABC}] = a^{(4)}(\text{Ж}_{ABC})^4 + b^{(4)}(\text{Ж}_{ABC})^6 + c^{(4)}(\text{Ж}_{ABC})^8 \tag{57}$$

To ensure that $P_3[\text{Ж}_{ABC}]$ is an odd function, eqn (56) includes only odd powers of Ж_{ABC}. To ensure that $P_4[\text{Ж}_{ABC}]$ is an even function, eqn (57) includes only even powers of Ж_{ABC}. Substituting eqn (49) into (56) and (57) gives

$$P_3[1-\varsigma] = (a^{(3)} + b^{(3)} + c^{(3)}) - (3a^{(3)} + 5b^{(3)} + 7c^{(3)})\varsigma + (3a^{(3)} + 10b^{(3)} + 21c^{(3)})\varsigma^2 + \ldots \tag{58}$$

$$P_4[1-\varsigma] = (a^{(4)} + b^{(4)} + c^{(4)}) + (4a^{(4)} + 6b^{(4)} + 8c^{(4)})\varsigma + (6a^{(4)} + 15b^{(4)} + 28c^{(4)})\varsigma^2 + \ldots \tag{59}$$

Asymptotically matching these expansions to eqn (55) gives the following eqn systems:

$$\begin{pmatrix} a^{(3)} + b^{(3)} + c^{(3)} = 1 \\ 3a^{(3)} + 5b^{(3)} + 7c^{(3)} = 2.5 \\ 3a^{(3)} + 10b^{(3)} + 21c^{(3)} = 2.25 \end{pmatrix} \tag{60}$$

$$\begin{pmatrix} a^{(4)} + b^{(4)} + c^{(4)} = 1 \\ 4a^{(4)} + 6b^{(4)} + 8c^{(4)} = 2.5 \\ 6a^{(4)} + 15b^{(4)} + 28c^{(4)} = 2.25 \end{pmatrix} \tag{61}$$

These have the unique solutions

$$(a^{(3)}, b^{(3)}, c^{(3)}) = ((3/2), (-3/4), (1/4)) \tag{62}$$

$$(a^{(4)}, b^{(4)}, c^{(4)}) = ((5/2), (-9/4), (3/4)) \tag{63}$$

In summary, these polynomials have the forms:

$$P_1[\text{Ж}] = \left(\text{Ж} + 3(\text{Ж})^3\right)/4 \tag{64}$$

$$P_2[\text{Ж}] = \left(3(\text{Ж})^2 + (\text{Ж})^4\right)/4 \tag{65}$$

$$P_3[\text{Ж}] = \left(6(\text{Ж})^3 - 3(\text{Ж})^5 + (\text{Ж})^7\right)/4 \tag{66}$$

$$P_4[\text{Ж}] = \left(10(\text{Ж})^4 - 9(\text{Ж})^6 + 3(\text{Ж})^8\right)/4 \tag{67}$$

These are plotted in Fig. 3.

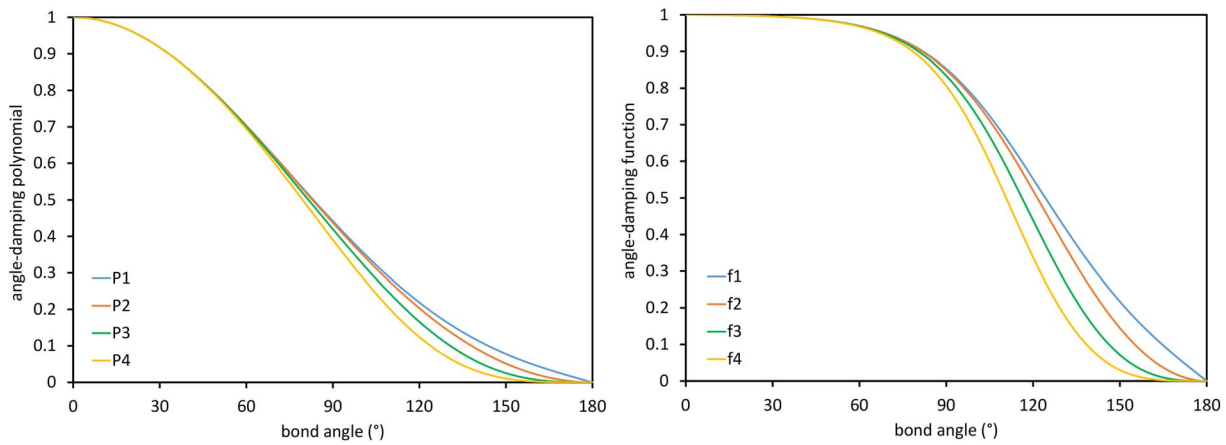

Fig. 3 Values of the polynomials P_1, P_2, P_3, and P_4 (left panel) and elementary angle-damping functions f_1, f_2, f_3, and f_4 (right panel) plotted as a function of the bond angle. As the bond angle value gets smaller, all four curves approach the same curve; that is, they are asymptotically matched.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

This asymptotic matching will ensure that $f_{n\geq 1}^{\mathrm{ABC}} \approx f_1^{\mathrm{ABC}}$ in the acute angle regime if we choose

$$f_n^{\mathrm{ABC}} = Q[P_n[\text{Ж}_{\mathrm{ABC}}]] \tag{68}$$

where the function Q is yet to be determined. To ensure that $Q[P_{\mathrm{odd}_n}[\text{Ж}]]$ is an odd function of Ж, it manifestly follows that Q must be an odd function. To achieve the boundary conditions

$$P_{n\geq 1}[\text{Ж}_{\mathrm{ABC}} = 0] = f_{n\geq 1}|_{\text{Ж}_{\mathrm{ABC}}=0} = 0 \tag{69}$$

$$P_{n\geq 1}[\text{Ж}_{\mathrm{ABC}} = 1] = f_{n\geq 1}|_{\text{Ж}_{\mathrm{ABC}}=1} = 1 \tag{70}$$

it manifestly follows that

$$Q[0] = 0 \tag{71}$$

$$Q[1] = 1 \tag{72}$$

Moreover, the function Q should monotonically increase from $Q[0] = 0$ to $Q[1] = 1$. Comparing eqn (41)–(46) shows that the leading-order term in the Maclaurin series expansion of $Q[x]$ should be first-order in x. Arguably, the simplest function that meets all of these constraints is

$$Q[x] = \frac{\tanh[Kx]}{\tanh[K]} \tag{73}$$

The value of K can be assigned by enforcing

$$\tanh\left[KP_1\left[\underbrace{\cos[(\pi/2)/2]}_{\sqrt{\frac{1}{2}}}\right]\right] \approx 1 \tag{74}$$

via asymptotic matching to reproduce the near-perpendicular-angle behavior described above. We asymptotically match the first three non-zero terms in the Maclaurin series expansion

$$\tanh[x] \approx x - \frac{x^3}{3} + \frac{2x^5}{15} = 1 \tag{75}$$

$x_{\mathrm{root}} = 1.244460035526845\ldots$ is the only real-valued root of eqn (75). The other four roots are complex-valued. Evaluating

$$P_1\left[\sqrt{\frac{1}{2}}\right] = \frac{1}{4}\sqrt{\frac{1}{2}} + \frac{3}{4}\left(\sqrt{\frac{1}{2}}\right)^3 = \sqrt{\frac{1}{2}}(5/8) \tag{76}$$

we obtain

$$K = \frac{x_{\mathrm{root}}}{\sqrt{\frac{1}{2}}(5/8)} = 2.815891616117388\ldots \tag{77}$$

Why asymptotically match the first three non-zero terms (see eqn (75)) instead of one, two, four, or five non-zero terms? Computational tests showed that asymptotically matching an even number of non-zero terms (*i.e.*, 2 or 4 terms) led to the real-valued $x_{\mathrm{root}} < 0$ which is not desirable. Therefore, an odd number of non-zero terms must be included in the asymptotic matching. Matching only one non-zero term (which would give $x_{\mathrm{root}} = 1$) approximates the tanh[x] function as linear which does not account for any curvature effects of the tanh[x] function. Thus, it is desirable to include at least three non-zero terms in the asymptotic matching. Arguable, we should only include the leading-order curvature terms, because we do not want to overfit by including too many terms in the asymptotic matching. Note that the extreme case of overfitting resulting from including an infinitely large odd total number of terms in the asymptotic matching yields $K \to \infty$, which is not a workable solution. Including precisely three non-zero terms in the asymptotic matching has the desirable effect of including some curvature in the tanh[x] expansion without any overfitting.

Putting this altogether, we finally have

$$f_n^{\mathrm{ABC}} = \frac{\tanh[KP_n[\text{Ж}_{\mathrm{ABC}}]]}{\tanh_K} \tag{78}$$

$$f_{n_\mathrm{eq}}^{\mathrm{ABC}} = \frac{\tanh[KP_n[\text{Ж}_{\mathrm{ABC}}^{\mathrm{eq}}]]}{\tanh_K} \tag{79}$$

where

$$\tanh_K = \tanh[K] = 0.992861208914406\ldots \tag{80}$$

with the value of K given in eqn (77). The expressions for f_n^{BCD} and $f_{n_\mathrm{eq}}^{\mathrm{BCD}}$ are obtained by replacing Ж_{ABC} and $\text{Ж}_{\mathrm{ABC}}^{\mathrm{eq}}$ with Ж_{BCD} and $\text{Ж}_{\mathrm{BCD}}^{\mathrm{eq}}$, respectively, in the above equations.

2.4 Principles used to derive the angle-damped torsion model potentials

The ADDT, ADCO, and ADLD model potentials describe dihedral torsion with angle damping. The ADDT model potential applies when neither included equilibrium bond angle is linear (*i.e.*, $\theta_{\mathrm{ABC}}^{\mathrm{eq}} \neq \pi$ and $\theta_{\mathrm{BCD}}^{\mathrm{eq}} \neq \pi$). The ADCO model potential applies when the dihedral torsion potential is an even function of the dihedral value

$$U_{\mathrm{ABCD}}^{\mathrm{ADCO}}[\phi] = U_{\mathrm{ABCD}}^{\mathrm{ADCO}}[-\phi] \tag{81}$$

and neither included equilibrium bond angle is linear (*i.e.*, $\theta_{\mathrm{ABC}}^{\mathrm{eq}} \neq \pi$ and $\theta_{\mathrm{BCD}}^{\mathrm{eq}} \neq \pi$). The ADLD model potential applies when one or both included equilibrium bond angles is linear (*i.e.*, $\theta_{\mathrm{ABC}}^{\mathrm{eq}} = \pi$ or $\theta_{\mathrm{BCD}}^{\mathrm{eq}} = \pi$); in this case, $\phi_{\mathrm{ABCD}}^{\mathrm{eq}}$ cannot be defined.

Before deriving the explicit forms of the ADDT, ADCO, and ADLD model potentials, we must first infer the underlying principles that should govern their forms. I propose the following principles:

(a) To be mathematically and physically self-consistent, every angle-damped dihedral torsion potential should satisfy the combined angle-dihedral coordinate branch equivalency conditions. (See Section 2.1 above.)

(b) The dependence on contained bond angles (*e.g.*, θ_{ABC} and θ_{BCD} for dihedral ABCD) should be formulated using algebraic combinations of elementary angle-damping functions, where each elementary angle-damping function decreases monotonically with increasing bond angle from a value of 1 (for bond angle = 0) to 0 (for bond angle = π). (See Sections 2.2 and 2.3 above.)

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

(c) For rigid values of the contained bond angles (*i.e.*, when the ABC and BCD bond angles are held constant), the dependence on dihedral value ϕ_{ABCD} should comprise a linear combination of torsion modes that is mathematically equivalent to a Fourier series expansion in ϕ_{ABCD}.

(d) Due to the trigonometric identities in eqn (26) and (27), a Fourier series in ϕ_{ABCD} can be constructed equivalently using basis functions of the form $\{\cos[n\phi_{ABCD} - \psi_n^{ABCD}]\}$ or $\{\cos[n\phi_{ABCD}],\sin[n\phi_{ABCD}]\}$. For ADDT but not for ADLD, the Fourier series could also be equivalently constructed using basis functions of the form $\{\cos[n(\phi_{ABCD} - \phi_{ABCD}^{eq})],\sin[n(\phi_{ABCD} - \phi_{ABCD}^{eq})]\}$. Expanding

$$\begin{aligned}&\text{function}[\theta_{ABC},\theta_{BCD}]D_n\cos\left[n\phi_{ABCD} - \psi_n^{ABCD}\right]\\&\quad= \text{function}[\theta_{ABC},\theta_{BCD}](A_n\cos[n\phi_{ABCD}] + B_n\sin[n\phi_{ABCD}])\\&\quad= \text{function}[\theta_{ABC},\theta_{BCD}](\alpha_n\cos[n(\phi_{ABCD} - \phi_{ABCD}^{eq})]\\&\qquad+\beta_n\sin[n(\phi_{ABCD} - \phi_{ABCD}^{eq})])\end{aligned} \tag{82}$$

reveals that the combined angle-damping factor for the sine mode of multiplicity n must be proportional to the combined angle-damping factor for the cosine mode of multiplicity n. Thus, aside from a constant coefficient, sine and cosine modes sharing the same multiplicity n must have identical combined angle-damping factors.

(e) Since ϕ_{ABCD} is undefined when either $\theta_{ABC} = \pi$ or $\theta_{BCD} = \pi$, the combined angle-damping factor multiplying any non-constant function of ϕ_{ABCD} must go to zero whenever either bond angle is linear (*i.e.*, whenever $\theta_{ABC} = \pi$ or $\theta_{BCD} = \pi$). This effectively zeros out that term whenever either bond angle is linear.

(f) Since the square of a real-valued number is non-negative, and since the potential energy of a displaced geometry should not be lower than the ground-state geometry's potential energy, it follows that a useful potential energy model can be constructed using the ansatz

$$\begin{aligned}&U[\text{displaced_geom}] - U[\text{opt_geom}]\\&\qquad= \sum_i\left((\text{form_1}_i)^2 - \text{form_2}_i\right) \geq 0\end{aligned} \tag{83}$$

or similar ansatzes that use squared quantities to ensure the relative potential energy is non-negative. An individual term in the summation is non-negative when

$$\text{form_2}_i \leq (\text{form_1}_i)^2 \tag{84}$$

which is not a strict requirement, because only the combined sum (over i values) in eqn (83) must be non-negative. To ensure the right-hand side of eqn (83) is zero at the optimized geometry, each individual term is defined to be zero at the optimized geometry:

$$\text{form_1}_i[\text{opt_geom}] = 0 \tag{85}$$

$$\text{form_2}_i[\text{opt_geom}] = 0 \tag{86}$$

I call this extremely useful and powerful approach 'completing the squares'.

(g) Within the 'completing the squares' approach, it is often preferable to construct the individual form_1$_i$ and form_2$_i$ terms using the 'most modest and least biased' approach. This means that among different options for constructing the model potential, we preferably choose the option that increases the potential energy by the smaller amount and with less correlation bias. (A specific example of this is discussed in Section 4.1 below.)

(h) For the ADDT and ADCO model potentials, it is convenient to express each combined angle-damping factor as a ratio that equals one when both contained bond angles equal their equilibrium values. Practically, this means each combined angle-damping factor in the ADDT and ADCO model potentials is expressed as some algebraic combination of the ratios $(f_j^{ABC}/f_{j_eq}^{ABC})$ and $(f_j^{BCD}/f_{j_eq}^{BCD})$. These ratios should not be used in the ADLD model potential, because $f_{j\geq1_eq}^{ABC} = 0$ and/or $f_{j\geq1_eq}^{BCD} = 0$ when at least one of the contained equilibrium bond angles is linear.

Without loss of generality, an angle-damped dihedral torsion potential can be expanded as a multivariate Fourier series of the form

$$\begin{aligned}U_{ABCD}^{Class_B}[\theta_{ABC},\theta_{BCD},\phi_{ABCD}] &= A_0W_0[\theta_{ABC},\theta_{BCD}]\\&\quad+\sum_{n=1}^{n_{max}}W_n[\theta_{ABC},\theta_{BCD}](A_n\cos[n\phi_{ABCD}]\\&\quad+B_n\sin[n\phi_{ABCD}])\end{aligned} \tag{87}$$

The form shown in eqn (87) spans the entire function space of continuously differentiable Class B torsion model potentials. Note that following point (d) above, the angle-damping function $W_n[\theta_{ABC},\theta_{BCD}]$ appearing before the $\cos[n\phi_{ABCD}]$ term must be the same as the one appearing before the $\sin[n\phi_{ABCD}]$ term. The combined angle-dihedral coordinate branch equivalency condition applied to $W_n[\theta_{ABC},\theta_{BCD}]\cos[n\phi_{ABCD}]$ and the analogous sine term requires that

$$\begin{aligned}W_n[(2\pi - \theta_{ABC}),\theta_{BCD}] &= W_n[\theta_{ABC},(2\pi - \theta_{BCD})]\\&= (-1)^nW_n[\theta_{ABC},\theta_{BCD}]\end{aligned} \tag{88}$$

3. The mirror image descriptor $S_{instance}$

A reflection operation transforms a chemical geometry into its mirror image. For a molecule whose atom-in-material nuclear positions are expressed in Cartesian coordinates

$$\left\{\vec{R}_A\right\} = \left\{\left(\vec{X}_A,\vec{Y}_A,\vec{Z}_A\right)\right\} \tag{89}$$

reflection can be achieved by changing the sign of the X-coordinate of each and every atom to give

$$\left\{\vec{R}_A^{\,reflected}\right\} = \left\{\left(-\vec{X}_A,\vec{Y}_A,\vec{Z}_A\right)\right\} \tag{90}$$

Eqn (90) corresponds to reflection about the $X = 0$ plane. Reflection can be performed about any plane in 3D Cartesian coordinate space.

Values of the dihedral ABCD before and after reflection are related by

$$\sin\left[\phi_{ABCD}^{before_reflection}\right] = -\sin\left[\phi_{ABCD}^{after_reflection}\right] \tag{91}$$

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

If the dihedral value is zero, then it is mapped onto itself upon refection. If the dihedral value is π (*i.e.*, 180°), then it is mapped onto itself upon reflection. Every other dihedral value is mapped upon reflection onto the dihedral value of the same absolute value but opposite sign. This follows directly from the definition of the directed dihedral given in the Introduction above. In the absence of externally applied fields, mirror image geometries have the same energy.

For this reason, it is convenient to define dihedral types such that all dihedral instances belonging to the same dihedral type: (1) are composed of the same atom types, bond types, and angle types, and (2) have the same absolute value of equilibrium dihedral value (within a tolerance).[18,19] Dihedral instances belonging to the same dihedral type may have different dihedral value signs. (*i.e.*, they may be local mirror images of each other).[18,19] Each dihedral type has its own set of torsion force constant values that apply to all dihedral instances in that dihedral type. According to this scheme, dihedral model potentials should be defined in such a way that mirror image dihedral instances belonging to the same dihedral type are accurately described by the same set of torsion force constant values.

This can be accomplished by defining the mirror image descriptor

$$-1 \leq S_{\text{instance}} \leq +1 \tag{92}$$

in such a way that

$$S_{\text{instance}}\left[\left\{\vec{R}_{\text{A}}\right\}\right] = -S_{\text{instance}}\left[\left\{\vec{R}_{\text{A}}^{\text{reflected}}\right\}\right] \tag{93}$$

The general strategy is as follows. The torsion potential is decomposed into individual torsion modes such that each torsion mode potential has unchanged absolute value when the geometry is reflected. If the torsion mode would change sign upon reflection, it is multiplied by S_{instance} so that the product is invariant when the geometry is reflected. Each of these torsion modes has its own force constant value that is invariant upon reflection.

For example, suppose that $\phi_{\text{eq}} = -30°$ in the first instance and $\phi_{\text{eq}} = +30°$ in the second instance of two dihedrals of the same type. Then, a dihedral displacement to $\phi = 10°$ in the first instance is chemically equivalent to a dihedral displacement to $\phi = -10°$ in the second instance, and these should produce the same dihedral torsion potential values. Because cosine is an even function, for cosine modes we have

$$\cos[m(\phi - \phi_{\text{eq}})] = \cos[m(-\phi - (-\phi_{\text{eq}}))] \tag{94}$$

which give equal contributions to the torsion potential. In this example, $\cos[m(10°-(-30°))] = \cos[m(-10°-30°)] = \cos[m40°]$. Because sine is an odd function, we have to multiply it by the corresponding value of S_{instance} to recover equal contributions to the torsion potential upon reflection:

$$\sin[m(\phi - \phi_{\text{eq}})] = -\sin[m(-\phi - (-\phi_{\text{eq}}))] \tag{95}$$

In this case, $S_{\text{instance}} = -1$ in the first instance and $S_{\text{instance}} = +1$ in the second instance to give sine function terms $-\sin[m(10° - (-30°))] = \sin[m(-10° - 30°)] = -\sin[m40°]$.

The case for which $\sin[\phi_{\text{eq}}] \rightarrow 0$ deserves some further clarification. In this case, both mirror images have the same equilibrium dihedral value, so the two distinct mirror images cannot be distinguished by the sign of the equilibrium dihedral's value. We are thus faced with several viable alternative choices. One possible solution is to turn off the sine modes by having $S_{\text{instance}} \rightarrow 0$ when $\sin[\phi_{\text{eq}}] \rightarrow 0$; this has the effect of making the torsion potential symmetrical about $\phi = \phi_{\text{eq}}$. In option # 1, this is done in a discrete way using the signum function:

$$S_{\text{instance}} = \text{signum}\left[\sin\left[\phi_{\text{eq}}\right]\right] = \begin{cases} 0 & \text{if } \sin\left[\phi_{\text{eq}}\right] = 0 \\ +1 & \text{if } \sin\left[\phi_{\text{eq}}\right] > 0 \\ -1 & \text{if } \sin\left[\phi_{\text{eq}}\right] < 0 \end{cases} \tag{96}$$

In option # 2, this is done in a smoothed continuous way using

$$S_{\text{instance}} = \tanh[10^{D} \sin[\phi_{\text{eq}}]] \tag{97}$$

where the smoothing exponent D controls the decimal place for which the smoothing occurs. For example, setting $D = 2$ means the second decimal place (*i.e.*, 0.01 radians) is where the smoothing occurs, while setting $D = 1$ means the first decimal place (*i.e.*, 0.1 radians) is where the smoothing occurs. Option # 2 has the key advantage that the value of S_{instance} is continuously differentiable with respect to continuous changes in ϕ_{eq}. In option # 3, we retain the value of $S_{\text{instance}} = +1$ or -1 to distinguish the two mirror image structures even when $\sin[\phi_{\text{eq}}] \rightarrow 0$. This option is the most comprehensive, but it requires developing an alternate method to distinguish the two mirror images. For example, one could use pattern-matching algorithms to distinguish between the two mirror images.

For simplicity, this article used option #1 (see eqn (96)) to compute the value of S_{instance} used in the ADDT and CADT model potentials. However, these two model potentials are compatible with any of the three options described in the previous paragraph. For the ADLD model potential, option #3 must be used as explained in Section 9 below, because ϕ_{eq} is not defined for linear dihedrals. The ADCO and CACO model potentials do not use any S_{instance}.

4. Angle-damped dihedral torsion (ADDT) model potential

4.1 Derivation including torsion offset potential

If neither contained equilibrium bond angle is linear, eqn (87) can be equivalently re-written (with the help of trigonometric identities in eqn (26) and (27)) as

$$U_{\text{ABCD}}^{\text{ADDT}}[\theta_{\text{ABC}}, \theta_{\text{BCD}}, \phi_{\text{ABCD}}] = \alpha_0 H_0[\theta_{\text{ABC}}, \theta_{\text{BCD}}] + \sum_{n=1}^{n_{\max}} H_n[\theta_{\text{ABC}}, \theta_{\text{BCD}}]\Big(-a_\phi^n \cos[n(\phi_{\text{ABCD}} - \phi_{\text{ABCD}}^{\text{eq}})] + b_\phi^n S_{\text{instance}} \sin[n(\phi_{\text{ABCD}} - \phi_{\text{ABCD}}^{\text{eq}})]\Big) \tag{98}$$

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

For convenience, we regroup the terms that equal zero when $\phi_{ABCD} = \phi^{eq}_{ABCD}$:

$$U^{ADDT}_{ABCD}[\theta_{ABC},\theta_{BCD},\phi_{ABCD}] = \sum_{n=1}^{n_{max}} a^n_\phi (J_n[\theta_{ABC},\theta_{BCD}] - H_n[\theta_{ABC},\theta_{BCD}]) + \sum_{n=1}^{n_{max}} H_n[\theta_{ABC},\theta_{BCD}]\left(a^n_\phi(1-\cos[n(\phi_{ABCD} - \phi^{eq}_{ABCD})]) + b^n_\phi S_{instance}\sin[n(\phi_{ABCD} - \phi^{eq}_{ABCD})]\right) \quad (99)$$

Comparing eqn (98) to (99) shows that

$$\alpha_0 H_0[\theta_{ABC},\theta_{BCD}] = \sum_{n=1}^{n_{max}} a^n_\phi (J_n[\theta_{ABC},\theta_{BCD}]) \quad (100)$$

I call $J_n[\theta_{ABC},\theta_{BCD}] - H_n[\theta_{ABC},\theta_{BCD}]$ the 'torsion offset potential (TOP) for mode n', because it adds a theta-dependent offset potential (aka 'potential shift') that is independent of the dihedral's value.

The combined angle-dihedral coordinate branch equivalency condition requires that

$$H_n[(2\pi - \theta_{ABC}),\theta_{BCD}] = H_n[\theta_{ABC},(2\pi - \theta_{BCD})] = (-1)^n H_n[\theta_{ABC},\theta_{BCD}] \quad (101)$$

$$J_n[(2\pi - \theta_{BCD}),\theta_{BCD}] = J_n[\theta_{ABC},(2\pi - \theta_{BCD})] = J_n[\theta_{ABC},\theta_{BCD}] \quad (102)$$

To ensure the potential energy is non-negative for displaced geometries,

$$J_n[\theta_{ABC},\theta_{BCD}] \geq 0 \quad (103)$$

$$J_n[\theta^{eq}_{ABC},\theta^{eq}_{BCD}] = 1 \quad (104)$$

will be constructed using a 'completing the squares approach'. Following principle (h) of Section 2.4:

$$H_n[\theta^{eq}_{ABC},\theta^{eq}_{BCD}] = 1 \quad (105)$$

Explicit formulas for J_n and H_n are derived as follows:

(1) To satisfy the combined angle-dihedral coordinate branch equivalency condition, we could construct a torsion offset potential (TOP) having the following form:

$$TOP_n[\theta_{ABC},\theta_{BCD}] = J_n[\theta_{ABC},\theta_{BCD}] - H_n[\theta_{ABC},\theta_{BCD}] \quad (106)$$

$$J_n[\theta_{ABC},\theta_{BCD}] = \frac{1}{4}\left(\left(\frac{f^{ABC}_n f^{ABC}_{\lfloor n/2\rfloor_eq}}{f^{ABC}_{n_eq} f^{ABC}_{\lfloor n/2\rfloor}}\right)^2 + \left(\frac{f^{ABC}_{\lfloor n/2\rfloor}}{f^{ABC}_{\lfloor n/2\rfloor_eq}}\right)^2\right)\left(\left(\frac{f^{BCD}_n f^{BCD}_{\lfloor n/2\rfloor_eq}}{f^{BCD}_{n_eq} f^{BCD}_{\lfloor n/2\rfloor}}\right)^2 + \left(\frac{f^{BCD}_{\lfloor n/2\rfloor}}{f^{BCD}_{\lfloor n/2\rfloor_eq}}\right)^2\right) \quad (107)$$

$$H_n[\theta_{ABC},\theta_{BCD}] = \frac{f^{ABC}_n f^{BCD}_n}{f^{ABC}_{n_eq} f^{BCD}_{n_eq}} \quad (108)$$

Here, $\lfloor x \rfloor$ denotes floor[x], which is the largest integer less than or equal to x. Eqn (106)–(108) follows a 'completing the squares' strategy that guarantees that

$$TOP_n[\theta_{ABC},\theta_{BCD}] \geq 0 \quad (109)$$

Proof: Since

$$\frac{1}{2}\left(\frac{f^{ABC}_n f^{ABC}_{\lfloor n/2\rfloor_eq}}{f^{ABC}_{n_eq} f^{ABC}_{\lfloor n/2\rfloor}}\right)^2 + \frac{1}{2}\left(\frac{f^{ABC}_{\lfloor n/2\rfloor}}{f^{ABC}_{\lfloor n/2\rfloor_eq}}\right)^2 - \frac{f^{ABC}_n}{f^{ABC}_{n_eq}} = \frac{1}{2}\left(\frac{f^{ABC}_n f^{ABC}_{\lfloor n/2\rfloor_eq}}{f^{ABC}_{n_eq} f^{ABC}_{\lfloor n/2\rfloor}} - \frac{f^{ABC}_{\lfloor n/2\rfloor}}{f^{ABC}_{\lfloor n/2\rfloor_eq}}\right)^2 \geq 0 \quad (110)$$

it immediately follows that

$$\frac{1}{2}\left(\frac{f^{ABC}_n f^{ABC}_{\lfloor n/2\rfloor_eq}}{f^{ABC}_{n_eq} f^{ABC}_{\lfloor n/2\rfloor}}\right)^2 + \frac{1}{2}\left(\frac{f^{ABC}_{\lfloor n/2\rfloor}}{f^{ABC}_{\lfloor n/2\rfloor_eq}}\right)^2 \geq \frac{f^{ABC}_n}{f^{ABC}_{n_eq}} \quad (111)$$

and similarly for the BCD angle. Substituting eqn (111) and the analogous eqn for BCD yields eqn (109).

(2) Consider the following generalization of eqn (110)

$$\frac{1}{2}\left(\frac{f^{ABC}_n f^{ABC}_{j_eq}}{f^{ABC}_{n_eq} f^{ABC}_{j}}\right)^2 + \frac{1}{2}\left(\frac{f^{ABC}_{j}}{f^{ABC}_{j_eq}}\right)^2 - \frac{f^{ABC}_n}{f^{ABC}_{n_eq}} = \frac{1}{2}\left(\frac{f^{ABC}_n f^{ABC}_{j_eq}}{f^{ABC}_{n_eq} f^{ABC}_{j}} - \frac{f^{ABC}_{j}}{f^{ABC}_{j_eq}}\right)^2 \geq 0 \quad (112)$$

where $0 \leq j \leq n$. In the limit $\mathcal{K}^{eq}_{ABC}$ approaches a small positive number (*i.e.*, $\mathcal{K}^{eq}_{ABC} \to \varepsilon$), then $1/f^{ABC}_{j_eq}$ becomes proportional to ε^{-j}, and $f^{ABC}_{j_eq}/f^{ABC}_{n_eq}$ becomes proportional to ε^{j-n}. To keep the TOP from becoming excessively large, it is desirable therefore to formulate the TOP such that $max[j,(n-j)]$ is as small as feasible. This follows principle (g) of Section 2.4 in order to construct the TOP in a way that produces the smallest energy increase. Immediately, this yields j = floor[$n/2$] or ceiling[$n/2$]. Due to the asymptotic matching between $f^{ABC}_{j_eq}$ and $f^{ABC}_{n_eq}$, there is a partial cancellation of amplitudes within the $f^{ABC}_{j_eq}/f^{ABC}_{n_eq}$. Therefore, to minimize the magnitude of the TOP, for odd n the smaller integer (*i.e.*, floor[n/2]) should be assigned to the $1/f^{ABC}_{j_eq}$ while the larger integer (*i.e.*, ceiling[$n/2$]) should be assigned to the $f^{ABC}_{j_eq}/f^{ABC}_{n_eq}$ term. For even n, floor[$n/2$] and ceiling[$n/2$] are obviously equal. Thus we have j = floor[$n/2$] which yields eqn (107) and (110) above.

(3) The specific form of eqn (107) was chosen to avoid correlation bias. Whenever $(f^{ABC}_n/f^{ABC}_{n_eq}) \neq 1$, the ABC bond angle has been displaced away from its equilibrium value; consequently, the TOP should be significantly positive regardless of whether $(f^{BCD}_n/f^{BCD}_{n_eq})$ is < 1, ~1, or >1. Similarly, the TOP should be significantly positive whenever $(f^{BCD}_n/f^{BCD}_{n_eq}) \neq 1$, regardless of whether $(f^{ABC}_n/f^{ABC}_{n_eq})$ is <1, ~1, or >1. The form of eqn (107) satisfies this requirement.

(4) Most importantly, these choices for J_n and H_n guarantee that the minimum TOP_n value occurs for

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$\mathrm{TOP}_n[\theta^{\mathrm{eq}}_{\mathrm{ABC}},\theta^{\mathrm{eq}}_{\mathrm{BCD}}] = 0 \tag{113}$$

and that the minimum value for $(1 - \cos[n(\phi_{\mathrm{ABCD}} - \phi^{\mathrm{eq}}_{\mathrm{ABCD}})])$ is 0 and occurs when $\phi_{\mathrm{ABCD}} = \phi^{\mathrm{eq}}_{\mathrm{ABCD}}$. If a single torsion mode is active (*i.e.*, smart selected), then it is generally convenient to impose the following constraint to ensure the energy increases for all displacements away from the equilibrium ABC and BCD bond angles and for all displacements away from the equilibrium ABCD dihedral value:

$$a^n_\phi \geq 0 \tag{114}$$

What is the physical significance of the TOP? As an example, consider the HNCO molecule. For reasons explained in a companion article, as the NCO bond angle approaches linear (*i.e.*, π, 180°) the force exerted by the angle-bending potential must become zero.[20] However, Fig. 9 shows the force exerted within the molecule along the angle-scan energy curve is markedly not zero as the NCO bond angle approaches linear (*i.e.*, π, 180°). This is not a paradox but rather a manifestation of the new energy term described by the TOP. Section 10.4 below shows the TOP gives rise to a new and unique physical phenomenon called 'slip torsion' that would not exist without it.

The following limits, which retain the leading term in the Taylor series expansion, should be used to avoid division by zero in all equations containing the ratios $f^{\mathrm{ABC}}_n/f^{\mathrm{ABC}}_{\lfloor n/2\rfloor}$ and/or $f^{\mathrm{BCD}}_n/f^{\mathrm{BCD}}_{\lfloor n/2\rfloor}$ (*e.g.*, eqn (107) and others) as one or both bond angles approach linearity:

$$\lim_{Ж_{\mathrm{ABC}}\to 0}\frac{f^{\mathrm{ABC}}_n}{f^{\mathrm{ABC}}_{\lfloor n/2\rfloor}} = \lim_{Ж_{\mathrm{ABC}}\to 0}\frac{P_n[Ж_{\mathrm{ABC}}]}{P_{\lfloor n/2\rfloor}[Ж_{\mathrm{ABC}}]} = \mu^n_{\lfloor n/2\rfloor}(Ж_{\mathrm{ABC}})^{n-\lfloor n/2\rfloor} \tag{115}$$

$$\lim_{Ж_{\mathrm{BCD}}\to 0}\frac{f^{\mathrm{BCD}}_n}{f^{\mathrm{BCD}}_{\lfloor n/2\rfloor}} = \lim_{Ж_{\mathrm{BCD}}\to 0}\frac{P_n[Ж_{\mathrm{BCD}}]}{P_{\lfloor n/2\rfloor}[Ж_{\mathrm{BCD}}]} = \mu^n_{\lfloor n/2\rfloor}(Ж_{\mathrm{BCD}})^{n-\lfloor n/2\rfloor} \tag{116}$$

$$\mu^2_1 = 3,\ \mu^3_1 = 6,\ \mu^4_2 = 10/3 \tag{117}$$

I recommend using the leading order expansion shown in eqn (115) iff $Ж_{\mathrm{ABC}} < 0.001$. I recommend using the leading order expansion shown in eqn (116) iff $Ж_{\mathrm{BCD}} < 0.001$. For example, if $Ж_{\mathrm{ABC}} < 0.001$ then $f^{\mathrm{ABC}}_4/f^{\mathrm{ABC}}_2$ to leading order equals $(10/3)(Ж_{\mathrm{ABC}})^2$.

4.2 Derivation of the first seven independent torsion modes for rotatable dihedrals

To ensure the quantum-mechanically-computed ground-state geometry is an equilibrium geometry of the fitted classical forcefield, the following constraint should be imposed:

$$\left.\frac{\partial U^{\mathrm{ADDT}}_{\mathrm{ABCD}}[\theta^{\mathrm{eq}}_{\mathrm{ABC}},\theta^{\mathrm{eq}}_{\mathrm{BCD}},\phi_{\mathrm{ABCD}}]}{\partial\phi_{\mathrm{ABCD}}}\right|_{\phi_{\mathrm{ABCD}}=\phi^{\mathrm{eq}}_{\mathrm{ABCD}}} = 0 \tag{118}$$

Substituting eqn (99), (104), and (105) into (118) yields the following constraint

$$\sum_{n=1}^{n_{\max}} n b^n_\phi = 0 \tag{119}$$

The constraint shown in eqn (119) has the effect of reducing the number of independent sine modes from $n_{\max}$ to $(n_{\max} - 1)$, and this reduces the total number of independent rotatable dihedral modes (cosine plus sine) from $2n_{\max}$ to $(2n_{\max} - 1)$. One possible strategy to handle this reduced number of degrees of freedom is to use a linear regression algorithm that can incorporate any form of linear constraints (such as eqn (119)) when optimizing the force constant values. However, the glmnet implementation of the LASSO method handles bounds on the model parameters rather than general linear constraints.[21–23] Since we use the LASSO method to address the multicollinearity and feature selection problems when optimizing the force constants,[18] the sine torsion modes should be written in a way that makes them independent of each other without the need for a separate linear constraint.

This can be accomplished as follows. We choose $n_{\max} = 4$, because it gives an excellent compromise between accuracy and computational costs. In this case, there are $n_{\max} = 4$ independent cosine modes, which we label as modes 1 to 4. There are $(n_{\max} - 1) = 3$ independent sine modes, which we label as modes 5, 6, and 7. For the first $n_{\max} = 4$ modes, we can use the cosine torsion modes shown in eqn (123), because these already satisfy the zero torsion force condition at the equilibrium geometry. The remaining $(n_{\max} - 1) = 3$ independent torsion modes must be constructed as orthogonal linear combinations of the sine torsion modes, where each independent torsion mode has a single independent *k*-value (aka 'force constant value').

First, we define

$$U^{\sin}_n = \sin[n(\phi - \phi_{\mathrm{eq}})] \tag{120}$$

The linear combinations $3U^{\sin}_1 - U^{\sin}_3$ and $2U^{\sin}_2 - U^{\sin}_4$ yield zero torsion force at the equilibrium geometry. Moreover, these two linear combinations are orthogonal to each other; that is, they have zero overlap integral between them. The overlap integral between two real-valued functions $g_i[\phi_{\mathrm{ABCD}}]$ and $g_j[\phi_{\mathrm{ABCD}}]$ is defined as

$$\langle g_i[\phi_{\mathrm{ABCD}}]|g_j[\phi_{\mathrm{ABCD}}]\rangle = \int_{-\pi}^{\pi} g_i[\phi_{\mathrm{ABCD}}]g_j[\phi_{\mathrm{ABCD}}]\mathrm{d}\phi_{\mathrm{ABCD}} \tag{121}$$

The seventh mode takes a bit more ingenuity to construct. Within the subspace of functions spanned by linear combinations of $U^{\sin}_1$ and $U^{\sin}_3$, the linear combination $U^{\sin}_1 + 3U^{\sin}_3$ is orthogonal to $3U^{\sin}_1 - U^{\sin}_3$. Within the subspace of functions spanned by linear combinations of $U^{\sin}_2$ and $U^{\sin}_4$, the linear combination $U^{\sin}_2 + 2U^{\sin}_4$ is orthogonal to $2U^{\sin}_2 - U^{\sin}_4$. These can be combined to form the linear combination $U^{\sin}_1 + 3U^{\sin}_3 - U^{\sin}_2 - 2U^{\sin}_4$ that yields zero force at the equilibrium geometry and is orthogonal to the other six independent torsion modes.

Finally, it is useful to scale each of the independent torsion modes so that its root-mean-squared deviation from its average value is $\sqrt{1/2}$. This puts all of the different independent torsion

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$P_m^{\mathrm{DT}}[\phi] = \begin{cases} -\cos\left[m\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right] & \text{for } m = 1 \text{ to } 4 \\ \dfrac{\left(3\sin\left[\phi-\phi_{\mathrm{eq}}\right]-\sin\left[3\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right)}{\sqrt{10}} & \text{for } m = 5 \\ \dfrac{\left(2\sin\left[2\left(\phi-\phi_{\mathrm{eq}}\right)\right]-\sin\left[4\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right)}{\sqrt{5}} & \text{for } m = 6 \\ \dfrac{\left(\sin\left[\phi-\phi_{\mathrm{eq}}\right]-\sin\left[2\left(\phi-\phi_{\mathrm{eq}}\right)\right]+3\sin\left[3\left(\phi-\phi_{\mathrm{eq}}\right)\right]-2\sin\left[4\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right)}{\sqrt{15}} & \text{for } m = 7 \end{cases} \tag{122}$$

modes on the same normalization scale. Considering only the dihedral dependence, this gives the first seven independent torsion modes of the rotatable dihedral torsion potential:

$$G_m^{\mathrm{DT}}[\phi] = \begin{cases} \left(1-\cos\left[m\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right) & \text{for } m = 1 \text{ to } 4 \\ P_m^{\mathrm{DT}}[\phi] & \text{for } m = 5 \text{ to } 7 \end{cases} \tag{123}$$

where ϕ is the current dihedral angle, and ϕ_{eq} is the reference dihedral angle from the equilibrium state of the structure. $P_m^{\mathrm{DT}}[\phi]$ is defined to have an average value of zero:

$$P_m^{\mathrm{DT}}[\phi] = G_m^{\mathrm{DT}}[\phi] - G_m^{\mathrm{DT,avg}} \tag{124}$$

$$G_m^{\mathrm{DT,avg}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} G_m^{\mathrm{DT}}[\phi]d\phi \tag{125}$$

Fig. 4 plots each of these seven modes for expanding the dihedral torsion potential function.

Including the torsion offset potential and the angle-damping factors, this yields the following seven ADDT modes:

$$U_{\mathrm{mode_1}}^{\mathrm{ADDT}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] = k_\phi^1\left(J_1[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] - H_1[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}]\cos\left[\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right) \tag{126}$$

$$U_{\mathrm{mode_2}}^{\mathrm{ADDT}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] = k_\phi^2\left(J_2[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] - H_2[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}]\cos\left[2\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right) \tag{127}$$

$$U_{\mathrm{mode_3}}^{\mathrm{ADDT}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] = k_\phi^3\left(J_3[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] - H_3[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}]\cos\left[3\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right) \tag{128}$$

$$U_{\mathrm{mode_4}}^{\mathrm{ADDT}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] = k_\phi^4\left(J_4[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] - H_4[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}]\cos\left[4\left(\phi-\phi_{\mathrm{eq}}\right)\right]\right) \tag{129}$$

$$U_{\mathrm{mode_5}}^{\mathrm{ADDT}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] = \frac{k_\phi^5}{\sqrt{10}}S_{\mathrm{instance}}\left(3\sin\left[\phi-\phi_{\mathrm{eq}}\right]H_1[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] - \sin\left[3\left(\phi-\phi_{\mathrm{eq}}\right)\right]H_3[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}]\right) \tag{130}$$

$$U_{\mathrm{mode_6}}^{\mathrm{ADDT}}[\text{Ж}_{ABC}, \text{Ж}_{BCD}, \phi_{ABCD}] = \frac{k_\phi^6}{\sqrt{5}}S_{\mathrm{instance}}\left(2\sin\left[2\left(\phi-\phi_{eq}\right)\right]H_2[\theta_{ABC}, \theta_{BCD}] - \sin\left[4\left(\phi-\phi_{eq}\right)\right]H_4[\theta_{ABC}, \theta_{BCD}]\right) \tag{131}$$

$$U_{\mathrm{mode_7}}^{\mathrm{ADDT}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] = \frac{k_\phi^7}{\sqrt{15}}S_{\mathrm{instance}}\begin{pmatrix}\sin\left[\phi-\phi_{\mathrm{eq}}\right]H_1[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] - \sin\left[2\left(\phi-\phi_{\mathrm{eq}}\right)\right]H_2[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] \\ +3\sin\left[3\left(\phi-\phi_{\mathrm{eq}}\right)\right]H_3[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] - 2\sin\left[4\left(\phi-\phi_{\mathrm{eq}}\right)\right]H_4[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}]\end{pmatrix} \tag{132}$$

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

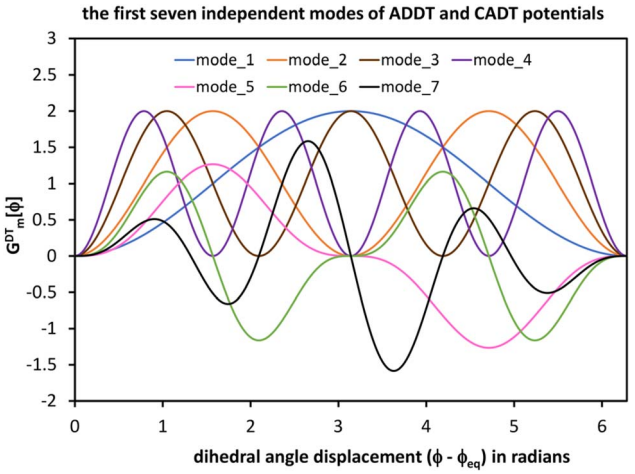

Fig. 4 Plot showing the dihedral torsion potential $G_m^{\mathrm{DT}}[\phi]$ for each mode *versus* the dihedral displacement. Modes 1 to 4 are cosine modes, and modes 5 to 7 are sine modes.

After dihedral mode smart selection (see Section 7), this yields

$$U^{\mathrm{ADDT}}_{\mathrm{ABCD}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] - U^{\mathrm{ADDT}}_{\mathrm{ABCD}}[\text{Ж}^{\mathrm{eq}}_{\mathrm{ABC}}, \text{Ж}^{\mathrm{eq}}_{\mathrm{BCD}}, \phi^{\mathrm{eq}}_{\mathrm{ABCD}}] = \sum_{j=1}^{N^{\mathrm{ABCD}}_{\mathrm{active_modes}}} U^{\mathrm{ADDT}}_{\mathrm{mode_}mj}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] \quad (133)$$

where $N^{\mathrm{ABCD}}_{\mathrm{active_modes}}$ is the number of active modes for dihedral ABCD.

5. Constant amplitude dihedral torsion (CADT) model potential

As shown in Section 10.2 below, angles with $\theta^{\mathrm{eq}}_{\mathrm{ABC}} < 130°$ are typically much stiffer than angles with $\theta^{\mathrm{eq}}_{\mathrm{ABC}} > 160°$. As a result of this increased stiffness and also that they are farther from 180° to begin with, it is relatively unlikely that angles having $\theta^{\mathrm{eq}}_{\mathrm{ABC}} < 130°$ will reach 180° during a typical classical molecular dynamics simulation. Also, the relatively large increase in energy from $\theta^{\mathrm{eq}}_{\mathrm{ABC}} < 130°$ to $\theta^{\mathrm{eq}}_{\mathrm{ABC}} = 180°$ makes reaching 180° not 'thermally accessible' during classical Monte Carlo simulations. Depending on the situation, there may or may not be a ground state crossover as the ABC angle changes from $\theta_{\mathrm{ABC}} < 130°$ to $\theta_{\mathrm{ABC}} > 160°$. For example, carbon atoms having sp^2 hybridization typically have bond angles of ~120° while carbon atoms having sp^1 hybridization typically have bond angles of ~180°. Consequently, angles having $\theta^{\mathrm{eq}}_{\mathrm{ABC}} < 130°$ are in some sense 'far removed' from the dynamics near $\theta_{\mathrm{ABC}} \approx 180°$.

Why set this angle threshold value at 130°? Since sp^2 hybridization (which typically yields ~120° bond angles) is fundamentally different than sp^1 hybridization (which typically yields ~180° bond angles), we infer the angle threshold value should be >120° and <180°. Since bond angles >135° are closer to 180° (parallel bonds) than to 90° (perpendicular bonds), it follows that the angle threshold should probably be set to a value ≤ 135°.

Now, sp^2 hybridized angles will exhibit some statistical fluctuations giving some sp^2 hybridized equilibrium bond angle values slightly larger than 120° and also some others slightly smaller than 120°. Because the angle threshold value should not divide this group, it follows that the angle threshold value should be greater than or equal to approximately 128°. Now, between the feasible range of approximately 128° to 135°, the precise value of the angle threshold value is a judgement call. An angle threshold value of 130° was chosen as a round number that maximizes the size of the region treated by the ADDT potential.

Accordingly, if both $\theta^{\mathrm{eq}}_{\mathrm{ABC}} < 130°$ and $\theta^{\mathrm{eq}}_{\mathrm{BCD}} < 130°$, then it is reasonable to make the simplifications

$$f_n^{\mathrm{ABC}} \approx f_{n_\mathrm{eq}}^{\mathrm{ABC}} \approx f_{1}^{\mathrm{ABC}} \approx f_{1_\mathrm{eq}}^{\mathrm{ABC}} \quad (134)$$

$$f_n^{\mathrm{BCD}} \approx f_{n_\mathrm{eq}}^{\mathrm{BCD}} \approx f_{1}^{\mathrm{BCD}} \approx f_{1_\mathrm{eq}}^{\mathrm{BCD}} \quad (135)$$

for $n = 1, 2, 3, 4$. Eqn (134) results both from the asymptotic matching (as described in Section 2.3 above) and from the relatively stiff ABC bond angle which causes $f_n^{\mathrm{ABC}} \approx f_{n_\mathrm{eq}}^{\mathrm{ABC}}$. Eqn (135) results from analogous properties for the BCD bond angle. This yields the approximation

$$J_n[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] \approx H_n[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] \approx 1 \quad (136)$$

These simplifications produce the following constant amplitude dihedral torsion (CADT) potential energy modes:

$$U^{\mathrm{CADT}}_{\mathrm{mode_}m}[\phi_{\mathrm{ABCD}}] = \begin{cases} k_\phi^m\left(1 - \cos\left[m\left(\phi - \phi_{\mathrm{eq}}\right)\right]\right) & \text{for } m = 1 \text{ to } 4 \\ k_\phi^m S_{\mathrm{instance}} P_m^{\mathrm{DT}}[\phi] & \text{for } m = 5 \text{ to } 7 \end{cases} \quad (137)$$

The CADT potential has a force discontinuity if either $\theta_{\mathrm{ABC}} = 180°$ or $\theta_{\mathrm{BCD}} = 180°$, but those angle values are energetically unlikely to be reached during typical molecular dynamics or Monte Carlo simulations when both $\theta^{\mathrm{eq}}_{\mathrm{ABC}} < 130°$ and $\theta^{\mathrm{eq}}_{\mathrm{BCD}} < 130°$.

As explained in the companion article, dihedral types sharing the same middle bond instances can be 'pruned' so that one representative dihedral type is retained per middle bond type in the forcefield.[18] This dihedral pruning preserves symmetry equivalency.[18] As explained in the companion article, this should preferably be done so that the retained representative dihedral type simultaneously has a relatively modest number of instances and relatively large values of $(180° - \theta^{\mathrm{eq}}_{\mathrm{ABC}})$ and $(180° - \theta^{\mathrm{eq}}_{\mathrm{BCD}})$.[18] For most materials, the practical consequence is that the vast majority of dihedrals retained after pruning will use the CADT potentials instead of the ADDT potentials.

As explained in Section 7, dihedral mode smart selection for the CADT potential follows exactly the same procedure and equations as for the ADDT potential. This yields

$$U^{\mathrm{CADT}}_{\mathrm{ABCD}}[\phi_{ABCD}] - U^{\mathrm{CADT}}_{\mathrm{ABCD}}[\phi^{eq}_{ABCD}] = \sum_{j=1}^{N^{\mathrm{ABCD}}_{\mathrm{active_modes}}} U^{\mathrm{CADT}}_{\mathrm{mode_mj}}[\phi_{ABCD}] \quad (138)$$

where $N^{\mathrm{ABCD}}_{\mathrm{active_modes}}$ is the number of active modes for dihedral ABCD.

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

After parameterization, the CADT potential can be equivalently rewritten as the dihedral potential Fourier series:

$$U^{\mathrm{CADT}}_{\mathrm{ABCD}}[\phi_{\mathrm{ABCD}}] - U^{\mathrm{CADT}}_{\mathrm{ABCD}}[\phi^{\mathrm{eq}}_{\mathrm{ABCD}}] = \text{Ħ} - \sum_{n=1}^{4} k^n_\phi \cos\left[n\left(\phi - \phi_{\mathrm{eq}}\right)\right] + S_{\mathrm{instance}} \sum_{n=1}^{4} b^n_\phi \sin\left[n\left(\phi - \phi_{\mathrm{eq}}\right)\right] \quad (139)$$

where

$$\text{Ħ} = \sum_{m=1}^{4} k^m_\phi \quad (140)$$

$$b^1_\phi = \frac{3}{\sqrt{10}} k^5_\phi + \frac{1}{\sqrt{15}} k^7_\phi \quad (141)$$

$$b^2_\phi = \frac{2}{\sqrt{5}} k^6_\phi - \frac{1}{\sqrt{15}} k^7_\phi \quad (142)$$

$$b^3_\phi = \frac{-1}{\sqrt{10}} k^5_\phi + \frac{3}{\sqrt{15}} k^7_\phi \quad (143)$$

$$b^4_\phi = \frac{-1}{\sqrt{5}} k^6_\phi - \frac{2}{\sqrt{15}} k^7_\phi \quad (144)$$

Depending on the particular situation, some terms in eqn (139) may be zero (*i.e.*, omitted) as the consequence of dihedral mode smart selection. In this work, the above series runs up to $n_{\mathrm{max}} = 4$ and spans the full subspace associated with those dihedral potential modes subject to the constraint that $U^{\mathrm{CADT}}_{\mathrm{ABCD}}[\phi^{\mathrm{eq}}_{\mathrm{ABCD}}] = 0$ is a stationary point (*i.e.*, has first derivative equal to zero). It is possible (albeit not usually necessary) to include higher than $n = 4$ terms in the CADT potential and Fourier series expansion.

6. The ADCO and CACO torsion model potentials

Due to symmetry, some dihedrals have a torsion potential that is an even function of the dihedral value as shown in eqn (81). For this dihedral type, it is advantageous to use a model torsion potential that has this same symmetry. Considering only the dihedral dependence, the cosine-only torsion model potentials have the following independent torsion modes:

$$P^{\mathrm{CO}}_n[\phi] = \cos[n\phi] \quad (145)$$

$$G^{\mathrm{CO}}_n[\phi] = \cos[n\phi] - \cos[n\phi^{\mathrm{training}}_{\mathrm{eq}}] \quad (146)$$

Including the torsion offset potential and the angle-damping factors yields the angle-damped cosine only (ADCO) model potential:

$$U^{\mathrm{ADCO}}_{\mathrm{mode_}n}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] = k_{\mathrm{ADCO}} c^{\mathrm{CO}}_n \left(H_n \cos[n\phi] - J_n \cos\left[n\phi^{\mathrm{training}}_{\mathrm{eq}}\right] \right) \quad (147)$$

where k_{ADCO} is an adjustable force constant and c^{CO}_n is the smart selection coefficient for mode n. For each ADCO dihedral type, there is one adjustable force constant k_{ADCO} irrespective of how many ADCO modes were smart selected; the ratios between the individual ADCO modes are controlled by $\{c^{\mathrm{CO}}_n\}$. Here, $\phi^{\mathrm{training}}_{\mathrm{eq}}$ is the equilibrium dihedral value in the QM-optimized low-energy ground-state geometry of the training dataset. After dihedral mode smart selection (see Section 7), this yields

$$U^{\mathrm{ADCO}}_{\mathrm{ABCD}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] - U^{\mathrm{ADCO}}_{\mathrm{ABCD}}\left[\text{Ж}^{\mathrm{eq}}_{\mathrm{ABC}}, \text{Ж}^{\mathrm{eq}}_{\mathrm{BCD}}, \phi^{\mathrm{training}}_{\mathrm{eq,ABCD}}\right] = \sum_{j=1}^{N^{\mathrm{ABCD}}_{\mathrm{active_modes}}} U^{\mathrm{ADCO}}_{\mathrm{mode_n_j}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] \quad (148)$$

where $N^{\mathrm{ABCD}}_{\mathrm{active_modes}}$ is the number of active modes for dihedral ABCD.

Just as described in Section 5 above, if both $\theta^{\mathrm{eq}}_{\mathrm{ABC}} < 130°$ and $\theta^{\mathrm{eq}}_{\mathrm{BCD}} < 130°$, then it is preferable to use the constant amplitude approximation shown in eqn (134)–(136). This simplification yields the constant amplitude cosine only (CACO) model potential:

$$U^{\mathrm{CACO}}_{\mathrm{mode_}n}[\phi_{\mathrm{ABCD}}] = k_{\mathrm{CACO}} c^{\mathrm{CO}}_n (\cos[n\phi] - \cos[n\phi^{\mathrm{training}}_{\mathrm{eq}}]) \quad (149)$$

where k_{CACO} is an adjustable force constant and c^{CO}_n is the smart selection coefficient for mode n. For each CACO dihedral type, there is one adjustable force constant k_{CACO} irrespective of how many CACO modes were smart selected; the ratios between the individual CACO modes are controlled by $\{c^{\mathrm{CO}}_n\}$. After dihedral mode smart selection (see Section 7), this yields

$$U^{\mathrm{CACO}}_{\mathrm{ABCD}}[\phi_{\mathrm{ABCD}}] - U^{\mathrm{CACO}}_{\mathrm{ABCD}}\left[\phi^{\mathrm{training}}_{\mathrm{eq}}\right] = \sum_{j=1}^{N^{\mathrm{ABCD}}_{\mathrm{active_modes}}} U^{\mathrm{CACO}}_{\mathrm{mode_}n_j}[\phi_{\mathrm{ABCD}}] \quad (150)$$

where $N^{\mathrm{ABCD}}_{\mathrm{active_modes}}$ is the number of active modes for dihedral ABCD.

If $\sin[\phi_{\mathrm{eq}}] \neq 0$, due to series truncation (*i.e.*, only including the most important torsion modes) the ADDT and CADT model potentials typically do not yield strictly even torsion potentials even when the molecule itself has such an underlying molecular symmetry. For example, the CADT model potential does not yield a strictly even torsion potential for the hydrogen peroxide (HOOH) molecule, even though the HOOH molecule itself has such an underlying molecular symmetry. Because the CADT model potential has mismatched symmetry for this molecule, it yields a different position of the equilibrium dihedral value on the positive and negative dihedral sides. (This is explicitly shown in Section 10.8 and Table 16 below.) The ADCO and CACO model potentials address this problem by imposing even function symmetry on the torsion model potential. However, the equilibrium dihedral value of the ADCO and CACO model potentials (aka $\phi^{\mathrm{FF}}_{\mathrm{eq}}$) may not exactly match that of the QM-optimized training geometry:

$$\phi^{\mathrm{FF}}_{\mathrm{eq}} \approx \phi^{\mathrm{training}}_{\mathrm{eq}} \quad (151)$$

due to truncation of the ADCO or CACO model potential. As more modes are included in the ADCO and CACO model

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

potentials, ϕ_{eq}^{FF} should asymptotically approach $\phi_{eq}^{training}$ when the true dihedral torsion potential is an even function and the training dataset is a complete torsion scan. As more torsion modes are included in the ADDT and CADT model potentials, these should asymptotically approach the same torsion potentials as the ADCO and CACO model potentials, respectively, when the true dihedral torsion potential is an even function.

Why do the ADCO and CACO model potentials include the products $k_{ADCO}c_n^{CO}$ and $k_{CACO}c_n^{CO}$, respectively, instead of having an independently adjustable force constant value for each torsion mode? The specific reason for this is to more accurately reproduce $\phi_{eq}^{FF} \approx \phi_{eq}^{training}$. By fitting the $\{c_n\}$ explicitly to a complete torsion scan for a chosen dihedral instance during smart mode selection, this optimizes the ratios between these coefficients to more accurately reproduce $\phi_{eq}^{FF} \approx \phi_{eq}^{training}$, because the exact potential energy for such a torsion scan precisely corresponds to a Fourier series expansion. The k_{ADCO} or k_{CACO} value is subsequently computed during force constant optimization involving the full training dataset. If the force constant for each ADCO or CACO torsion mode were to be optimized independently to the full training dataset, this would likely result in less accurate values for $\phi_{eq}^{FF} \approx \phi_{eq}^{training}$, because the exact potential energy for such a training dataset does not necessarily correspond to a Fourier series expansion.

7. Selecting torsion model potentials and torsion modes

7.1 Selecting the best torsion model potential for each dihedral type

Fig. 5 summarizes the overall process for constructing and validating dihedral torsion potentials. The process begins with defining atom types in the material. The second-neighbor-based atom types of Chen and Manz[24] are recommended; however, this is not the only feasible choice of atom types. As explained in a companion article, the next step defines the list of active internal coordinate types and instances.[18] This list of active internal coordinates includes the bond stretches, (optionally) Urey–Bradley stretches, angle bends, dihedral torsions, (optionally) cross-terms, and (optionally) concurrence terms.[18] Dihedral pruning can be used to reduce excessive internal coordinate redundancy.[18]

Each dihedral type should be classified as rotatable, hindered, nonrotatable, or linear. A companion article gives a particular method to classify each dihedral type as non-rotatable, rotatable, hindered, or linear.[18,19] According to this particular classification, a non-rotatable dihedral has a middle bond that is part of a bonded ring. The middle bond for a hindered dihedral is not part of any bonded ring; however, a hindered dihedral cannot be rigidly rotated over the entire range $-\pi < \phi_{ABCD} \leq \pi$ without changing the material's bond connectivity.[18,19] For example, two or more atoms may sterically collide with each other thereby preventing the hindered dihedral from accessing that part of configuration space.[18]

The next step is to select an appropriate torsion model potential for each dihedral type. Fig. 6 illustrates the recommended procedure for doing this. Although the ADDT (or ADCO if $U_{ABCD}^{torsion}[\phi] = U_{ABCD}^{torsion}[-\phi]$) model potential can be applied to dihedral types for which (θ_{ABC}^{eq} and θ_{BCD}^{eq}) < 130°, the greater simplicity of the CADT (or CACO if $U_{ABCD}^{torsion}[\phi] = U_{ABCD}^{torsion}[-\phi]$) model potentials favors their use in this case. Although classical nonreactive forcefields employing dihedral torsion model potentials have been used for several decades, the procedure shown in Fig. 6 is a transformative improvement over prior approaches.

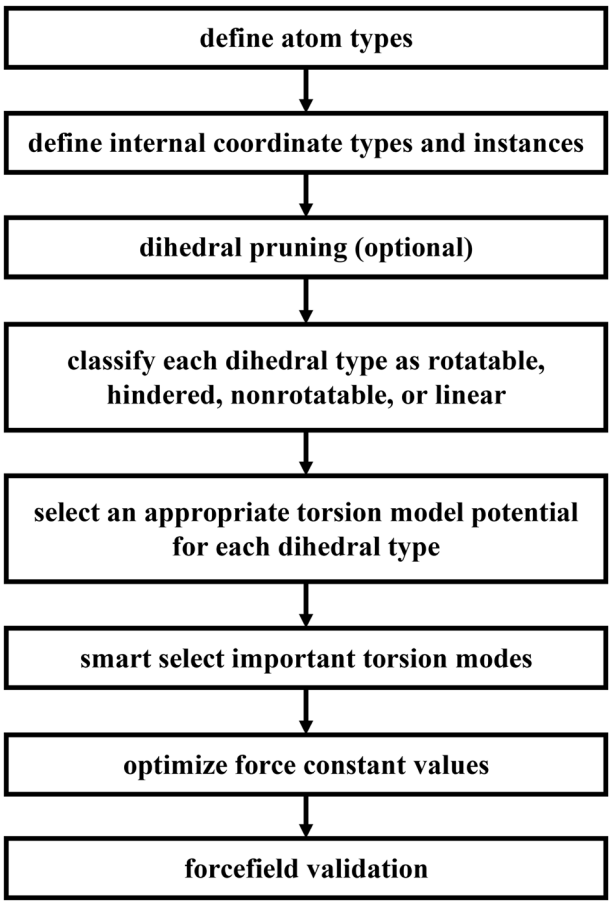

Fig. 5 Overall sequence of steps for defining dihedral types, classifying them, and selecting appropriate torsion model potentials and modes to construct and validate a flexible forcefield.

After the kind of torsion model potential has been selected for each dihedral type, the next step is to smart select the

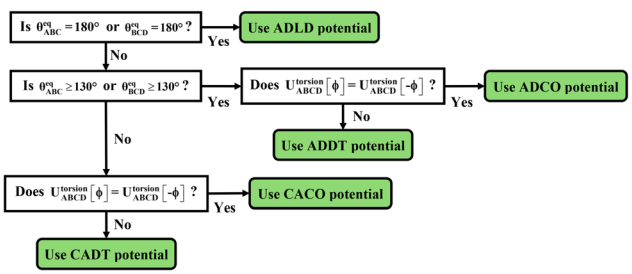

Fig. 6 Flowchart illustrating the procedure to select an appropriate torsion model potential for each dihedral type.

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

important torsion modes. Here, the term 'smart selection' means the important torsion modes are kept while the unimportant torsion modes are omitted from the torsion model potential. For rotatable dihedral types, Section 7.2 explains a procedure for rotatable dihedral mode smart selection using torsion scans. For linear dihedral types, Section 9.2 explains a procedure for selecting specific modes from the linear dihedral torsion potential. Due to the limited range of motion of nonrotatable dihedral types, these can typically be described by the first cosine mode of the ADDT (if (θ^{eq}_{ABC} or θ^{eq}_{BCD}) $\geq 130°$) or CADT (if (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $< 130°$) model potential.[18]

Hindered dihedral types require some special consideration. As explained in a companion article, a dihedral type is classified as 'hindered' iff (a) it is not part of a bonded ring and (b) rigid rotation about its middle bond causes one part of the structure to sterically collide with another part of the structure. Closer examination reveals that such collisions may have two limiting cases. In case # 1, the collision is 'unavoidable', because the colliding groups are structurally tethered in ways that do not allow them to easily pass each other. In case # 2, the collision is 'accidental' and potentially avoidable. Consider long chain-like molecules comprising polymers, large biomolecules (*e.g.*, proteins, RNA, enzymes, fatty acids, polysaccharides, phospholipids, *etc.*), hydrocarbon chains (*e.g.*, petroleum), *etc.* In these cases, one end of the chain may 'accidently' collide with another part of the chain when rigidly rotated (about a dihedral middle bond) due to the particular manner in which the flexible chain has been 'folded'. Let us define a 'range of dihedral rotation' as the connected range of ϕ_{ABCD} values that are not sterically prevented; this refers to the range of ϕ_{ABCD} values that can be reached *via* continuous *relaxed* displacements away from ϕ^{eq}_{ABCD} avoiding steric collision. If this range of rotation about the hindered dihedral's middle bond is severely limited (such as sometimes occurs in case # 1), then it may be adequate to describe this particular hindered dihedral type by the first cosine mode of the ADDT (if (θ^{eq}_{ABC} or θ^{eq}_{BCD}) $\geq 130°$) or CADT (if (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $< 130°$) model potential.[18] On the other if this range of rotation about the hindered dihedral's middle bond is large (such as sometimes occurs in case # 2), then it may be desirable to retain several torsion modes of the ADDT, CADT, ADCO, or CACO model potential to describe this hindered dihedral.

The next step is to optimize the values of all of the force constants in the flexibility model. For reasons explained in companion articles, these should be optimized simultaneously by fitting to a training dataset, rather than optimized sequentially one-at-a-time.[18,20] Finally, the parameterized forcefield should be validated using an appropriate validation dataset that is separate from and independent of the training dataset.

7.2 Rotatable dihedral mode smart selection using torsion scan curves

A relaxed torsion scan curve changes the value of the dihedral ϕ_{ABCD} from $-\pi$ to π in small increments while allowing all other independent geometric parameters (*e.g.*, bond lengths and bond angles) to reoptimize. Each geometry in a relaxed torsion scan is computed from constrained geometry optimization that produces the lowest energy possible for that particular constrained dihedral value.

A rigid torsion scan curve changes the value of the dihedral ϕ_{ABCD} from $-\pi$ to π in small increments while keeping all of the bond lengths and bond angles fixed at their reference values. Unless otherwise specified, these reference values are the corresponding bond length and bond angle values in the material's optimized (low-energy) ground-state geometry. A rigid torsion scan is a series of single-point energy calculations performed at rigid geometries. There are no changes in the values of the angle-damping factors during rigid torsion scans. In general, a complete rigid torsion scan curve is possible if the dihedral is rotatable, but not if the dihedral is non-rotatable or hindered. We did not use rigid torsion scan energy curves for non-rotatable or hindered dihedrals, because these dihedrals have a limited (aka 'restricted') range of motion.[18]

For a rotatable dihedral, the potential energy along a torsion scan curve can be modeled by projecting onto an orthonormal basis of independent torsion modes:

$$\int_{-\pi}^{\pi} \left(\frac{P_m^{DT}[\phi]}{\sqrt{\pi}}\right)\left(\frac{P_n^{DT}[\phi]}{\sqrt{\pi}}\right) d\phi = \delta_{m,n}^{Kronecker} \quad (152)$$

$$\int_{-\pi}^{\pi} \left(\frac{P_m^{CO}[\phi]}{\sqrt{\pi}}\right)\left(\frac{P_n^{CO}[\phi]}{\sqrt{\pi}}\right) d\phi = \delta_{m,n}^{Kronecker} \quad (153)$$

Let $E^{QM}_{RTS}[\phi]$ be the QM-computed energy of the material along the torsion scan curve for rotatable dihedral ABCD. This torsion scan curve is conducted using T dihedral values equally spaced over the range $(-\pi,\pi]$. Let $E^{QM_avg}_{RTS}$ be the average value

$$E_{RTS}^{QM_avg} = \frac{1}{T}\sum_{j=1}^{T} E_{RTS}^{QM}[\phi_j] \quad (154)$$

and w the self-overlap integral

$$w = \int_{-\pi}^{\pi} \left(E_{RTS}^{QM}[\phi] - E_{RTS}^{QM_avg}\right)^2 d\phi \approx \left(\frac{2\pi}{T}\right)\sum_{j=1}^{T} \left(E_{RTS}^{QM}[\phi_j] - E_{RTS}^{QM_avg}\right)^2 \quad (155)$$

The 'QM torsion norm' is the root-mean-squared value of $(E^{QM}_{RTS}[\phi] - E^{QM_avg}_{RTS})$:

$$\text{torsion_norm}_{QM} = \sqrt{\left(\frac{1}{T}\right)\sum_{j=1}^{T} \left(E_{RTS}^{QM}[\phi_j] - E_{RTS}^{QM_avg}\right)^2} = \sqrt{\frac{w}{2\pi}} \quad (156)$$

The QM torsion barrier is defined as the maximum energy minus the minimum energy along the torsion scan curve:

$$\text{torsion_barrier}_{QM} = \max_{-\pi<\phi\leq\pi}\left[E_{RTS}^{QM}[\phi]\right] - \min_{-\pi<\phi\leq\pi}\left[E_{RTS}^{QM}[\phi]\right] \quad (157)$$

Using a complete set of dihedral-torsion (DT) projectors, the QM potential for the torsion scan curve can be expanded in terms of the orthogonal basis set as

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi] = E_{\mathrm{RTS}}^{\mathrm{QM_avg}} + \sqrt{w}\sum_{m=1}^{\infty} c_m^{\mathrm{DT}} \frac{P_m^{\mathrm{DT}}[\phi]}{\sqrt{\pi}} \tag{158}$$

Since the cosine-only (CO) projectors are even functions of ϕ, the CO projectors can provide a complete expansion of the QM potential along the torsion scan curve iff the following symmetry descriptor

$$\mathrm{sym_value} = \frac{1}{2}\sqrt{\frac{\sum_{j=1}^{T}\left(E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi_j] - E_{\mathrm{RTS}}^{\mathrm{QM}}[-\phi_j]\right)^2}{\sum_{j=1}^{T}\left(E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi_j] - E_{\mathrm{RTS}}^{\mathrm{QM_avg}}\right)^2}} \tag{159}$$

equals zero. Iff sym_value equals zero, then in this case the QM potential can be expanded using the CO projectors as

$$E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi] = E_{\mathrm{RTS}}^{\mathrm{QM_avg}} + \sqrt{w}\sum_{m=1}^{\infty} c_m^{\mathrm{CO}} \frac{P_m^{\mathrm{CO}}[\phi]}{\sqrt{\pi}} \tag{160}$$

The model potential for the torsion scan curve can be expanded in terms of the orthogonal basis set as

$$E_{\mathrm{RTS}}^{\mathrm{model}}[\phi] = E_{\mathrm{RTS}}^{\mathrm{model_avg}} + \sqrt{w}\sum_{m=1}^{m_{\max}} c_m \frac{P_m[\phi]}{\sqrt{\pi}} \tag{161}$$

$$E_{\mathrm{RTS}}^{\mathrm{model}}[\phi] - E_{\mathrm{RTS}}^{\mathrm{model}}\left[\phi_{\mathrm{eq}}^{\mathrm{training}}\right] = \sqrt{w}\sum_{m=1}^{m_{\max}} c_m \frac{G_m[\phi]}{\sqrt{\pi}} \tag{162}$$

where

$$E_{\mathrm{RTS}}^{\mathrm{model_avg}} = \frac{1}{T}\sum_{j=1}^{T} E_{\mathrm{RTS}}^{\mathrm{model}}[\phi_j] \tag{163}$$

For the ADDT and CADT model potentials, $P_{\mathrm{m}} = P_{\mathrm{m}}^{\mathrm{DT}}$ is used in eqn (161) and (164), and $G_{\mathrm{m}} = G_{\mathrm{m}}^{\mathrm{DT}}$ is used in eqn (162). For the ADCO and CACO model potentials, $P_{\mathrm{m}} = P_{\mathrm{m}}^{\mathrm{CO}}$ is used in eqn (161) and (164), and $G_{\mathrm{m}} = G_{\mathrm{m}}^{\mathrm{CO}}$ is used in eqn (162). The coefficients are given by

$$\begin{aligned} c_m &= \int_{-\pi}^{\pi} \frac{P_m[\phi]}{\sqrt{\pi}}\left(\frac{E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi] - E_{\mathrm{RTS}}^{\mathrm{QM_avg}}}{\sqrt{w}}\right)\mathrm{d}\phi \\ &\approx \left(\frac{2\pi}{T}\right)\sum_{j=1}^{T}\frac{P_m[\phi_j]}{\sqrt{\pi}}\left(\frac{E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi_j] - E_{\mathrm{RTS}}^{\mathrm{QM_avg}}}{\sqrt{w}}\right) \end{aligned} \tag{164}$$

For the torsion modes included in the model potential, the expansion coefficients $\{c_{\mathrm{m}}\}$ are the same for the QM and model potentials along the torsion scan curve. The model potential neglects some of the less important (*i.e.*, negligible) torsion modes while the QM potential is formally expanded (see eqn (158)) as a nontruncated sum over all possible torsion modes.

The 'model torsion norm' is the root-mean-squared value of $(E_{\mathrm{RTS}}^{\mathrm{model}}[\phi_j] - E_{\mathrm{RTS}}^{\mathrm{model_avg}})$:

$$\mathrm{torsion_norm}_{\mathrm{model}} = \sqrt{\left(\frac{1}{T}\right)\sum_{j=1}^{T}\left(E_{\mathrm{RTS}}^{\mathrm{model}}[\phi_j] - E_{\mathrm{RTS}}^{\mathrm{model_avg}}\right)^2} \tag{165}$$

As derived in ESI Section S12,† the QM and model torsion norms are related by

$$\mathrm{torsion_norm}_{\mathrm{model}} = \sqrt{\frac{w}{2\pi}\sum_{m=1}^{m_{\max}}\left(c_m{}^2\right)} \tag{166}$$

$$\mathrm{torsion_norm}_{\mathrm{model}} = \mathrm{torsion_norm}_{\mathrm{QM}}\sqrt{\mathrm{SumCSq}} \tag{167}$$

where SumCSq is defined as

$$\mathrm{SumCSq} = \sum_{m=1}^{m_{\max}}\left(c_m{}^2\right) \tag{168}$$

Since $\{c_m\}$ are the coefficients for projection onto an orthonormal basis set, it follows that

$$0 \le \mathrm{SumCSq} \le 1 \tag{169}$$

As derived in ESI Section S12,† the full set of mode coefficients for the normalized QM potential along the torsion scan curve satisfies

$$\sum_{m=1}^{\infty}\left(\left(c_m^{\mathrm{DT}}\right)^2\right) = 1 \tag{170}$$

Hence, the SumCSq value can be interpreted as the fraction of the QM torsion scan curve that is recovered by the model potential. Iff SumCSq $\approx$ 1 and $E_{\mathrm{RTS}}^{\mathrm{model_avg}} \approx E_{\mathrm{RTS}}^{\mathrm{QM_avg}}$, then the model potential provides an adequate approximation of the QM potential along the torsion scan curve:

$$E_{\mathrm{RTS}}^{\mathrm{model}}[\phi] \approx E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi] - E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi_{\mathrm{eq}}^{\mathrm{training}}] \tag{171}$$

Here, we consider the general case in which $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ may potentially be different in value than $E_{\mathrm{RTS}}^{\mathrm{QM_avg}}$. This leads to three scenarios. Scenario #1: this scenario chooses the value of $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ such that

$$E_{\mathrm{RTS}}^{\mathrm{model}}[\phi_{\mathrm{eq}}^{\mathrm{training}}] = 0 \tag{172}$$

which makes the left and right sides of eqn (171) exactly equal to each other at the point $\phi = \phi_{\mathrm{eq}}^{\mathrm{training}}$. For the ADDT and CADT model potentials, this scenario corresponds to the choice

$$E_{\mathrm{RTS}}^{\mathrm{model_avg}} = \frac{\sqrt{w}}{\sqrt{\pi}}\sum_{m=1}^{m_{\max}} c_m^{\mathrm{DT}} G_m^{\mathrm{DT,avg}} \tag{173}$$

For the ADCO and CACO model potentials, this scenario corresponds to the choice

$$E_{\mathrm{RTS}}^{\mathrm{model_avg}} = -\frac{\sqrt{w}}{\sqrt{\pi}}\sum_{n=1}^{n_{\max}} c_n^{\mathrm{CO}} \cos\left[n\phi_{\mathrm{eq}}^{\mathrm{training}}\right] \tag{174}$$

Scenario #2: this scenario chooses the value of $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ such that

$$E_{\mathrm{RTS}}^{\mathrm{model_avg}} = E_{\mathrm{RTS}}^{\mathrm{QM_avg}} - E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi_{\mathrm{eq}}^{\mathrm{training}}] \tag{175}$$

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

which makes the averages of the left and right sides of eqn (171) exactly equal to each other. Using a complete set of orthonormal DT projectors to expand the QM potential along the torsion scan curve yields the untruncated expansion:

$$E_{\mathrm{RTS}}^{\mathrm{QM_avg}} - E_{\mathrm{RTS}}^{\mathrm{QM}}\left[\phi_{\mathrm{eq}}^{\mathrm{training}}\right] = \frac{\sqrt{w}}{\sqrt{\pi}} \sum_{m=1}^{\infty} c_m^{\mathrm{DT}} G_m^{\mathrm{DT,avg}} \tag{176}$$

Scenario #3: this scenario (denoted as "other") encompasses any situations in which $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ is chosen to satisfy any criteria different from Scenarios #1 and #2 described above.

Comparing eqn (173), (175), and (176), the difference in $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ for scenarios #1 and #2 using DT projectors is

$$E_{\mathrm{RTS}}^{\mathrm{model_avg}}[\text{Scenario } \#2] - E_{\mathrm{RTS}}^{\mathrm{model_avg}}[\text{Scenario } \#1] = \frac{\sqrt{w}}{\sqrt{\pi}} \sum_{m=(m_{\max}+1)}^{\infty} c_m^{\mathrm{DT}} G_m^{\mathrm{DT,avg}} \tag{177}$$

Examining eqn (177), $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ exactly coincides for the two scenarios if the modal coefficients c_m are zero for all values of $m > m_{\max}$. If the modal coefficients c_m^{DT} are nearly zero for all values of $m > m_{\max}$, then $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ approximately coincides for the two scenarios.

The situation is slightly more complicated for the CO projectors than for the DT projectors. If *sym*_value (see eqn (159)) equals zero, then in this case the average QM potential can be expanded as

$$E_{\mathrm{RTS}}^{\mathrm{QM_avg}} - E_{\mathrm{RTS}}^{\mathrm{QM}}\left[\phi_{\mathrm{eq}}^{\mathrm{training}}\right] = -\frac{\sqrt{w}}{\sqrt{\pi}} \sum_{n=1}^{\infty} c_n^{\mathrm{CO}} \cos\left[n\phi_{\mathrm{eq}}^{\mathrm{training}}\right] \tag{178}$$

Comparing eqn (174), (175), and (178), the difference in $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ for scenarios #1 and #2 using CO projectors when *sym*_value = 0 is

$$E_{\mathrm{RTS}}^{\mathrm{model_avg}}[\text{Scenario } \#2] - E_{\mathrm{RTS}}^{\mathrm{model_avg}}[\text{Scenario } \#1] = -\frac{\sqrt{w}}{\sqrt{\pi}} \sum_{n=(n_{\max}+1)}^{\infty} c_n^{\mathrm{CO}} \cos\left[n\phi_{\mathrm{eq}}^{\mathrm{training}}\right] \tag{179}$$

When *sym*_value ≠ 0, eqn (178) and (179) do not apply. If *sym*_value is large, then the CO projectors might give a large difference in $E_{\mathrm{RTS}}^{\mathrm{model_avg}}$ value between scenarios #1 and #2.

For a torsion scan curve, the *R*-squared value is computed *via* eqn (17) using the following definitions:

$$\mathrm{SST} = \sum_{j=1}^{T} \left(E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi_j] - E_{\mathrm{RTS}}^{\mathrm{QM_avg}}\right)^2 \tag{180}$$

$$\mathrm{SSE} = \sum_{j=1}^{T} \left(E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi_j] - E_{\mathrm{RTS}}^{\mathrm{QM}}\left[\phi_{\mathrm{eq}}^{\mathrm{training}}\right] - E_{\mathrm{RTS}}^{\mathrm{model}}[\phi_j]\right)^2 \tag{181}$$

As derived in ESI Section S12,† this *R*-squared value can be equivalently rewritten as

$$R\text{-squared} = R^2 = \mathrm{SumCSq} - \frac{T}{\mathrm{SST}}\left(E_{\mathrm{RTS}}^{\mathrm{QM_avg}} - E_{\mathrm{RTS}}^{\mathrm{QM}}\left[\phi_{\mathrm{eq}}^{\mathrm{training}}\right] - E_{\mathrm{RTS}}^{\mathrm{model_avg}}\right)^2 \tag{182}$$

Examining eqn (182),

$$R\text{-squared} \leq \mathrm{SumCSq} \tag{183}$$

For Scenario #2, combining eqn (175) and (182) shows that

$$R\text{-squared} = R^2 = \mathrm{SumCSq} \tag{184}$$

Eqn (184) holds in Scenario #2 irrespective of whether DT or CO projectors are used and irrespective of the value of *sym*_value, but of course the value of SumCSq (and hence of *R*-squared) is impacted by which projectors are used.

As derived in ESI Section S12,† for Scenario #1 eqn (182) is rewritten as

$$R\text{-squared} = R^2 = \mathrm{SumCSq} - 2\left(\sum_{m=(m_{\max}+1)}^{\infty} c_m^{\mathrm{DT}} G_m^{\mathrm{DT,avg}}\right)^2 \tag{185}$$

Thus when using Scenario #1, *R*-squared is close to SumCSq when c_m^{DT} for every omitted torsion mode is close to zero. Iff *sym*_value = 0, then an analogous derivation using the CO projectors for scenario #1 yields

$$R\text{-squared} = R^2 = \mathrm{SumCSq} - 2\left(\sum_{n=(n_{\max}+1)}^{\infty} c_n^{\mathrm{CO}} \cos\left[n\phi_{\mathrm{eq}}^{\mathrm{training}}\right]\right)^2 \tag{186}$$

When computing the *R*-squared value as described above, all *T* geometries along the torsion scan curve had equal observation weights. In classical molecular dynamics or Monte Carlo simulations employing the NPT, NVT, μPT, or μVT thermodynamic ensembles, the lower-energy geometries should appear more often (*i.e.*, have higher observation weights) than higher-energy geometries. When every geometry along the torsion scan curve has equal observation weights, scenario #2 described above has a higher *R*-squared value (see eqn (184)) than scenario #1 (see eqn (185)). Scenario #1 (see eqn (172)) exactly matches the relative energy of the QM and model potentials at the training dataset's optimized ground-state geometry. For this reason, scenario #1 typically performs better than scenario #2 when employing the NPT, NVT, μPT, or μVT thermodynamic ensembles, because scenario #1 gives smaller errors than scenario #2 for the lower-energy geometries that receive relatively higher observation weights in such ensembles. For this reason, my ADDT, CADT, ADCO, and CACO model potentials are typically constructed according to scenario #1 rather than according to scenario #2. Except where otherwise indicated, the ADDT, CADT, ADCO, and CACO model potentials were constructed according to scenario #1 throughout this work.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

In this work and a companion article,[19] we used the following smart selection thresholds. If *sym*_value ≤ 0.01, this means $U_{ABCD}^{torsion}[\phi] = U_{ABCD}^{torsion}[-\phi]$ within the tolerance, so the ADCO or CACO model potential was used. In this case, an ADCO or CACO mode was kept if abs[c_m] > 0.001. Keeping the ADCO or CACO coefficients greater than this 'very tight' cutoff helps ϕ_{eq}^{FF} more closely approach $\phi_{eq}^{training}$. If 0.01 < *sym*_value ≤ 0.1, the ADDT or CADT model potential was used, and an ADDT or CADT mode was kept if abs[c_m] > 0.01. This case corresponds to the situation in which $U_{ABCD}^{torsion}[\phi]$ is approximately but not strictly equal to $U_{ABCD}^{torsion}[-\phi]$, so it is beneficial to use a 'tight' cutoff (*i.e.*, abs[c_m] > 0.01) for retaining torsion modes to achieve a balance between accuracy and conciseness. This 'tight' cutoff helps the ADDT or CADT model potential to more accurately reproduce the position of the alternate local energy minimum $\phi_{eq}^{alternate} \approx -\phi_{eq}$. If 0.1 < *sym*_value, the ADDT or CADT model potential was used, and an ADDT or CADT mode was kept if abs[c_m] > 0.1. This case corresponds to the situation in which $U_{ABCD}^{torsion}[\phi]$ is not approximately equal to $U_{ABCD}^{torsion}[-\phi]$, so conciseness of the torsion modes is preferred. Examining eqn (168), this 'normal' cutoff (*i.e.*, *abs*[c_m] > 0.1) neglects a torsion mode if it affects the SumCSq value by ≤ 0.01.

The coefficients k_ϕ^m, c_m, and $\{b_\phi^n\}$ have the following distinct meanings. k_ϕ^m is the force constant for torsion mode m. $\sqrt{w}c_m$ is the projection coefficient of mode m for a torsion scan of one dihedral instance, where $\sqrt{w}$ is the torsion scan's normalization coefficient and c_m is the normalized function's projection onto torsion mode m. In contrast to the coefficients $\{c_m\}$ which refer to an individual torsion scan, the force constants $\{k_\phi^m\}$ apply to all motions of the system. Because of multicollinearity between some flexibility terms, there is no universal relationship between $\{\sqrt{w}c_m\}$ and $\{k_\phi^m\}$. As shown above, ADDT and CADT torsion mode 5 is composed of two sine functions, torsion mode 6 is composed of two sine functions, and torsion mode 7 is composed of four sine functions. The coefficients $\{b_\phi^n\}$ are the coefficients for individual sine functions, and these $\{b_\phi^n\}$ are computed from the force constants $\{k_\phi^m\}$ of ADDT or CADT torsion modes m = 5, 6, and 7 using the formulas shown in eqn (141)–(144).

8. How do the CADT and CACO model potentials compare to previously published class A torsion potentials?

The Class A (*i.e.*, 'dihedral-only') torsion potential of dihedral ϕ_{ABCD} can be expanded as a Fourier series expressed in either of two equivalent forms:

$$U_{ABCD}^{dihedral\text{-}only}[\phi_{ABCD}] = D_0 + \sum_{n=1}^{n_{max}} D_n \cos\left[n\phi_{ABCD} - \psi_n^{ABCD}\right] \quad (187)$$

$$U_{ABCD}^{dihedral\text{-}only}[\phi_{ABCD}] = A_0 + \sum_{n=1}^{n_{max}} \left(A_n \cos[n\phi_{ABCD}] + B_n \sin[n\phi_{ABCD}]\right) \quad (188)$$

where D_0, D_n (for n = 1, 2, 3, …), and ψ_n^{ABCD} are adjustable (aka 'fitted') parameters that define the potential. Here, n_{max} is the maximum value of n that is considered in the model. This Fourier series becomes formally complete in the limit $n_{max} \to \infty$. Comparing eqn (26), (187), and (188) confirms that

$$A_0 = D_0 \quad (189)$$

$$A_{n\geq 1} = D_n \cos[\psi_n^{ABCD}] \quad (190)$$

$$B_n = D_n \sin[\psi_n^{ABCD}] \quad (191)$$

If neither contained equilibrium bond angle is linear (*i.e.*, $\theta_{ABC}^{eq} \neq \pi$ and $\theta_{BCD}^{eq} \neq \pi$), then it is possible to define the equilibrium dihedral value (ϕ_{ABCD}^{eq}), which is the value of ϕ_{ABCD} in the material's optimized ground-state (*i.e.*, low-energy) geometry. In this case, the Fourier series can be re-written as

$$U_{ABCD}^{dihedral\text{-}only}[\phi_{ABCD}] = \alpha_0 + \sum_{n=1}^{n_{max}} \Big(-a_\phi^n \cos[n(\phi_{ABCD} - \phi_{ABCD}^{eq})] + b_\phi^n S_{instance} \sin[n(\phi_{ABCD} - \phi_{ABCD}^{eq})] \Big) \quad (192)$$

which is equivalent to eqn (187) and (188). Comparing eqn (26), (187), and (192) confirms that

$$\alpha_0 = D_0 \quad (193)$$

$$a_\phi^n = -D_n \cos[\psi_n^{ABCD} - n\phi_{ABCD}^{eq}] \quad (194)$$

$$b_\phi^n S_{instance} = D_n \sin[\psi_n^{ABCD} - n\phi_{ABCD}^{eq}] \quad (195)$$

Defining the variable substitutions

$$n_{max} = 4 \quad (196)$$

$$k_\phi^n = a_\phi^n \text{ for } n = 1,2,3,4 \quad (197)$$

$$\text{Ћ} = \alpha_0 \quad (198)$$

and substituting these into eqn (192) yields the Fourier series expansion shown in eqn (139). Applying the boundary condition

$$\lim_{\phi_{ABCD} \to \phi_{ABCD}^{eq}} U_{ABCD}^{dihedral\text{-}only}[\phi_{ABCD}] = U_{ABCD}^{dihedral\text{-}only}[\phi_{ABCD}^{eq}] \quad (199)$$

yields eqn (140). The boundary condition

$$\left.\frac{dU_{ABCD}^{dihedral\text{-}only}[\phi_{ABCD}]}{d\phi_{ABCD}}\right|_{\phi_{ABCD}=\phi_{ABCD}^{eq}} = 0 \quad (200)$$

applies, because all of the atom-in-material forces are zero at the material's optimized ground-state geometry. This boundary condition is automatically satisfied for the constant Ћ and the cosine modes in eqn (139). Applying the boundary condition of eqn (200) to the Fourier series shown in eqn (139) yields eqn (119). The constraint shown in eqn (119) has the effect of reducing the number of independent sine modes from n_{max} to (n_{max} − 1). Consider n_{max} = 4 as an example. In this case, there are n_{max} = 4 independent cosine modes and (n_{max} − 1) = 3 independent sine modes, which totals to 7 seven rotatable dihedral torsion modes.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

From this, we conclude that the CADT potential described in Section 5 completely spans the degrees of freedom associated with the Fourier series expansion, which can be equivalently written in any of the following forms:

(i) Eqn (137) and (138)

(ii) Eqn (139)–(144), which is the same as eqn (192) with constraint (200),

(iii) Eqn (187) with constraint (200)

(iv) Eqn (188) with constraint (200)

The Fourier expansion form shown in eqn (187) is extremely inconvenient, because it requires non-linear regression to optimize the ψ_n^{ABCD} parameters.[25] Rewriting the Fourier expansion as shown in eqn (137), (138) or eqn (188) or eqn (192) avoids this problem.

Which parts of this Fourier series expansion were not captured by specific Class A torsion potentials used in the prior literature? Some prior literature[7,26,27] used an expansion of the form

$$U_{\mathrm{ABCD}}^{\mathrm{cos\text{-}pwrs}}[\phi_{\mathrm{ABCD}}] = \Lambda_0 + \Lambda_1 \cos[\phi_{\mathrm{ABCD}}] + \Lambda_2 \cos^2[\phi_{\mathrm{ABCD}}] + \Lambda_3 \cos^3[\phi_{\mathrm{ABCD}}] + \ldots \quad (201)$$

which is an even function of ϕ_{ABCD}. An expansion of powers of cosine is equivalent to an expansion involving various cosine multiplicities:

$$U_{\mathrm{ABCD}}^{\mathrm{even}}[\phi_{\mathrm{ABCD}}] = \Gamma_0 + \Gamma_1 \cos[\phi_{\mathrm{ABCD}}] + \Gamma_2 \cos[2\phi_{\mathrm{ABCD}}] + \Gamma_3 \cos[3\phi_{\mathrm{ABCD}}] + \ldots \quad (202)$$

which is functionally equivalent to the torsion potential form used in some common implementations of the OPLS-AA,[28] AMBER,[8,29] CHARMM[30] and some other[5] forcefields. For example, the OPLS-AA forcefield uses the following 'cosine only' torsion potential:[28]

$$U_{\mathrm{torsion}}^{\text{OPLS-AA}} = \sum_i \left(\frac{V_{i,1}}{2}(1+\cos[\phi_i]) + \frac{V_{i,2}}{2}(1-\cos[2\phi_i]) + \frac{V_{i,3}}{2}(1+\cos[3\phi_i]) + \frac{V_{i,4}}{2}(1-\cos[4\phi_i]) \right) \quad (203)$$

These 'cosine only' torsion model potentials are manifestly incomplete, because they omit all of the odd-function contributions to the torsion potential.

Burger *et al.*[6] used a dihedral potential of the form

$$U_{\mathrm{dihedrals}} = \frac{1}{2} \sum_{i=1}^{\mathrm{dihedrals}} \sum_{n=1}^{N_i} V_{i,n}(1+\cos[n\phi_i - \psi_i]) \quad (204)$$

This type of model potential is incomplete and overly restrictive, because it uses the same ψ_i value for all N_i modes of dihedral ϕ_i. Hopkins and Roitberg[25] used a dihedral model potential of the form:

$$U_{\mathrm{dihedrals}} = \sum_{i=1}^{\mathrm{dihedrals}} \sum_{n=1}^{n_{\max}} k_{i,n}\left(1+\cos\left[n\phi_i - \psi_{i,n}\right]\right) \quad (205)$$

They also rewrote this in the form of eqn (188) except that they omitted the constant (A_0) term. Comparing eqn (187) to (205) shows that eqn (205) provides a nearly complete Fourier series expansion of the torsion potential, except that eqn (205) is missing one degree of freedom in the constant part of the potential. This does not affect the forces.[25] Radom *et al.*[3] investigated a torsion potential that was a truncation of the Fourier series shown in eqn (188).

Some prior literature appears at first glance to use dihedral model potentials containing $\cos[n\phi_i - \psi_{i,n}]$ terms. However, some of these forcefields impose the too-severe restriction that $\psi_{i,n} \in \{0,180^\circ\}$, which means that $\cos[n\phi_i - \psi_{i,n}] \rightarrow \pm\cos[n\phi_i]$.[5] In this case, the expansion is actually equivalent to the form shown in eqn (202), which is not a Fourier series expansion because it omits the odd-function contributions.

Some forcefields recommend using $\psi_{i,n} \in \{0,180^\circ\}$ even if they do not strictly impose it. "In CHARMM it is possible to use any value for the phase;[31] however, it is strongly suggested that values of 0 and 180° be used as the parameters are then appropriate for different stereoisomers associated with a given dihedral."[32] However, this too-severe restriction often prevents the torsion model from accurately reproducing the torsion scan curve. My ADDT and CADT model potentials completely avoid this problem, because they allow different stereoisomers to be described by the same ADDT and CADT force constant values without omitting the odd-function contributions.

The CACO torsion model potential is functionally equivalent to eqn (202) with the following caveats. A least-squares fit of eqn (202) to a torsion scan curve optimizes the constant potential intercept Γ_0 to match the average dihedral potential over the dihedral scan curve. The CACO torsion model potential adjusts the constant offset such that the model potential is zero when $\phi = \phi_{\mathrm{eq}}^{\mathrm{training}}$:

$$U_{\mathrm{ABCD}}^{\mathrm{CACO}}[\phi_{\mathrm{eq}}^{\mathrm{training}}] = 0 \quad (206)$$

Functionally, the CACO torsion model potential can be derived from eqn (202) as

$$U_{\mathrm{ABCD}}^{\mathrm{CACO}}[\phi] = U_{\mathrm{ABCD}}^{\mathrm{even}}[\phi] - U_{\mathrm{ABCD}}^{\mathrm{even}}[\phi_{\mathrm{eq}}^{\mathrm{training}}] \quad (207)$$

with the additional caveat that the force constant for each CACO mode is expressed as $k_{\mathrm{CACO}}c_n$. These particular choices ensure that CACO reproduces relative energies more accurately for the lower energy geometries (which appear more frequently during classical molecular dynamics and Monte Carlo simulations) and that $\phi_{\mathrm{eq}}^{\mathrm{FF}} \approx \phi_{\mathrm{eq}}^{\mathrm{training}}$.

9. Angle-damped linear dihedral (ADLD) model potential

9.1 Model potential for dihedrals containing at least one linear equilibrium bond angle

A dihedral containing one or two linear (or close to linear) equilibrium bond angles is called a 'linear dihedral'. In this context, 'close to linear' means $\pi - \theta_{\mathrm{ABC}}^{\mathrm{eq}} < \varepsilon$ or $\pi - \theta_{\mathrm{BCD}}^{\mathrm{eq}} < \varepsilon$, where ε is a tolerance (*e.g.*, $\varepsilon = 0.03$ radians was used in this work and the companion article[18,19]). Please note that the classification of whether or not a dihedral is a 'linear dihedral' is based on the values of $\theta_{\mathrm{ABC}}^{\mathrm{eq}}$ and $\theta_{\mathrm{BCD}}^{\mathrm{eq}}$ not the values of θ_{ABC} and θ_{BCD}. For a linear dihedral, $\phi_{\mathrm{ABCD}}^{\mathrm{eq}}$ is undefined and $f_{n\geq 1_\mathrm{eq}}^{\chi} = 0$ for whichever equilibrium bond angle is linear.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

The new ADLD model potential is derived starting from my general Class B torsion potential shown in eqn (87). The combined angle-dihedral coordinate branch equivalency condition applied to $H_n[\theta_{ABC},\theta_{BCD}](A_n\cos[n\phi_{ABCD}] + B_n\sin[n\phi_{ABCD}])$ requires that eqn (88) be satisfied. This can be accomplished by choosing the angle-damping functions multiplying cosine and sine terms in the ADLD potential to be

$$Q_n[\theta_{ABC},\theta_{BCD}] = f^{ABC}_{\lceil n/2\rceil} f^{ABC}_{\lceil n/2\rceil} f^{BCD}_{\lceil n/2\rceil} f^{BCD}_{\lceil n/2\rceil} \tag{208}$$

Note that these formulas involve no division by the equilibrium values. Here, $\lceil x\rceil$ denotes ceiling[x], which is the smallest integer greater than or equal to x. As explained in a previous section, $\lfloor x\rfloor$ denotes floor[x], which is the largest integer less than or equal to x.

For the ADLD model potential, an angle-dependent offset potential is required to ensure that the energy increases as the structure is displaced from its ground-state geometry. For an even mode, one could envision something like

$$Q_{2j}[\theta_{ABC},\theta_{BCD}]\left(\sqrt{(a_{2j})^2+(b_{2j})^2} + a_{2j}\cos[2j\phi_{ABCD}] + b_{2j}S_{instance}\sin[2j\phi_{ABCD}]\right) \geq 0 \tag{209}$$

to ensure that the energy contribution is always non-negative. In eqn (209), the term that does not depend on ϕ_{ABCD} is the angle-dependent offset potential.

Because they are odd functions of ϕ_{ABCD}, the $\sin[n\phi_{ABCD}]$ terms involve the sign $S_{instance} \in \{0,-1,+1\}$. Different dihedral instances of the same dihedral type can have different values of $S_{instance}$. Because ϕ_{eq} is undefined for linear dihedrals, the value of $S_{instance}$ cannot be assigned using the sign of ϕ_{eq} for linear dihedrals. For linear dihedrals, the value of $S_{instance}$ must be determined by detecting local mirror-image environments (*i.e.*, local chiral enantiomer environments) for different dihedral instances of the same dihedral type. Note that $S_{instance} = 0$ implies a mirror-image (*i.e.*, reflection) symmetry of the torsion potential.

Eqn (209) is not convenient, because it requires nonlinear regression to optimize the a_{2j} and b_{2j} values. To facilitate linear regression, it is convenient to instead expand each even mode as

$$U^{ADLD}_{2j} = (f^{ABC}_j f^{BCD}_j)^2(k^j_{LD1}(1 - \cos[2j\phi_{ABCD}]) + k^j_{LD2}(1 + \cos[2j\phi_{ABCD}]) + k^j_{LD3}S_{instance}\sin[2j\phi_{ABCD}]) \tag{210}$$

The coefficients k^j_{LD1}, k^j_{LD2}, and k^j_{LD3} in eqn (210) can be optimized using linear regression, and the constraints

$$k^j_{LD1} \geq 0 \tag{211}$$

$$k^j_{LD2} \geq 0 \tag{212}$$

should be imposed during this linear regression. Eqn (209) can always be represented as a special case of eqn (210) by setting

$$k^j_{LD3} = b_{2j} \tag{213}$$

$$k^j_{LD1} = \frac{1}{2}\left(+\sqrt{(a_{2j})^2+(b_{2j})^2} - a_{2j}\right) \geq 0 \tag{214}$$

$$k^j_{LD2} = \frac{1}{2}\left(+\sqrt{(a_{2j})^2+(b_{2j})^2} + a_{2j}\right) \geq 0 \tag{215}$$

For odd modes, $Q_{2j-1}[\theta_{ABC},\theta_{BCD}]$ multiplied by only a constant does not satisfy the combined angle-dihedral coordinate branch equivalency condition. To ensure that the energy increases as the structure is displaced from its ground-state geometry, instead of $Q_{2j-1}[\theta_{ABC},\theta_{BCD}]$ we can use an angle-dependent offset potential of the form

$$\frac{1}{2}\left(\left(f^{ABC}_j f^{BCD}_{j-1}\right)^2 + \left(f^{ABC}_{j-1} f^{BCD}_j\right)^2\right) \tag{216}$$

which does satisfy the combined angle-dihedral coordinate branch equivalency condition. The offset in eqn (216) arises from the following 'completing the squares':

$$\begin{aligned}&\frac{1}{2}\left(f^{ABC}_j f^{BCD}_{j-1} - f^{ABC}_{j-1} f^{BCD}_j\right)^2 \\ &\quad + Q_{2j-1}[\theta_{ABC},\theta_{BCD}](1 \pm \cos[(2j-1)\phi_{ABCD}]) \\ &\quad = \underbrace{\frac{1}{2}\left(\left(f^{ABC}_j f^{BCD}_{j-1}\right)^2 + \left(f^{ABC}_{j-1} f^{BCD}_j\right)^2\right)}_{\text{offset}} \\ &\quad \pm f^{ABC}_j f^{ABC}_{j-1} f^{BCD}_j f^{BCD}_{j-1}\cos[(2j-1)\phi_{ABCD}]\end{aligned} \tag{217}$$

Because

$$0 \leq |\cos[(2j-1)\phi_{ABCD}]| \leq 1 \tag{218}$$

it directly follows that

$$\begin{aligned}&\frac{1}{2}\left(\left(f^{ABC}_j f^{BCD}_{j-1}\right)^2 + \left(f^{ABC}_{j-1} f^{BCD}_j\right)^2\right) \\ &\quad - f^{ABC}_j f^{ABC}_{j-1} f^{BCD}_j f^{BCD}_{j-1}|\cos[(2j-1)\phi_{ABCD}]| \\ &\quad \geq \left(\frac{1}{2}\left(\left(f^{ABC}_j f^{BCD}_{j-1}\right)^2 + \left(f^{ABC}_{j-1} f^{BCD}_j\right)^2\right) - f^{ABC}_j f^{ABC}_{j-1} f^{BCD}_j f^{BCD}_{j-1}\right) \\ &\quad = \frac{1}{2}\left(f^{ABC}_j f^{BCD}_{j-1} - f^{ABC}_{j-1} f^{BCD}_j\right)^2 \geq 0\end{aligned} \tag{219}$$

Somewhat analogous to the even modes, each odd mode is expanded using three force constants amenable to linear regression:

$$\begin{aligned}U^{ADLD}_{2j-1} = &\frac{1}{2}\left(\left(f^{ABC}_j\right)^2\left(f^{BCD}_{j-1}\right)^2 + \left(f^{ABC}_{j-1}\right)^2\left(f^{BCD}_j\right)^2\right)(k^j_{LD4} + k^j_{LD5}) \\ &+ f^{ABC}_j f^{ABC}_{j-1} f^{BCD}_j f^{BCD}_{j-1}\left(\left(-k^j_{LD4} + k^j_{LD5}\right)\cos[(2j-1)\phi_{ABCD}] + k^j_{LD6}S_{instance}\sin[(2j-1)\phi_{ABCD}]\right)\end{aligned} \tag{220}$$

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

To ensure that the k^j_{LD4} and k^j_{LD5} terms never decrease in energy as the geometry undergoes an infinitesimal displacement away from the optimized ground-state geometry, combining eqn (219) and (220) shows the following constraints should be enforced during linear regression

$$k^j_{\mathrm{LD4}} \geq 0 \tag{221}$$

$$k^j_{\mathrm{LD5}} \geq 0 \tag{222}$$

Each $\sin[2j\phi_{\mathrm{ABCD}}]$ and $\sin[(2j-1)\phi_{\mathrm{ABCD}}]$ term decreases in energy as the dihedral undergoes an infinitesimal displacement away from the optimized ground-state geometry in one direction and increases in energy as the dihedral is displaced away from the optimized ground-state geometry in the opposite direction. Because the optimized ground-state geometry is an energy minimum, this means the $\sin[2j\phi_{\mathrm{ABCD}}]$ and $\sin[(2j-1)\phi_{\mathrm{ABCD}}]$ terms cannot appear by themselves in the optimized forcefield, but rather they can only occur if a force constant for one or more $\cos[2j\phi_{\mathrm{ABCD}}]$ or $\cos[(2j-1)\phi_{\mathrm{ABCD}}]$ modes is non-zero. In this case, the force constant(s) for the cosine mode(s) must be large enough to ensure that any active sine modes cannot lower the isolated bonded cluster's energy below that of its optimized ground-state geometry. Instead of enforcing this as an explicit constraint on the force constants, we rely on the training dataset fitting to obtain (asymptotically close to) this behavior during linear regression that optimizes the values of force constants.

Putting this altogether gives

$$\begin{aligned}
U^{\mathrm{ADLD}}_{\mathrm{ABCD}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] &= \sum_{j=1}^{\infty}\left(U^{\mathrm{ADLD}}_{2j} + U^{\mathrm{ADLD}}_{2j-1}\right)\\
&= \sum_{j=1}^{\infty}\left(f_j^{\mathrm{ABC}}\right)^2\left(f_j^{\mathrm{BCD}}\right)^2\left(k^j_{\mathrm{LD1}}(1-\cos[2j\phi_{\mathrm{ABCD}}])\right.\\
&\quad + k^j_{\mathrm{LD2}}(1+\cos[2j\phi_{\mathrm{ABCD}}]) + k^j_{\mathrm{LD3}}S_{\mathrm{instance}}\sin[2j\phi_{\mathrm{ABCD}}])\\
&\quad + \sum_{j=1}^{\infty} f_j^{\mathrm{ABC}} f_{j-1}^{\mathrm{ABC}} f_j^{\mathrm{BCD}} f_{j-1}^{\mathrm{BCD}}\left(\left(-k^j_{\mathrm{LD4}} + k^j_{\mathrm{LD5}}\right)\cos[(2j-1)\phi_{\mathrm{ABCD}}]\right.\\
&\quad \left. + k^j_{\mathrm{LD6}}S_{\mathrm{instance}}\sin[(2j-1)\phi_{\mathrm{ABCD}}]\right) + \frac{1}{2}\sum_{j=1}^{\infty}\left(\left(f_j^{\mathrm{ABC}}\right)^2\left(f_{j-1}^{\mathrm{BCD}}\right)^2\right.\\
&\quad \left.+\left(f_{j-1}^{\mathrm{ABC}}\right)^2\left(f_j^{\mathrm{BCD}}\right)^2\right)\left(k^j_{\mathrm{LD4}} + k^j_{\mathrm{LD5}}\right)
\end{aligned} \tag{223}$$

subject to the constraints

$$k^j_{\mathrm{LD1}}, k^j_{\mathrm{LD2}}, k^j_{\mathrm{LD4}}, k^j_{\mathrm{LD5}} \geq 0 \tag{224}$$

This can also be re-written in the equivalent form

$$\begin{aligned}
U^{\mathrm{ADLD}}_{\mathrm{ABCD}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] &= \sum_{n=1}^{\infty} U_n^{\mathrm{ADLD}}\\
&= \sum_{n=1}^{\infty} Q_n\left(\left(k_n^{+\cos} - k_n^{-\cos}\right)\cos[n\phi_{\mathrm{ABCD}}] + k_n^{\sin}S_{\mathrm{instance}}\sin[n\phi_{\mathrm{ABCD}}]\right)\\
&\quad + \frac{1}{2}\sum_{n=1}^{\infty}\left(\left(f^{\mathrm{ABC}}_{\lceil n/2\rceil}\right)^2\left(f^{\mathrm{BCD}}_{\lfloor n/2\rfloor}\right)^2 + \left(f^{\mathrm{ABC}}_{\lfloor n/2\rfloor}\right)^2\left(f^{\mathrm{BCD}}_{\lceil n/2\rceil}\right)^2\right)\left(k_n^{+\cos} + k_n^{-\cos}\right)
\end{aligned} \tag{225}$$

subject to the constraints

$$k_n^{+\cos}, k_n^{-\cos} \geq 0 \tag{226}$$

9.2 Selecting specific modes from the linear dihedral torsion potential

A particular linear dihedral normally involves only a small subset of the terms shown in eqn (223). The particular terms involved depend on the type of linear dihedral. A single-linear dihedral contains exactly one linear (or close to linear) equilibrium bond angle: $(\pi - \theta^{\mathrm{eq}}_{\mathrm{ABC}} < \varepsilon)$ xor $(\pi - \theta^{\mathrm{eq}}_{\mathrm{BCD}} < \varepsilon)$. A double-linear dihedral contains two linear (or close to linear) equilibrium bond angles: $(\pi - \theta^{\mathrm{eq}}_{\mathrm{ABC}} < \varepsilon)$ and $(\pi - \theta^{\mathrm{eq}}_{\mathrm{BCD}} < \varepsilon)$. Here, we consider four cases.

Case # 1 is a symmetry-induced single-linear dihedral. Case # 1 corresponds to the situation in which the linear equilibrium bond angle is a ω-fold axis of rotational symmetry for some whole number $1 < \omega < \infty$. This symmetry eliminates all of the terms in eqn (223) except those for which n is wholly divisible by ω. For example, if $\omega = 2$ then only the n = even (*i.e.*, the $n = 2j$ terms) survive and all of the odd terms (*i.e.*, the $n = (2j-1)$ terms) are eliminated. If $n = 3$, then only the terms for which n (which is $2j$ or $(2j-1)$) is wholly divisible by 3 (*i.e.*, $n = 3, 6, 9, \ldots$) survive. Depending on the situation, the bent equilibrium bond angle may be contained in a mirror plane (Case 1a) or not (Case 1b). If the bent equilibrium bond angle is contained within a mirror plane (Case 1a), then this mirror symmetry eliminates all of the odd functions (*i.e.*, the sine terms) in eqn (223). Examples of Case 1a include boranecarbonitrile (H_2B–C≡N) and isocyanoborane (H_2B–N≡C) which have a 2-fold axis of rotational symmetry and acetonitrile (H_3C–C≡N) which has a 3-fold axis of rotational symmetry. (Since $n = 1$ is not wholly divisible by any number $\omega > 1$, it follows that the $n = (2(j=1)-1) = 1$ terms are always eliminated in Cases 1a and 1b.)

Case # 2 is a symmetry-induced double-linear dihedral. Case # 2 corresponds to the situation in which the linear ABCD equilibrium axis (along which the two linear equilibrium bond angles ABC and BCD reside) is either contained within two mirror planes (Case 2a) or a ω-fold axis of rotational symmetry (Case 2b). Either of these two symmetries (*i.e.*, Case 2a or Case 2b) ensures that the atom-in-material forces are zero on atoms A, B, C, and D when they are arranged along this linear axis. Case 2a: If two mirror planes contain the linear ABCD equilibrium axis, then this mirror symmetry eliminates all of the odd functions (*i.e.*, the sine terms) in eqn (223). Examples of this case include the acetylene (H–C≡C–H) and fulminic acid (H–C≡N–O) molecules. In this case, the torsion potential expands as

$$\begin{aligned}
&U^{\mathrm{case_2}}_{\mathrm{linear_dihedral}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}]\\
&\quad = \left(\frac{1}{2}\left(k^1_{\mathrm{LD4}} + k^1_{\mathrm{LD5}}\right)\left(\left(f_1^{\mathrm{ABC}}\right)^2 + \left(f_1^{\mathrm{BCD}}\right)^2\right)\right.\\
&\quad\quad \left. + f_1^{\mathrm{ABC}} f_1^{\mathrm{BCD}}\left(-k^1_{\mathrm{LD4}} + k^1_{\mathrm{LD5}}\right)\cos[\phi_{\mathrm{ABCD}}]\right) + \text{h.o.t.}
\end{aligned} \tag{227}$$

where h.o.t. stands for 'higher-order terms'. In practice, only the leading-order terms explicitly written in eqn (227), which are the $n = (2(j=1)-1) = 1$ terms, are needed if the non-equilibrium

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

bond angle fluctuations are modest in amplitude. This simplification arises because for small $\not{K}_{ABC}$ the angle-damping factor f_j^{ABC} is proportional to $(\not{K}_{ABC})^j$. In Case 2b, none of the terms in eqn (223) are necessarily automatically eliminated by symmetry; however, depending on the circumstances some of the terms in eqn (223) may turn out not to be significant. The C–C–N–O double-linear dihedral in the acetonitrile N-oxide (H_3C–C≡N–O) molecule is an example of Case 2b.

Case # 3 is an accidental single-linear dihedral. In this context, the term 'accidental' means that the single-linear dihedral arises due to a close to linear equilibrium bond angle that is not caused by any intrinsic symmetry. Depending on one's perspective, this case could be regarded as either 'rare' or 'not rare'. Specifically, when examining a single material, it is comparatively 'rare' that the balancing of forces would fortuitously lead to an after-dihedral-pruning close to linear equilibrium bond angle not caused by any intrinsic symmetry. (Before dihedral pruning, close to linear equilibrium bond angles would occur frequently. The vast majority of these would be removed from the active dihedrals list by dihedral pruning.) However, when examining a large database of materials, the sheer large number of materials investigated makes it probable that a few of these materials will contain an accidental single-linear dihedral after dihedral pruning. Since the force for small (*i.e.*, infinitesimal) displacements of the linear bond angle must be zero, it follows (see ESI† Section S8) that the $n = (2(j = 1) - 1) = 1$ terms are eliminated in eqn (223), while higher-order terms can potentially contribute to the torsion potential.

Case # 4 is an accidental double-linear dihedral. This case corresponds to the situation in which a double-linear dihedral arises due to a balance of forces that causes a linear ABCD moiety not caused by intrinsic symmetry. In this case, none of the terms in eqn (223) are necessarily automatically eliminated by symmetry; however, depending on the circumstances some of the terms in eqn (223) may turn out not to be significant. If considering relatively local symmetries of the chemical group rather than global symmetries of the entire chemical system, this case would be rare after dihedral pruning.

Depending on one's perspective, an acetylene molecule adsorbed in a MOF could be considered either as Case #2 or Case #4 if it has a double-linear dihedral. If one considers an acetylene molecule's symmetry without including the MOF adsorbent, this double-linear dihedral would be classified as Case #2. If one considers the global symmetry of the MOF plus adsorbed acetylene molecule, the lower symmetry of this combined system could result in classifying this double-linear dihedral as Case #4. Ultimately, this demonstrates that the classification into various cases is a judgement call that depends on how large a chemical subunit is chosen when evaluating the symmetry properties. A reasonable compromise would be to consider a local region that extends ~3 bonds in all directions around the edges of the dihedral being considered.

10. Results and discussion

10.1 CCSD quantum chemistry calculations of small molecules

All quantum chemistry calculations in this section were performed in Gaussian16 (ref. 33) using the CCSD method with def2-TZVPD[34] basis sets. (In this article, CCSD not CCSD(T) calculations were used.) For molecules containing no elements heavier than neon, all electrons were correlated in the coupled-cluster calculation. For molecules containing one or more elements heavier than neon, the FreezeNobleGasCore keyword was used, which applies the coupled-cluster correlation to the valence shell electrons only on all atoms. Geometries were optimized to the following convergence criteria: (1) the maximum force is less than 0.00045 hartrees/bohr; (2) the root-mean squared (RMS) force is less than 0.0003 hartrees/bohr; (3) the maximum displacement is less than 0.0018 bohr; and (4) the RMS displacement is less than 0.0012 bohr. As described below, some of the calculations constrained one or more internal coordinates to generate energy scans.

Values of calculated geometric parameters in the CCSD/def2-TZVPD optimized ground-state structures of 13 molecules containing no linear bond angles are shown in Table 1 and of five molecules containing one or more linear bond angles are shown in Table 2. For each molecule, optimized bond lengths and angles are listed for the particular dihedral that was subsequently studied in detail as described below. Fig. 7 shows the optimized geometries.

Table 1 Key values of calculated geometric parameters in the CCSD/def2-TZVPD optimized ground-state structure of each molecule. These dihedrals contain two nonlinear equilibrium bond angles

Molecule	Atoms in dihedral	ϕ_{eq} (°)	Bond 1 (Å)	Bond 2 (Å)	Bond 3 (Å)	Angle 1 (°)	Angle 2 (°)
$(CClFH)_2$	HCCH	180.0	1.090 (HC)	1.524 (CC)	1.090 (CH)	111.2 (HCC)	111.2 (CCH)
C(OH)ClFH	FCOH	−64.70	1.358 (FC)	1.355 (CO)	0.967 (OH)	111.1 (FCO)	109.0 (COH)
Ethane	HCCH	180.0	1.093 (HC)	1.525 (CC)	1.093 (CH)	111.3 (HCC)	111.3 (CCH)
FSSF	FSSF	87.42	1.625 (FS)	1.913 (SS)	1.625 (SF)	107.2 (FSS)	107.2 (SSF)
Glyoxal	OCCO	180.0	1.201 (OC)	1.520 (CC)	1.201 (CO)	121.3 (OCC)	121.3 (CCO)
H_2O_2	HOOH	111.1	0.967 (HO)	1.438 (OO)	0.967 (OH)	100.8 (HOO)	100.8 (OOH)
HNCO	HNCO	180.0	1.006 (HN)	1.211 (NC)	1.160 (CO)	123.6 (HNC)	173.0 (NCO)
HNCS	HNCS	180.0	1.005 (HN)	1.199 (NC)	1.572 (CS)	131.7 (HNC)	174.3 (NCS)
HONC	HONC	180.0	0.969 (HO)	1.327 (ON)	1.168 (NC)	105.1 (HON)	173.3 (ONC)
HSNC	HSNC	180.0	1.343 (HS)	1.665 (SN)	1.175 (NC)	95.4 (HSN)	173.8 (SNC)
IF_3ClOH	ClIOH	180.0	2.343 (ClI)	1.861 (IO)	0.972 (OH)	83.5 (ClIO)	107.1 (IOH)
N_2O_2	ONNO	0.00	1.151 (ON)	1.872 (NN)	1.151 (NO)	103.4 (ONN)	103.4 (NNO)
PF_4OH	FPOH	180.0	1.543 (FP)	1.617 (PO)	0.963 (OH)	88.0 (FPO)	113.1 (POH)

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 2 Key values of calculated geometric parameters in the CCSD/def2-TZVPD optimized ground-state structure of each molecule. These dihedrals contain one or two linear equilibrium bond angles

Molecule	Atoms in dihedral	Bond 1 (Å)	Bond 2 (Å)	Bond 3 (Å)	Angle 1 (°)	Angle 2 (°)
Acetonitrile	HCCN	1.090 (HC)	1.462 (CC)	1.152 (CN)	109.9 (HCC)	180.0 (CCN)
Acetylene	HCCH	1.064 (HC)	1.202 (CC)	1.064 (CH)	180.0 (HCC)	180.0 (CCH)
H_2BCN	HBCN	1.185 (HB)	1.533 (BC)	1.156 (CN)	118.8 (HBC)	180.0 (BCN)
H_2BNC	HBNC	1.186 (HB)	1.432 (BN)	1.175 (NC)	118.3 (HBN)	180.0 (BNC)
HCNO	HCNO	1.062 (HC)	1.152 (CN)	1.202 (NO)	180.0 (HCN)	180.0 (CNO)

For the ONNO molecule, CCSD/def2-TZVPD geometry optimization was performed for both the *cis* and *trans* conformations in both the singlet and triplet spin states. The computed relative energies (in eV) were 0.00 (*cis*, singlet), 0.11 (*trans*, singlet), 0.48 (*cis*, triplet) and 0.53 (*trans*, triplet). East previously studied the electronic states of this molecule in cis and trans geometries with complete active space self-consistent field (CAS-SCF), multireference configuration (MRCISD), and other electronic structure methods.[35] Those computations indicated the cis geometry with singlet spin state to be the ONNO molecule's ground state.[35]

10.2 Detailed torsion analysis for rotatable dihedrals using ADDT and CADT model potentials

Table 3 summarizes torsion mode analysis for these 13 molecules containing nonlinear dihedrals. Both rigid and (fully or partly) relaxed torsion scans are presented. For HNCO, HNCS, HONC, and HSNC, fully relaxed torsion scans could not be computed, because due to 'slip torsion' (see Section 10.4) some constrained dihedral values yielded a relaxed bond angle that was too close to linear for the optimizer to converge. (The dihedral value becomes indeterminate when the bond angle becomes linear.) For each of these four molecules, a partly relaxed torsion scan was thus performed holding the two bond angles rigid but allowing the bond lengths to relax as the constrained dihedral was scanned. As shown in Table 3, these partly relaxed torsion scans yielded results virtually equal to the rigid torsion scans. Fully relaxed dihedral scans were performed for the other nine molecules. In Table 3, the columns labeled 'mode 1', 'mode 2', *etc.* display the coefficients (*i.e.*, c_i values) for these individual modes. As demonstrated by the SumCSq $\approx$ 1.00 values in Table 3, my new seven-mode dihedral torsion model potential yielded superb fits for all of these molecules.

Fig. 8 plots the quantum chemistry results and model fits (including all seven torsion modes) for the rigid and fully relaxed torsion scans. In some cases (*e.g.*, ethane, glyoxal, and H_2O_2), the relaxed torsion scan gave nearly the same results as the rigid torsion scan. In other cases (*e.g.*, $(CClFH)_2$, IF_3ClOH, N_2O_2, and PF_4OH), there was a huge difference between the relaxed and rigid torsion scans. FSSF and C(OH)ClFH showed modest non-negligible differences between relaxed and rigid torsion scans.

Angle-scan energy curves were QM-computed for the 10 molecules from this set that contained at least one atom with a coordination number equal to 2. Fig. 9 compares these QM-computed results to a model potential. The model potential was the sum of the ADDT torsion offset potential (computed using the rigid torsion parameters from Table 3) and the following angle-bending model potential[20] with the k_{angle} value displayed on each graph:

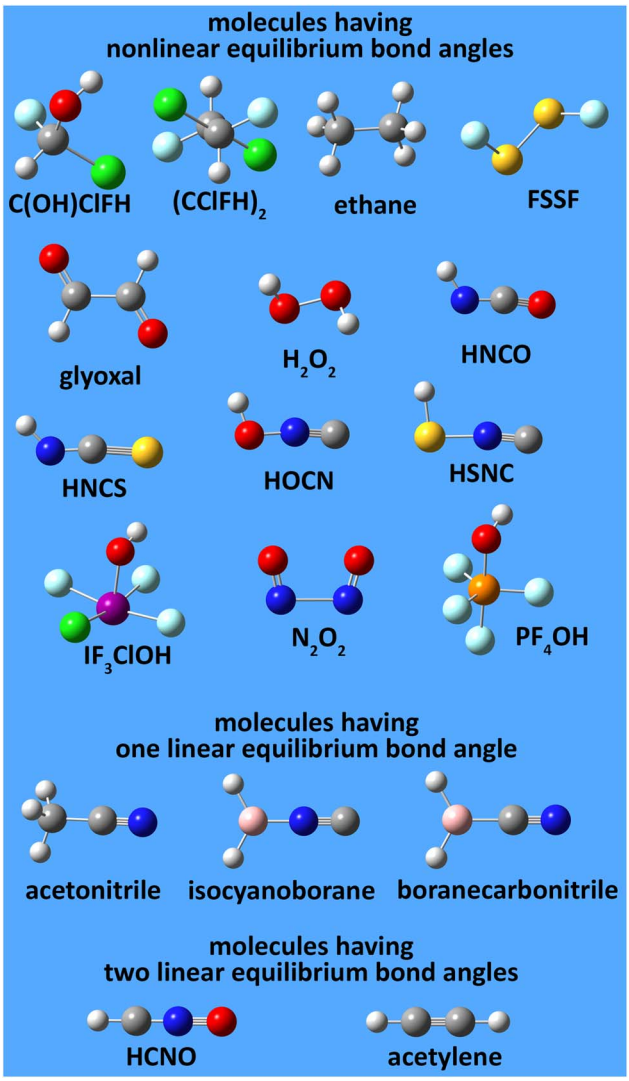

Fig. 7 Optimized geometries of molecules used to study dihedral torsion.

$$U_{\mathrm{Manz_bend}}[\theta] = k_{\mathrm{angle}} \frac{2(\cos\theta - \cos\theta_{\mathrm{eq}})^2}{\sin^2\theta + 3\sin^2\theta_{eq}\left(\frac{\tanh[2\sin[\theta/2]]}{\tanh[2\sin[\theta_{\mathrm{eq}}/2]]}\right)} \quad (228)$$

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

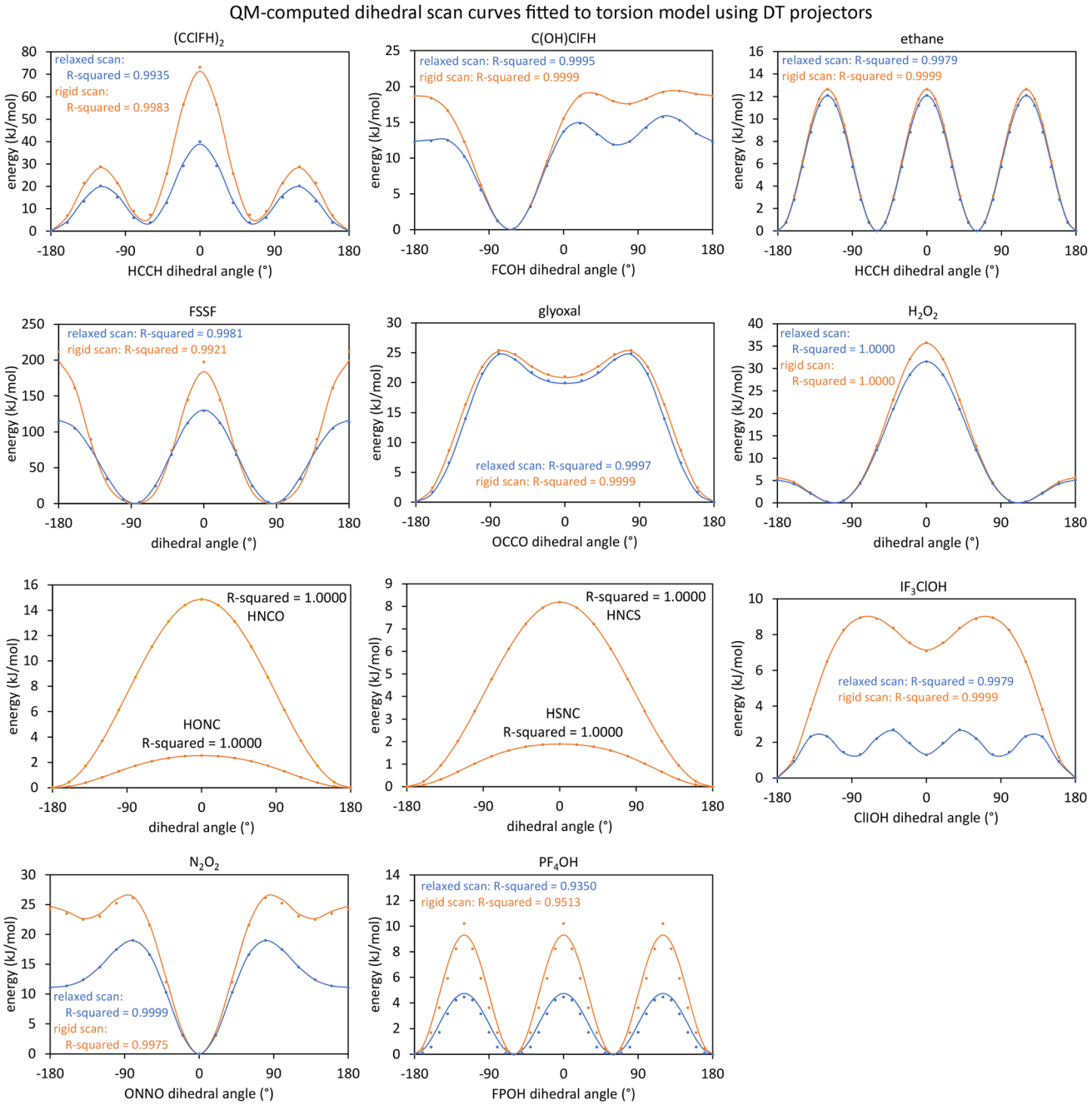

Fig. 8 Example dihedral scan energy curves for several molecules. The *y*-axis plots the energy relative to the low energy conformation. The filled circles show the QM-computed (CCSD/def2-TZVPD) values. The solid lines show the fitted model potential of eqn (161) and (162) using the DT projectors (eqn (122) and (123)) with the parameters from Table 3. Bond angles and bond lengths were held fixed to generate the results shown in orange. Bond angles and bond lengths were relaxed to generate the results (where available) shown in blue. The SumCSq values are listed in Table 3.

In this model potential, the torsion offset potential and angle-bending potential were only included for the one angle of interest and not for any other angles. Since the torsion parameters (listed in Table 3) were already computed from the torsion mode analysis, k_{angle} was the only parameter freely adjusted to generate the model curves displayed in Fig. 9. Agreement between the QM-computed results and the model potential was generally good; however, as the bond angle became acute (<90°) steric repulsion often caused the quantum-mechanically-computed results to rise higher in energy than the model potential. For HNCO, HNCS, HONC, and HSNC, the angle-scan energy curve was completely described by the ADDT torsion offset potential with negligible contribution (*i.e.*, $k_{angle} = 0$) from the angle-bending model potential. For C(OH)ClFH, FSSF, HOOH, IF_3ClOH, ONNO, and PF_4OH, the angle-scan energy curve was almost completely described by the angle-bending model potential with only a tiny contribution from the torsion offset potential. In summary, these results show the

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

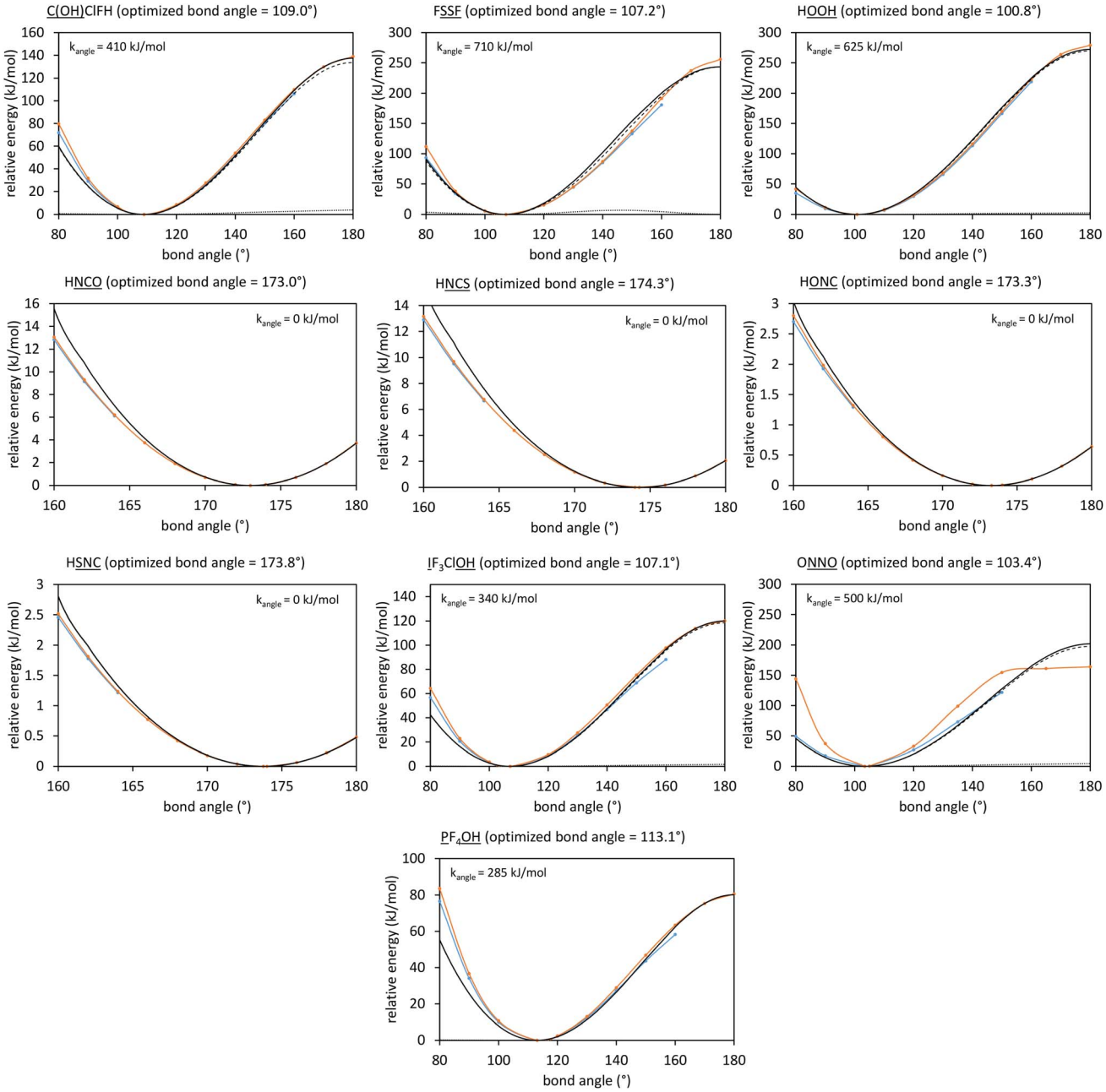

Fig. 9 Angle-scan energy curves for several polyatomic molecules. The orange curves show the QM-computed (CCSD/def2-TZVPD) values holding the bond lengths (and other geometric parameters) fixed as the constrained angle varied, while the blue points (CCSD/def2-TZVPD) relaxed all geometric parameters except the constrained angle. (Blue points were not available in some instances due to the bond angle being too close to linear for either the relaxed torsion scan or the relaxed angle-bending energy to be computed.) The model potential (solid black line) is the sum of the angle-bending model potential (dashed black line) and the ADDT torsion offset potential (dotted black line). The angle-bending force constant is displayed treating radians as dimensionless units. The ADDT torsion offset potential was computed using the rigid torsion parameters from Table 3.

angle-bending model potential is typically much more significant than the torsion offset potential when the equilibrium bond angle is highly bent (*i.e.*, $(\pi - \theta_{eq})$ is large).

The torsion offset potential occurs for the cosine modes (*e.g.*, modes 1 to 4) of the ADDT model potential. There is no torsion offset potential for the sine modes of the ADDT and CADT model potentials. The torsion offset potential is also zero for the CADT cosine modes, because $TOP_n = J_n - H_n$ (see eqn (106)) and $J_n = H_n = 1$ (see eqn (136)) for the CADT model potential.

An angle-bending model potential must have a slope of zero at $\theta = \pi$ in order for its derivative (and hence force) to be continuous at $\theta = \pi$. As explained in prior literature, this restriction arises from the reflection symmetry of $\cos[\theta]$ about $\theta = \pi$.[10,20] In contrast to the bond angle, the directed dihedral is

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 3 Torsion mode analysis using the DT projectors for 13 molecules at the CCSD/def2-TZVPD level of theory. For the results without parentheses, all bond angles and bond lengths were held fixed at values taken from the ground-state structure, and the dihedral value was rigidly scanned to generate the energy curves. Where available, results shown inside parentheses allowed the bond angles and bond lengths to relax as the constrained dihedral value was scanned to generate the energy curves. For HNCO, HNCS, HONC, and HSNC, results in square brackets allowed the bond lengths to relax while the bond angles were held rigid

Molecule	Atoms in dihedral	ϕ_{eq} (°)	QM torsion barrier (kJ mol^{-1})	QM torsion norm (kJ mol^{-1})	c_1	c_2	c_3	c_4	c_5	c_6	c_7	SumCSq
$(CClFH)_2$	HCCH	180.0	73.15	19.40	0.5943	−0.3795	0.7048	−0.0690	0.0000	0.0000	0.0003	0.9987
		(180.0)	(39.98)	(10.42)	(0.5058)	(−0.2707)	(0.8123)	(−0.0807)	(0.0000)	(0.0000)	(0.0002)	(0.9955)
C(OH)ClFH	FCOH	−64.70	19.39	6.55	0.8513	0.4885	0.1801	−0.0096	0.0197	0.0400	−0.0470	1.0000
		(−64.70)	(15.78)	(4.72)	(0.8028)	(0.4386)	(0.3742)	(−0.0092)	(0.0899)	(0.0856)	(−0.0863)	(0.9999)
Ethane	HCCH	180.0	12.65	4.47	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	−0.0001	1.0000
		(180.0)	(12.10)	(4.28)	(0.0000)	(0.0000)	(0.9996)	(0.0000)	(0.0000)	(0.0000)	(−0.0001)	(0.9992)
FSSF	FSSF	87.42	212.0	69.23	0.0041	0.9731	0.0014	−0.1768	0.0829	−0.0932	0.0372	0.9951
		(87.42)	(129.6)	(43.80)	(−0.0005)	(0.9902)	(0.0136)	(0.0035)	(−0.0414)	(−0.0798)	(0.0984)	(0.9986)
Glyoxal	OCCO	180.0	25.41	8.48	0.8090	0.5842	0.0599	−0.0255	0.0000	0.0000	0.0000	0.9999
		(180.0)	(24.86)	(8.52)	(0.8215)	(0.5658)	(0.0033)	(−0.0694)	(0.0000)	(0.0000)	(0.0000)	(0.9999)
H_2O_2	HOOH	111.1	35.75	11.99	0.2996	0.4077	−0.0446	−0.0009	−0.7454	0.3338	−0.2738	1.0000
		(111.1)	(31.56)	(10.68)	(0.3014)	(0.4028)	(−0.0322)	(0.0000)	(−0.7479)	(0.3257)	(−0.2835)	(1.0000)
HNCO	HNCO	180.0	14.86	5.25	1.0000	−0.0011	0.0004	0.0000	0.0000	0.0000	0.0000	1.0000
		[180.0]	[14.78]	[5.23]	[1.0000]	[0.0004]	[0.0003]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[1.0000]
HNCS	HNCS	180.0	8.19	2.90	1.0000	−0.0043	0.0002	0.0000	0.0000	0.0000	0.0000	1.0000
		[180.0]	[8.15]	[2.88]	[1.0000]	[-0.0030]	[0.0002]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[1.0000]
HONC	HONC	180.0	2.54	0.903	0.9950	0.1001	0.0001	0.0000	0.0000	0.0000	0.0000	1.0000
		[180.0]	[2.54]	[0.902]	[0.9950]	[0.1003]	[0.0002]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[1.0000]
HSNC	HSNC	180.0	1.89	0.675	0.9914	0.1307	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
		[180.0]	[1.89]	[0.673]	[0.9914]	[0.1310]	[0.0002]	[0.0001]	[0.0000]	[0.0000]	[0.0000]	[1.0000]
IF_3ClOH	ClIOH	180.0	8.95	2.91	0.7753	0.6228	0.0930	0.0488	0.0001	0.0001	0.0000	0.9999
		(180.0)	(2.70)	(0.70)	(0.4091)	(0.2565)	(0.2593)	(0.8356)	(0.0000)	(−0.0001)	(−0.0001)	(0.9987)
N_2O_2	ONNO	0.00	26.13	8.40	0.7431	0.5981	0.2981	−0.0254	0.0000	0.0000	0.0000	0.9994
		(0.00)	(18.99)	(5.38)	(0.3967)	(0.8543)	(0.3347)	(0.0272)	(0.0000)	(0.0000)	(0.0000)	(1.0000)
PF_4OH	FPOH	180.0	10.20	3.34	0.0000	0.0000	0.9862	0.0000	0.0007	0.0000	−0.0018	0.9727
		(180.0)	(4.46)	(1.69)	(0.0000)	(0.0000)	(0.9922)	(0.0000)	(0.0007)	(0.0000)	(−0.0018)	(0.9844)

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 4 Comparison of QM-computed (CCSD/def2-TZVPD) and ADDT model-predicted torsion norms for several molecules having at least one included equilibrium bond angle ≥ 130°. For each geometry, one bond angle and one dihedral were constrained with the other geometric parameters relaxed. The dihedral was constrained to a series of different values to generate the torsion curve at the constrained bond angle. ϕ_{min} is the dihedral value that has the lowest energy. The column labeled 'QM norm ratio' shows the ratio of the QM torsion norm for the constrained bond angle to the partly relaxed torsion norm displayed in Table 3 for the optimized bond angle (the 'optimized norm'). ADDT and CADT model predictions are listed in the last four columns

Molecule	θ_{constr} (°)	θ_{other} (°) at ϕ_{min}	ϕ_{min} (°)	QM torsion norm (kJ mol^{-1})	QM norm ratio	ADDT predicted ϕ_{min} (°)	ADDT predicted norm (kJ mol^{-1})	ADDT predicted norm ratio	CADT prediction equal to optimized norm (kJ mol^{-1})
HNCO	125.0 (NCO)	120.0 (HNC)	180.0	37.84	7.24	180.0	63.59	12.17	5.23
HNCO	140.0 (NCO)	122.2 (HNC)	180.0	28.99	5.55	180.0	38.98	7.46	5.23
HNCO	155.0 (NCO)	123.4 (HNC)	180.0	18.46	3.53	180.0	20.77	3.98	5.23
HNCO	165.0 (NCO)	123.9 (HNC)	180.0	11.13	2.13	180.0	11.47	2.20	5.23
HNCS	165.0 (NCS)	130.5 (HNC)	180.0	7.49	2.60	180.0	8.11	2.81	2.88
HONC	125.0 (ONC)	104.4 (HON)	180.0	8.10	8.98	180.0	12.04	13.35	0.90
HONC	140.0 (ONC)	105.1 (HON)	180.0	6.08	6.74	180.0	7.47	8.29	0.90
HONC	155.0 (ONC)	105.3 (HON)	180.0	3.58	3.97	180.0	3.91	4.33	0.90
HONC	165.0 (ONC)	105.2 (HON)	180.0	2.06	2.28	180.0	2.12	2.35	0.90
HSNC	125.0 (SNC)	95.7 (HSN)	180.0	5.46	8.12	180.0	10.48	15.58	0.67
HSNC	140.0 (SNC)	96.0 (HSN)	180.0	4.65	6.91	180.0	6.44	9.57	0.67
HSNC	155.0 (SNC)	95.8 (HSN)	180.0	2.86	4.25	180.0	3.27	4.87	0.67
HSNC	165.0 (SNC)	95.6 (HSN)	180.0	1.65	2.46	180.0	1.74	2.58	0.67

not required to have reflection symmetry. An angle-scan energy curve for a subgroup of 3 atoms within a larger bonded cluster is not required to have a slope of zero at $\theta = \pi$ for its derivative (and hence force) to be continuous at $\theta = \pi$, because this energy curve is the sum of an angle-bending potential and a torsion offset potential. Accordingly, any non-zero slope of the angle-scan energy curve at $\theta = \pi$ is assigned to the torsion offset potential and not to the angle-bending model potential. In the examples studied here, the QM-computed angle-scan energy curves have highly non-zero slopes at $\theta = \pi$ for HNCO, HNCS, HONC, and HSNC. As shown in Fig. 9, these nonzero slopes were practically perfectly described by the torsion offset potentials. Remarkably, this practically perfect agreement in slopes at $\theta = \pi$ occurred without using any freely adjustable parameters, because the torsion parameters had already been obtained from the torsion mode analysis as listed in Table 3.

Why is it useful to compute both rigid and relaxed angle-bending and dihedral–torsion scans? Comparing the rigid scan energy curve to the relaxed scan energy curve provides extremely valuable insights into the relative importance of some cross terms. First, consider the angle-scan curves shown in Fig. 9. Iff the relaxed scan curve is greatly below the rigid scan curve, then this indicates that changing bond lengths (or other geometric parameters) substantially lowers the energy at non-equilibrium angle values, and in this case bond-bend (or other) cross terms may be needed to construct an accurate forcefield. Iff the relaxed and rigid angle-scan curves are nearly identical, this suggests bond-bend cross terms are not required to construct an accurate forcefield model. Second, consider the dihedral-scan energy curves shown in Fig. 8. Iff the rigid and relaxed torsion scans have nearly identical energy profiles, then this suggests that bond-torsion and bend-torsion cross terms are not required to construct an accurate forcefield model. Iff the relaxed torsion scan curve is greatly below the rigid torsion scan curve, then this indicates that changing bond lengths, angle values, or other geometric parameters substantially lowers the energy at non-equilibrium dihedral values, and in this case bond-torsion, bend-torsion, and/or torsion–torsion cross terms may be needed to construct an accurate forcefield.

For purposes of optimizing flexibility model parameter values, one can choose to use either rigid torsion scans or relaxed torsion scans in the training dataset for rotatable dihedrals. Some quantum-chemistry software packages do not have built-in constrained geometry optimization algorithms that facilitate holding one dihedral's value constant while relaxing all other internal coordinates. For practical reasons, relaxed torsion scans would be extremely difficult to perform using those software packages; consequently, rigid torsion scans would be preferred in those cases. Because rigid torsion scans do not require constrained geometry optimization, they are computationally cheaper and easier to implement than relaxed torsion scans. For slip torsions, either rigid or 'partly relaxed' (*i.e.*, with at least one bond angle's value constrained) scans are required, because this allows sampling the full range of dihedral values. On the other hand, relaxed torsion scans may be preferred in some cases when constructing the training dataset for rotatable dihedrals, because relaxed torsion scans can provide a more accurate and realistic estimate of the torsion energy barrier than rigid torsion scans.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

How accurate is the ADDT model potential for predicting changes in the torsion norm due to non-equilibrium bond angle values? Two datasets were studied to explore this question. The first dataset contained dihedrals for which at least one of the contained equilibrium bond angles was ≥130°. The second dataset contained dihedrals for which both contained equilibrium bond angles were <130°. To generate these two datasets, one bond angle was constrained while the corresponding dihedral was scanned and all other geometric parameters were relaxed. These constrained CCSD/def2-TZVPD calculations were performed in Gaussian 16 using the opt = modredundant method.

For the first dataset, the bond angle chosen for study had an equilibrium value ≥130° in the optimized ground-state structure. Dihedral scans were attempted for HNCO, HNCS, HONC, and HSNC at constrained values of 125°, 140°, 155°, and 165° for the underlined bond angle. In some of these calculations, the other (*i.e.*, unconstrained) bond angle became too close to linear for the default geometry optimizer in Gaussian16 (ref. 33) to converge.

The ADDT model predictions shown in Tables 4, 6 and Fig. 10 were made using the following steps:

(1) (Partly or fully) relaxed torsion mode analysis was performed as shown in Table 3. This yielded the 'optimized' torsion norm and the 'optimized' c_1 to c_7 for each dihedral studied.

(2) CCSD/def2-TZVPD energies were computed at each constrained bond angle value for a series of uniformly spaced dihedral values over the range $-\pi < \phi \leq \pi$. In each of these calculations, one bond angle and one dihedral were held rigid while all other geometric parameters were relaxed. In some cases, molecular symmetry could be exploited. Taking HSNC as an example, QM calculations were performed for HSNC dihedral values of 0°, 20°, 40°, 60°, 80°, 100°, 120°, 140°, 160°, and 180° for each constrained SNC angle value of 125°, 140°, 155°, and 165°. The HSNC molecular symmetry shows the energy at a HSNC dihedral value of $\phi = \gamma$ equals that at a value of $\phi = -\gamma$.

(3) For each constrained bond angle value, $\phi_{\min}$ was identified as the ϕ value having the lowest QM-computed energy. In some cases, this occurred at a symmetry plane such as $\phi = 180°$ (for HSNC) or 0° (for N_2O_2). Otherwise, $\phi_{\min}$ was computed *via* a CCSD/def2-TZVPD calculation that held one bond angle rigid (at the constrained value) while all other geometric parameters were relaxed.

(4) Using results from (2), torsion mode analysis was then performed at each of the constrained bond angle values. This yielded the QM torsion norm and QM torsion mode coefficients for each constrained bond angle value.

(5) The ADDT model predictions were made as follows. Starting with the optimized torsion norm and coefficients as computed in 1), the potential was then rewritten as:

$$U^{\mathrm{ADDT}}_{\mathrm{ABCD}}[\text{Ж}_{\mathrm{ABC}}, \text{Ж}_{\mathrm{BCD}}, \phi_{\mathrm{ABCD}}] - U^{\mathrm{ADDT}}_{\mathrm{ABCD}}[\text{Ж}^{\mathrm{eq}}_{\mathrm{ABC}}, \text{Ж}^{\mathrm{eq}}_{\mathrm{BCD}}, \phi^{\mathrm{eq}}_{\mathrm{ABCD}}] = \sum_{n=1}^{4} \Big(k^n_\phi J_n[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] + H_n[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] \Big(-k^n_\phi \cos\big[n(\phi - \phi_{\mathrm{eq}})\big] + S_{\mathrm{instance}} b^n_\phi \sin\big[n(\phi - \phi_{\mathrm{eq}})\big] \Big) \Big) \quad (229)$$

The f^{ABC}_n, f^{BCD}_n, $f^{\mathrm{ABC}}_{n_\mathrm{eq}}$, and $f^{\mathrm{BCD}}_{n_\mathrm{eq}}$ values were computed using the θ_{constr}, θ_{other}, angle_1, and angle_2 values listed in Tables 1 and 4 or 6. The predicted norm and mode coefficients were then computed as

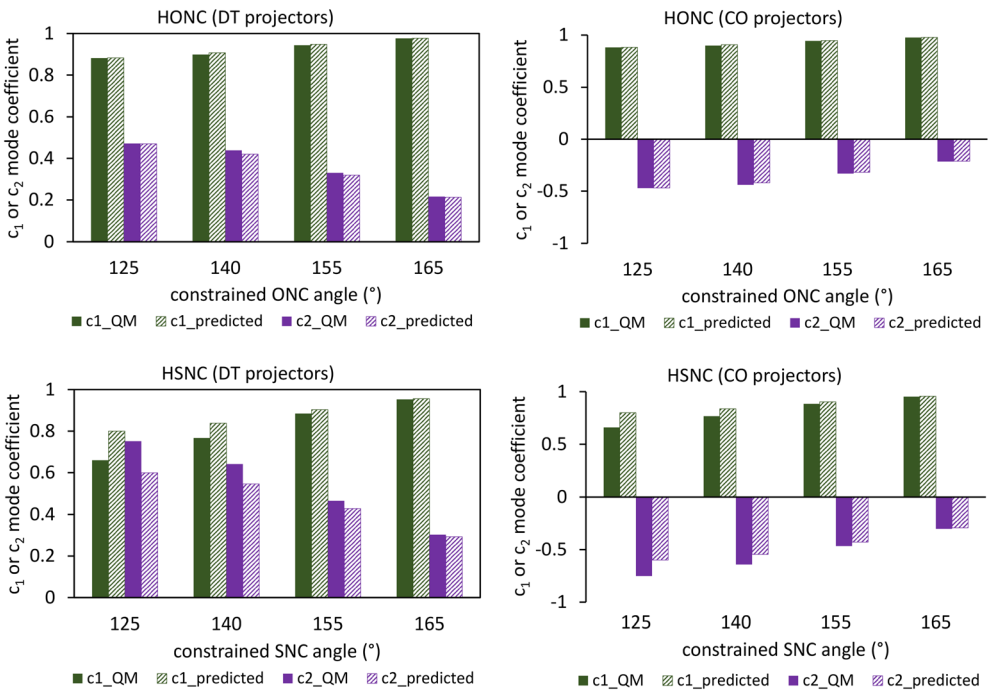

Fig. 10 Comparison of c_1 and c_2 mode coefficients for HONC (top panels) and HSNC (bottom panels) molecules at different constrained values of the ONC or SNC bond angle. Solid bars show the QM-extracted values using the DT projectors (left panels) and CO projectors (right panels). Diagonally-hashed bars show the values predicted using the ADDT model (left panels) and ADCO model (right panels).

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$\text{predicted_norm} = \sqrt{\frac{1}{2}\sum_{n=1}^{4}\left(H_n[\theta_{ABC},\theta_{BCD}]\right)^2\left(\left(k_\phi^n\right)^2+\left(b_\phi^n\right)^2\right)} \quad (230)$$

$$\text{predicted_}c_n = \frac{1}{\text{predicted_norm}}\left(\frac{H_n[\theta_{ABC},\theta_{BCD}]k_\phi^n}{\sqrt{2}}\right) \quad \text{for } n=1 \text{ to } 4 \quad (231)$$

The factors of $\sqrt{\frac{1}{2}}$ appear in eqn (230) and (231), because the root-mean-squared values of cosine and sine functions is $\sqrt{\frac{1}{2}}$. *Caution:* The relationship between the predicted torsion norm and the torsion force constants as shown in eqn (230) and (231) holds only when the relevant internal coordinates are non-redundant and there are no multicollinearity issues of the related flexibility terms. This condition was clearly met for the dihedral torsions studied here.

(6) For each value of the constrained bond angle, the ADDT predicted ϕ_{min} value was computed by calculating the full $U_{ABCD}^{ADDT}[\text{Ж}_{ABC}, \text{Ж}_{BCD}, \phi_{ABCD}]$ curve over the range $-180° < \phi_{ABCD} \leq 180°$ in 0.01° increments and identifying which value of ϕ_{ABCD} produced the minimum value of $U_{ABCD}^{ADDT}[\text{Ж}_{ABC}, \text{Ж}_{BCD}, \phi_{ABCD}]$.

For the first dataset, QM torsion norms for all of the converged calculations are listed in Table 4 and compared to predictions of the ADDT and CADT models. Here, the CADT model approximated the torsion norm as equal to the partly relaxed torsion norm listed in Table 3. As shown in Table 4, for these systems the QM torsion norm systematically increased as the constrained bond angle decreased in value. My ADDT model reproduced this trend, while the CADT model did not.

Table 5 summarizes the torsion mode coefficients extracted from these QM results. For HNCO and HNCS, only c_1 was larger in absolute value than 0.1. For HONC and HSNC, both c_1 and c_2 were larger in absolute value than 0.1, while c_3 to c_7 were smaller than this. For HONC and HSNC, c_2 systematically increased in magnitude as the constrained bond angle decreased. As shown in Fig. 10, my ADDT model predicted this trend with good quantitative accuracy.

The second dataset contained six molecules. In each molecule, the selected dihedral had both included equilibrium bond angles <130° in its optimized ground-state structure. For the same three bonded atoms as studied in Fig. 9, torsion mode analysis was performed at constrained bond angle values 10° larger and 10° smaller than the equilibrium value. As shown in Fig. 9, the energy penalty for changing the bond angle is much larger for the bond angles studied in dataset 2 compared to those in dataset 1. In other words, the bond angles in dataset 2 were much stiffer than those in dataset 1. Accordingly, during a molecular dynamics or Monte Carlo simulation, the bond angles in dataset 2 would be expected to change less compared to those in dataset 1. This explains why Table 6 for dataset 2 studied smaller changes in the constrained bond angle compared to Table 4 for dataset 1. As shown in Table 6, the QM torsion norm increased more often (~2/3 of the time) than decreased (~1/3 of the time) as the constrained bond angle decreased. The ADDT model predicted a systematic increase in the torsion norm as the constrained bond angle decreased, while the CADT model predicted unchanged amplitudes.

Table 7 lists summary statistics for the ADDT and CADT models. The mean $\log_{10}$ error (MLE) was defined as the average value of $\log_{10}$[predicted_norm/QM_norm]. The mean unsigned $\log_{10}$ error (MULE) was defined as the average value of abs [$\log_{10}$[predicted_norm/QM_norm]]. For dataset 1, the ADDT model performed much better than the CADT model. For dataset 2, both models performed acceptably and neither was substantially more accurate than the other. Given the much lower computational cost of the CADT model compared to the ADDT model, this clearly leads to the following conclusions. If the dihedral contains a bond angle whose equilibrium value is ≥ 130°, the ADDT model is preferred both due to its higher accuracy and also due to its higher stability, because it yields continuous derivatives even if the perturbed bond angle reaches linearity (*i.e.*, 180°) during a molecular dynamics or Monte Carlo simulation. If both contained bond angles in the dihedral have equilibrium values < 130°, then these bond angles are likely to be sufficiently stiff that it will be extremely rare for the perturbed bond angle to reach linearity (*i.e.*, 180°) during

Table 5 QM-computed torsion mode coefficients using the DT projectors for the same systems and constrained bond angles as analyzed in Table 4

Molecule	θ_{constr} (°)	c_1	c_2	c_3	c_4	c_5, c_6, c_7	SumCSq
HNCO	125.0	0.9994	0.0143	0.0308	−0.0011	0.0000	1.0000
HNCO	140.0	0.9999	0.0070	0.0119	−0.0012	0.0000	1.0000
HNCO	155.0	1.0000	0.0035	0.0035	0.0000	0.0000	1.0000
HNCO	165.0	1.0000	0.0019	0.0012	0.0001	0.0000	1.0000
HNCS	165.0	1.0000	−0.0040	−0.0004	0.0005	0.0000	1.0000
HONC	125.0	0.8816	0.4718	0.0089	0.0022	0.0000	1.0000
HONC	140.0	0.8985	0.4389	0.0072	0.0003	0.0000	1.0000
HONC	155.0	0.9438	0.3304	0.0026	0.0003	0.0000	1.0000
HONC	165.0	0.9764	0.2160	0.0010	0.0002	0.0000	1.0000
HSNC	125.0	0.6597	0.7514	−0.0107	0.0059	0.0000	1.0000
HSNC	140.0	0.7672	0.6414	0.0039	0.0013	0.0000	1.0000
HSNC	155.0	0.8847	0.4661	0.0031	0.0001	0.0000	1.0000
HSNC	165.0	0.9530	0.3029	0.0014	0.0000	0.0000	1.0000

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 6 Comparison of QM-computed (CCSD/def2-TZVPD) and ADDT model-predicted torsion norms for several molecules having both included equilibrium bond angles < 130°. For each geometry, one bond angle and one dihedral were constrained with the other geometric parameters relaxed. The dihedral was constrained to a series of different values to generate the torsion curve at the constrained bond angle. ϕ_{min} is the dihedral value that has the lowest energy. The column labeled 'QM norm ratio' shows the ratio of the QM torsion norm for the constrained bond angle to the fully relaxed torsion norm (the 'optimized norm') displayed in Table 3. ADDT and CADT model predictions are listed in the last four columns

Molecule	θ_{constr} (°)	θ_{other} (°) at ϕ_{min}	ϕ_{min} (°)	QM torsion norm (kJ mol^{-1})	QM norm ratio	ADDT predicted ϕ_{min} (°)	ADDT predicted norm (kJ mol^{-1})	ADDT predicted norm ratio	CADT prediction equal to optimized norm (kJ mol^{-1})
C(OH)ClFH	99.0 (COH)	110.5 (FCO)	−60	4.61	0.98	−65	5.55	1.18	4.72
C(OH)ClFH	119.0 (COH)	111.5 (FCO)	−68	4.86	1.03	−65	3.80	0.81	4.72
FSSF	97.2 (SSF)	106.0 (FSS)	±87[a]	39.83	0.91	±~85[a]	51.4	1.17	43.8
FSSF	117.2 (SSF)	108.2 (FSS)	±88[a]	60.04	1.37	±~86[a]	35.0	0.80	43.8
H_2O_2	90.8 (OOH)	100.6 (HOO)	±121[a]	14.45	1.35	±~111[a]	11.9	1.11	10.7
H_2O_2	110.8 (OOH)	101.0 (HOO)	±103[a]	8.41	0.79	±~112[a]	9.15	0.86	10.7
IF_3ClOH	97.1 (IOH)	84.4 (ClIO)	180	1.21	1.74	180	0.86	1.23	0.70
IF_3ClOH	117.1 (IOH)	83.0 (ClIO)	180	0.55	0.78	180	0.51	0.73	0.70
N_2O_2	93.4 (NNO)	98.7 (ONN)	0	7.79	1.45	0	6.50	1.21	5.38
N_2O_2	113.4 (NNO)	107.5 (ONN)	0	3.55	0.66	0	4.20	0.78	5.38
PF_4OH	103.1 (POH)	88.3 (FPO)	180 (−60, 60)[b]	2.04	1.21	60 (−60, 180)[b]	2.11	1.26	1.67[c]
PF_4OH	123.1 (POH)	88.0 (FPO)	180 (−60, 60)[b]	1.33	0.79	60 (−60, 180)[b]	1.19	0.71	1.67[c]

[a] This molecule's QM energy is unchanged upon mirror reflection which changes the dihedral's sign. For the FSSF and HOOH molecules, the ADDT model only approximately (not exactly) captures this reflection symmetry. [b] Due to the 3-fold symmetry of this molecule, there are chemically equivalent minima at $\phi = 60°$, 180°, and −60°. [c] For PF_4OH, the model-fitted fully relaxed norm from the seven fitted dihedral modes is 1.67 kJ mol^{-1} compared to 1.69 kJ mol^{-1} for a complete modal expansion; this reflects the SumCSq value of 0.98 as listed in Table 3.

a molecular dynamics or Monte Carlo simulation. In this case, the CADT model is preferred, because it is computationally cheaper than the ADDT model yet of comparable accuracy for these dihedrals.

10.3 Detailed torsion analysis for rotatable dihedrals using ADCO and CACO model potentials

The goal of this section is to repeat the analysis of Section 10.2 except using the ADCO and CACO model potentials in place of the ADDT and CADT model potentials. The key principle to keep in mind is that the CADT model potential is functionally equivalent to a full Fourier series expansion of the torsion potential as a function of the dihedral value ϕ subject to the constraint that

$$\left.\frac{dU^{CADT}_{ABCD}[\phi]}{d\phi}\right|_{\phi=\phi^{training}_{eq}} = 0 \quad (232)$$

In stark contrast, the CACO model potential spans the subset of even functions of the dihedral value ϕ while omitting all odd-function contributions. A constraint analogous to eqn (232) is not imposed for the CACO model potential; however, if *sym*_value = 0 and the CACO model potential is untruncated then fitting the CACO model potential to a torsion scan curve should cause $dU^{CACO}_{ABCD}[\phi]/d\phi|^{training}_{\phi=\phi eq}$ to asymptotically approach zero.

It directly follows that CO projectors (which are used by the CACO and ADCO model potentials) provide asymptotically close to functionally equivalent results to the DT projectors (which are used by the CADT and ADDT model potentials) for a torsion scan curve that has no odd-function contributions; that is, when *sym*_value = 0. When *sym*_value = 0, the SumCSq values from Table 3 computed using the DT projectors were nearly identical to the SumCSq values in Table 8 computed using the CO projectors. On the other hand, when *sym*_value ≠ 0 (*e.g.*, *sym*_value = 0.847 for the FCOH dihedral in the C(OH)ClFH molecule) then the SumCSq value using the CO projectors was much smaller than the SumCSq value using the DT projectors. When *sym*_value = 0, the model fits using the CO projectors were nearly identical to those using the DT projectors; this is visually illustrated by comparing the torsion scan curves in Fig. 8 (using the DT projectors) to those in ESI Fig. S1† (using the CO projectors). However, the CO projectors performed markedly worse than the DT projectors when *sym*_value ≠ 0 (*e.g.*, for the FCOH dihedral in the C(OH)ClFH molecule as shown in Fig. 11).

Table 7 Summary statistics comparing the ADDT model potential to the CADT model potential. MLE = mean $\log_{10}$ error; MULE = mean unsigned $\log_{10}$ error

	ADDT model		CADT model	
Dataset	MLE	MULE	MLE	MULE
An equil. angle ≥ 130°	0.098	0.098	−0.647	0.647
Both equil. angles < 130°	−0.035	0.087	−0.019	0.109

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 8 Torsion mode analysis using the CO projectors for 13 molecules at the CCSD/def2-TZVPD level of theory. For the results without parentheses, all bond angles and bond lengths were held fixed at values taken from the ground-state structure, and the dihedral value was rigidly scanned to generate the energy curves. Where available, results shown inside parentheses allowed the bond angles and bond lengths to relax as the constrained dihedral value was scanned to generate the energy curves. For HNCO, HNCS, HONC, and HSNC, results in square brackets allowed the bond lengths to relax while the bond angles were held rigid

Molecule	Atoms in dihedral	ϕ_{eq} (°)	QM torsion barrier (kJ mol^{-1})	QM torsion norm (kJ mol^{-1})	*sym*_value	c_1	c_2	c_3	c_4	SumCSq
$(CClFH)_2$	HCCH	180.0	73.15	19.40	0.000	0.5943	0.3795	0.7048	0.0690	0.9987
		(180.0)	(39.98)	(10.42)	(0.000)	(0.5058)	(0.2707)	(0.8123)	(0.0807)	(0.9955)
C(OH)ClFH	FCOH	−64.70	19.39	6.55	0.847	−0.3577	0.3475	0.1848	−0.0092	0.2829
		(−64.70)	(15.78)	(4.72)	(0.802)	(−0.2859)	(0.3551)	(0.3860)	(−0.0090)	(0.3569)
Ethane	HCCH	180.0	12.65	4.47	0.000	0.0000	0.0000	1.0000	0.0000	1.0000
		(180.0)	(12.10)	(4.28)	(0.000)	(0.0000)	(0.0000)	(0.9996)	(0.0000)	(0.9992)
FSSF	FSSF	87.42	212.0	69.23	0.000	−0.0908	0.9771	0.0101	0.1797	0.9953
		(87.42)	(129.6)	(43.80)	(0.000)	(0.0103)	(0.9943)	(0.1011)	(−0.0035)	(0.9990)
Glyoxal	OCCO	180.0	25.41	8.48	0.000	0.8090	−0.5842	0.0599	0.0255	0.9999
		(180.0)	(24.86)	(8.52)	(0.000)	(0.8215)	(−0.5658)	(0.0033)	(0.0694)	(0.9999)
H_2O_2	HOOH	111.1	35.75	11.99	0.000	0.8339	0.5495	0.0500	0.0094	1.0000
		(111.1)	(31.56)	(10.68)	(0.000)	(0.8390)	(0.5430)	(0.0361)	(0.0001)	(1.0000)
HNCO	HNCO	180.0	14.86	5.25	0.000	1.0000	0.0011	0.0004	0.0000	1.0000
		[180.0]	[14.78]	[5.23]	[0.000]	[1.0000]	[−0.0004]	[0.0003]	[0.0000]	[1.0000]
HNCS	HNCS	180.0	8.19	2.90	0.000	1.0000	0.0043	0.0002	0.0000	1.0000
		[180.0]	[8.15]	[2.88]	[0.000]	[1.0000]	[0.0030]	[0.0002]	[0.0000]	[1.0000]
HONC	HONC	180.0	2.54	0.903	0.000	0.9950	−0.1001	0.0001	0.0000	1.0000
		[180.0]	[2.54]	[0.902]	[0.000]	[0.9950]	[−0.1003]	[0.0002]	[0.0000]	[1.0000]
HSNC	HSNC	180.0	1.89	0.675	0.000	0.9914	−0.1307	0.0000	0.0000	1.0000
		[180.0]	[1.89]	[0.673]	[0.000]	[0.9914]	[−0.1310]	[0.0002]	[−0.0001]	[1.0000]
IF_3ClOH	ClIOH	180.0	8.95	2.91	0.000	0.7753	−0.6228	0.0930	−0.0488	0.9999
		(180.0)	(2.70)	(0.70)	(0.000)	(0.4091)	(−0.2565)	(0.2593)	(−0.8356)	(0.9987)
N_2O_2	ONNO	0.00	26.13	8.40	0.000	−0.7431	−0.5981	−0.2981	0.0254	0.9994
		(0.00)	(18.99)	(5.38)	(0.000)	(−0.3967)	(−0.8543)	(−0.3347)	(−0.0272)	(1.0000)
PF_4OH	FPOH	180.0	10.20	3.34	0.000	0.0000	0.0000	0.9862	0.0000	0.9727
		(180.0)	(4.46)	(1.69)	(0.000)	(0.0000)	(0.0000)	(0.9922)	(0.0000)	(0.9844)

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Using the trigonometric identities in eqn (26) and (27), the DT projectors for modes 1 to 4 can be re-written as

$$P_n^{\mathrm{DT}}[\phi] = -\cos[n(\phi - \phi_{\mathrm{eq}}^{\mathrm{training}})] = -\cos[n\phi_{\mathrm{eq}}^{\mathrm{training}}]\cos[n\phi] - \sin[n\phi_{\mathrm{eq}}^{\mathrm{training}}]\sin[n\phi] \quad (233)$$

Inserting this into eqn (164) gives

$$c_n^{\mathrm{DT}} = -\cos\left[n\phi_{\mathrm{eq}}^{\mathrm{training}}\right]\int_{-\pi}^{\pi}\frac{\cos[n\phi]}{\sqrt{\pi}}\left(\frac{E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi] - E_{\mathrm{RTS}}^{\mathrm{QM_avg}}}{\sqrt{w}}\right)\mathrm{d}\phi - \sin\left[n\phi_{\mathrm{eq}}^{\mathrm{training}}\right]\int_{-\pi}^{\pi}\frac{\sin[n\phi]}{\sqrt{\pi}}\left(\frac{E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi] - E_{\mathrm{RTS}}^{\mathrm{QM_avg}}}{\sqrt{w}}\right)\mathrm{d}\phi \quad (234)$$

When *sym*_value = 0, the odd-function components of the torsion potential are zero, which means that

$$\int_{-\pi}^{\pi}\frac{\sin[n\phi]}{\sqrt{\pi}}\left(\frac{E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi] - E_{\mathrm{RTS}}^{\mathrm{QM_avg}}}{\sqrt{w}}\right)\mathrm{d}\phi = 0 \quad (235)$$

so that eqn (234) reduces in this case to

$$c_n^{\mathrm{DT}} = -\cos\left[n\phi_{\mathrm{eq}}^{\mathrm{training}}\right]\int_{-\pi}^{\pi}\frac{\cos[n\phi]}{\sqrt{\pi}}\left(\frac{E_{\mathrm{RTS}}^{\mathrm{QM}}[\phi] - E_{\mathrm{RTS}}^{\mathrm{QM_avg}}}{\sqrt{w}}\right)\mathrm{d}\phi \quad (236)$$

Combining eqn (145), (164), and (236) gives

$$\Delta_n = c_n^{\mathrm{DT}} + c_n^{\mathrm{CO}}\cos[n\phi_{\mathrm{eq}}^{\mathrm{training}}] \quad (237)$$

equals zero when *sym*_value = 0. This is numerically confirmed for rigid torsion scans and relaxed (or partly relaxed) torsion scans of 12 molecules in Table S1 of the ESI.†

The torsion offset potential for the ADCO model potential can be computed by substituting $\phi = \phi_{\mathrm{eq}}^{\mathrm{training}}$ into the ADCO model potential:

$$\mathrm{TOP}[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] = k_{\mathrm{ADCO}}\sum_{n=1}^{n_{\max}} c_n^{\mathrm{CO}}\cos\left[n\phi_{\mathrm{eq}}^{\mathrm{training}}\right](H_n - J_n) \quad (238)$$

Comparing this to the TOP using the ADDT model potential

$$\mathrm{TOP}[\theta_{\mathrm{ABC}}, \theta_{\mathrm{BCD}}] = \sum_{n=1}^{4} k_\phi^n(J_n - H_n) \quad (239)$$

shows the two are equal iff

$$k_\phi^n = -k_{\mathrm{ADCO}}c_n^{\mathrm{CO}}\cos[n\phi_{\mathrm{eq}}^{\mathrm{training}}] \quad (240)$$

for $n = 1$ to 4. When *sym*_value = 0, then $\Delta_n = 0$ so eqn (240) holds in this case. The ADCO model potential and eqn (240) do not apply when *sym*_value ≠ 0.

The results shown in Table S1 of the ESI† confirm that eqn (240) is satisfied for all 12 molecules having *sym*_value = 0. Accordingly, the angle-scan curves shown in Fig. 9 are unchanged when replacing the ADDT model potential with the ADCO model potential for those molecules having *sym*_value = 0. This includes all molecules shown in Fig. 9 except C(OH)ClFH (for which *sym*_value = 0.847). In summary, the ADCO model potential gives angle-scan curves and TOP identical to those of the ADDT model potential for FSSF, HOOH, HNCO, HNCS, HONC, HSNC, IF_3ClOH, ONNO, and PF_4OH.

ESI Table S2† lists ADCO and CACO model predictions for four molecules (HNCO, HNCS, HONC, and HSNC) that have at least one equilibrium bond angle ≥ 130°. ESI Table S2† re-analyzes the systems and constrained bond angles that were previously analyzed using DT projectors in Tables 4 and 5. Although the mode coefficients using CO projectors were different from those using the DT projectors, the SumCSq values (*i.e.*, SumCSq = 1.0000) and predicted norms for each of these four molecules using the CO projectors were identical to those using the DT projectors.

Table 9 lists ADCO and CACO model predictions for five molecules (FSSF, H_2O_2, IF_3ClOH, N_2O_2, and PF_4OH) that have both equilibrium bond angles < 130° and *sym*_value = 0. For these five molecules, the ADCO and CACO predicted norms in Table 9 are identical to the ADDT and CADT predicted norms, respectively, in Table 6.

For each value of the constrained bond angle in Table 9 and ESI Table S2,† the ADCO predicted $\phi_{\min}$ value was computed by

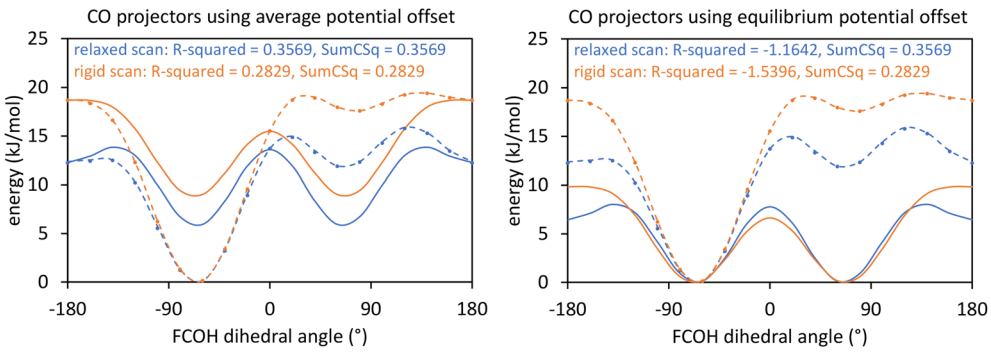

Fig. 11 When *sym*_value > 0, then the CO projectors do not provide a good description of the dihedral scan curve. This figure compares model fits using the DT projectors (dashed lines) to CO projectors using the average potential offset (solid lines in lefthand panel) and CO projectors using the equilibrium potential offset (solid lines in righthand panel) for the S enantiomer of C(OH)ClFH. The filled circles show the QM-computed (CCSD/def2-TZVPD) values. Bond angles and bond lengths were held fixed to generate the results shown in orange. Bond angles and bond lengths were relaxed to generate the results shown in blue. When using the DT projectors, the *R*-squared values were 0.9995 (relaxed scan) and 0.9999 (rigid scan), and the SumCSq values using the DT projectors were 0.9999 (relaxed scan) and 1.0000 (rigid scan).

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 9 Comparison of QM-computed (CCSD/def2-TZVPD) and ADCO model-predicted torsion norms for several molecules having both included equilibrium bond angles < 130° and *sym*_value = 0. ϕ_{min} is the dihedral value that has the lowest energy. The column labeled 'QM norm ratio' shows the ratio of the QM torsion norm for the constrained bond angle to the fully relaxed torsion norm. CACO and ADCO model predictions are listed in the last five columns. $\Delta\theta_{constr} = \theta_{constrained} - \theta_{eq}$

Molecule	$\Delta\theta_{constr}$ (°)	ϕ_{min} (°)	QM torsion norm (kJ mol^{-1})	QM norm ratio	CACO predicted ϕ_{min} (°)	CACO predicted norm (kJ mol^{-1})	ADCO predicted ϕ_{min} (°)	ADCO predicted norm (kJ mol^{-1})	ADCO predicted norm ratio
FSSF	0 (SSF)	±87.42	43.80	1.00	±85.9	43.8	±85.9	43.8	1.00
FSSF	−10 (SSF)	±86.73	39.83	0.91	±85.9	43.8	±85.5	51.4	1.18
FSSF	+10 (SSF)	±87.72	60.04	1.37	±85.9	43.8	±89.3	35.0	0.80
H_2O_2	0 (OOH)	±111.1	10.68	1.00	±111.3	10.7	±111.3	10.7	1.00
H_2O_2	−10 (OOH)	±120.8	14.45	1.35	±111.3	10.7	±120.8	11.9	1.11
H_2O_2	+10 (OOH)	±103.5	8.41	0.79	±111.3	10.7	±103.5	9.15	0.86
IF_3ClOH	0 (IOH)	180.0	0.70	1.00	180.0	0.70	180.0	0.70	1.00
IF_3ClOH	−10 (IOH)	180.0	1.21	1.74	180.0	0.70	180.0	0.86	1.23
IF_3ClOH	+10 (IOH)	180.0	0.55	0.78	180.0	0.70	180.0	0.51	0.73
N_2O_2	0 (NNO)	0.00	5.38	1.00	0.00	5.38	0.00	5.38	1.00
N_2O_2	−10 (NNO)	0.00	7.79	1.45	0.00	5.38	0.00	6.50	1.21
N_2O_2	+10 (NNO)	0.00	3.55	0.66	0.00	5.38	0.00	4.20	0.78
PF_4OH	0 (POH)	180 (−60, 60)[a]	1.69	1.00	180 (−60, 60)[a]	1.67[b]	180 (−60, 60)[a]	1.67[b]	1.00
PF_4OH	−10 (POH)	180 (−60, 60)[a]	2.04	1.21	180 (−60, 60)[a]	1.67[b]	180 (−60, 60)[a]	2.11	1.26
PF_4OH	+10 (POH)	180 (−60, 60)[a]	1.33	0.79	180 (−60, 60)[a]	1.67[b]	180 (−60, 60)[a]	1.19	0.71

[a] Due to the 3-fold symmetry of this molecule, there are chemically equivalent minima at $\phi = 60°$, 180°, and −60°. [b] For PF_4OH, the model-fitted fully relaxed norm from the four fitted CO dihedral modes is 1.67 kJ mol^{-1} compared to 1.69 kJ mol^{-1} for a complete modal expansion; this reflects the SumCSq value of 0.98 as listed in Table 8.

calculating the full $U^{ADCO}_{ABCD}[\text{Ж}_{ABC}, \text{Ж}_{BCD}, \phi_{ABCD}]$ curve over the range $-180° < \phi_{ABCD} \leq 180°$ in 0.01° increments and identifying which value of ϕ_{ABCD} produced the minimum value of $U^{ADCO}_{ABCD}[\text{Ж}_{ABC}, \text{Ж}_{BCD}, \phi_{ABCD}]$. The CACO predicted ϕ_{min} value remains constant as the bond angle changes, because the CACO model potential is not a function of bond angle. The CACO and ADCO predicted ϕ_{min} values are equal at the optimized bond angle value (*i.e.*, at $\Delta\theta_{constr} = 0$).

In Table 9 and ESI Table S2,† the ADCO predicted torsion norm was computed as

$$\text{predicted_norm} = k_{ADCO}\sqrt{\frac{1}{2}\sum_{n=1}^{4}\left(c_n^{CO}H_n[\theta_{ABC}, \theta_{BCD}]\right)^2} \quad (241)$$

The factor of $\sqrt{\frac{1}{2}}$ appears in eqn (241), because the root-mean-squared value of cosine functions is $\sqrt{\frac{1}{2}}$. Table 10 summarizes the mean $\log_{10}$ error (MLE) and mean unsigned $\log_{10}$ error (MULE) for the ADCO and CACO model predicted torsion norms. When (θ^{eq}_{ABC} or θ^{eq}_{BCD}) ≥ 130°, the ADCO model was much more accurate than the CACO model. When (θ^{eq}_{ABC} and θ^{eq}_{BCD}) < 130°, both the ADCO and CACO models performed acceptably and had similar accuracy to each other.

As shown in Fig. 10, the ADCO model correctly predicted trends for changes in c_n^{CO} values as the constrained bond angle changed in the HONC and HSNC molecules. The predicted coefficients were computed using

$$\text{predicted_}c_n = \frac{k_{ADCO}c_n^{CO}H_n[\theta_{ABC}, \theta_{BCD}]}{\text{predicted_norm}\sqrt{2}} \quad (242)$$

In summary, when *sym*_value = 0 then the ADDT and ADCO models provide nearly functionally equivalent results to each other if all seven ADDT modes and all four ADCO modes are included in the model potential. If some of the less important (but nonzero) modes are removed during smart selection so that the model is restricted to the more important ADDT or ADCO modes, then this introduces additional approximation that can make model results slightly different for the ADDT and ADCO model potentials. When *sym*_value = 0, then the CADT and CACO models provide nearly functionally equivalent results to each other if all seven CADT modes and all four CACO modes are included in the model potential. If some of the less important (but nonzero) modes are removed during smart selection so that the model is restricted to the more important CADT or CACO modes, then this introduces additional approximation that can make model results slightly different

Table 10 Summary statistics comparing the ADCO model potential to the CACO model potential for molecules having *sym*_value = 0. MLE = mean $\log_{10}$ error; MULE = mean unsigned $\log_{10}$ error

	ADCO model		CACO model	
Dataset	MLE	MULE	MLE	MULE
An equil. angle ≥ 130°	0.098	0.098	−0.647	0.647
Both equil. angles < 130°	−0.039	0.086	−0.022	0.128

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

for the CADT and CACO model potentials. Of course, the ADCO and CACO model potentials are not good approximations if *sym*_value differs substantially from zero.

10.4 Slip torsion and the indispensable torsion offset potential

By comparing model potential energy surfaces to the CCSD-calculated potential energy surface, Fig. 12 shows that including the torsion offset potential is absolutely required to describe angle-damped dihedral torsions. The CCSD-calculated potential energy surface used the underlying quantum chemistry data for the constrained torsion scans at NCO angle = 125°, 140°, 155°, and 165° listed in Table 4, plus the energy for linear (180°) bond angle. The model potential energy surfaces were generated using the torsion parameters listed in Tables 3 and 8 for the partly relaxed torsion scan. (For the lower right panel in Fig. 12, the added angle-bending potential used a greatly exaggerated value of k_{angle} = 2000 kJ mol^{-1} with my new angle-bending potential to make the effect of including an angle-bending potential more visible.) The CCSD potential energy surface was interpolated between data points to make it appear smoother. The model potential energy surfaces were computed in 1°(angle) and 5°(dihedral) increments. Agreement between the CCSD-calculated and ADDT/ADCO model potential energy surfaces was excellent. As shown in Fig. 12, potential energy surfaces omitting the torsion offset potential had incorrect shapes, and this makes them bad models. This

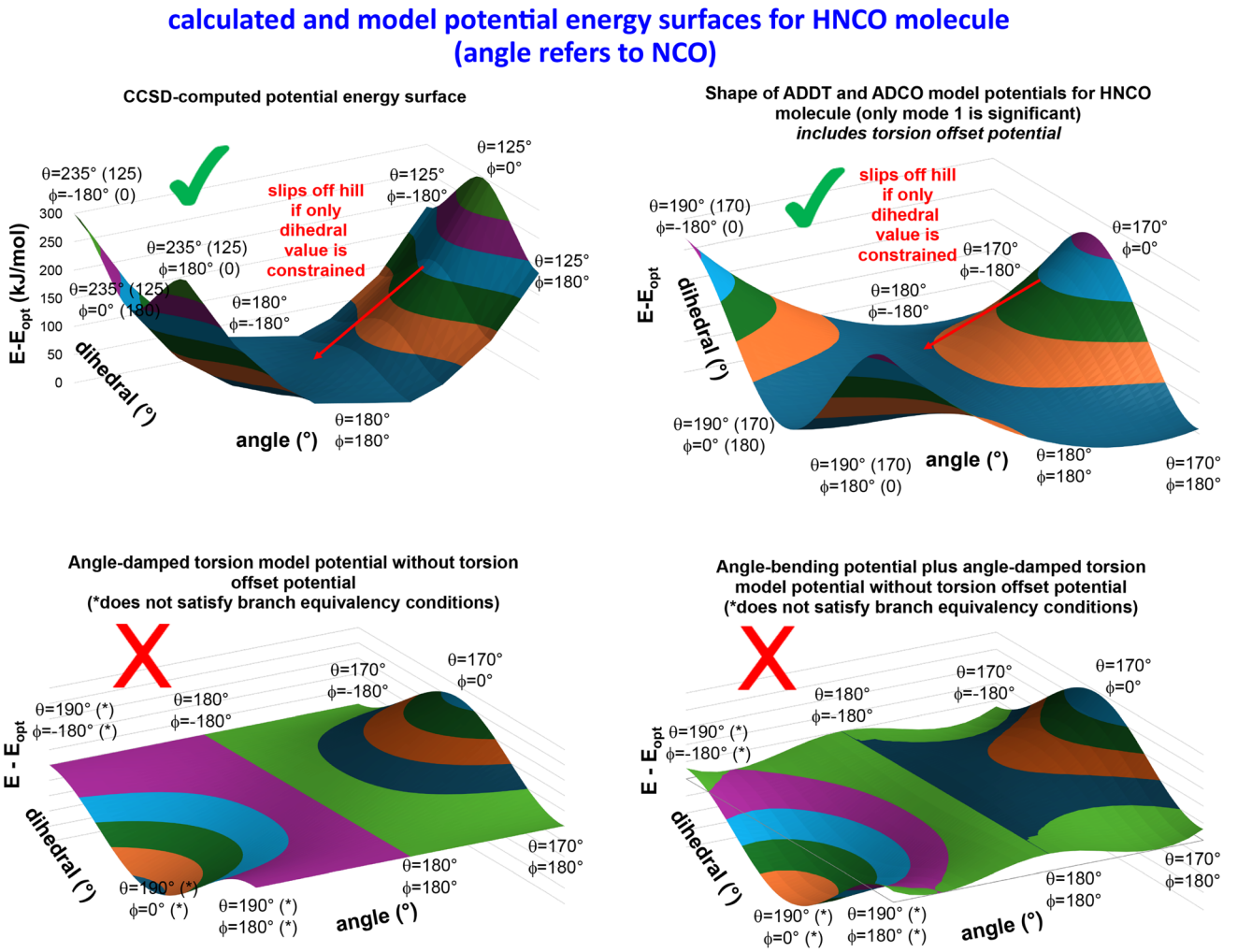

Fig. 12 Calculated and model potential energy surfaces of the HNCO molecule demonstrating that it is absolutely critical to include the torsion offset potential. This potential energy surface also illustrates the 'slip torsion' phenomenon. The CCSD-calculated potential energy surface (upper left panel) matches the ADDT and ADCO model potential energy surfaces (upper right panel) which include the mode 1 term $U^{ADDT}_{mode_1}[\aleph_{ABC}, \aleph_{BCD}, \phi_{ABCD}]$ or $U^{ADCO}_{mode_1}[\aleph_{ABC}, \aleph_{BCD}, \phi_{ABCD}]$ as the model potential. Both the CCSD-calculated and the ADDT and ADCO model potentials satisfy the angle-dihedral coordinate branch equivalency condition. For example, in the upper left panel the point marked (θ, ϕ) = (235°,-180°) is a branch replicate of the point marked (θ, ϕ) = (125°,0°). For comparison, the two lower panels show model potentials that do not include the torsion offset potential. The lower left panel shows the same potential model as the upper right panel, except the lower left panel excludes the torsion offset potential. This omission destroys the angle-dihedral coordinate branch equivalency and makes the model useless, because it is physically inconsistent. As shown in the lower right panel, adding an angle-bending potential does not rectify the physical inconsistency created by omitting the torsion offset potential.

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

conclusion still holds when any conceivable kind of angle-bending potential is included without including the torsion offset potential.

For this molecule, the ADDT and ADCO model potentials happened to coincide with each other for the following reasons. As shown in Table 3, the ADDT projection coefficients c_i are zero for all of the sine modes in this molecule. Since $\phi_{eq} = 180°$ for this molecule, it directly follows that $\cos[m(\phi - \phi_{eq})] = (-1)^m \cos[m\phi]$, and this means the coefficient for cosine mode m in this molecule has the same magnitude but not necessarily the same sign for the ADDT and ADCO model potentials. Examining Tables 3 and 8, mode 1 dominates (*i.e.*, $\text{abs}[c_1] = 1.0000$) for the partly relaxed torsion scan. For mode 2, $\text{abs}[c_2] = 0.0004$ is below the smart selection thresholds for the ADDT and ADCO model potentials. The remaining ADDT and ADCO modes have smaller magnitude coefficients that make them negligible.

In some materials, the angle-damped dihedral torsion (which includes the torsion offset potential) gives rise to a new physical phenomenon called 'slip torsion'. This is illustrated in Fig. 12 for the HNCO molecule. In this molecule, the optimized dihedral value is $\phi_{eq} = 180°$. Consider a geometry optimization in which the HNCO dihedral value is constrained but all other geometric parameters are allowed to relax. When the HNCO dihedral is constrained to a value of ϕ between −90° and +90°, the molecule is situated on the side of a potential energy hill. The molecule's energy can be lowered by sliding down the hill in the direction of increasing bond angle, θ_{NCO}. When the molecule slides down as far as $\theta_{NCO} = 180°$, then the HNCO dihedral value for this non-equilibrium linear bond angle becomes undefined and the constrained geometry calculation crashes (*i.e.*, abruptly terminates due to undefined dihedral value). True to form, this is exactly what happened when attempting to calculate fully relaxed torsion scans for the HNCO, HNCS, HOCN, and HSNC molecules. For these molecules, constraining only the dihedral value produced converged results when $-180° < \phi < -90°$ and $90° < \phi \leq 180°$, while the calculations did not converge when $-90° < \phi < 90°$. As explained in Section 10.2, converged results were achieved by performing partly relaxed dihedral scans in which both bond angles were constrained in addition to constraining the dihedral's value.

10.5 Potential energy surfaces for linear dihedrals

For five molecules with linear dihedrals, CCSD/def2-TZVPD calculations were performed at a series of constrained bond angles and constrained dihedral values to generate potential energy surfaces. These constrained geometry optimizations were performed in Gaussian 16 software using the opt = mod-redundant method. For the calculations used in this subsection, the bond lengths were not constrained. For the three molecules containing single-linear dihedrals, the angle for which $\theta_{eq} \neq \pi$ was not constrained, while the other angle was constrained to 130°, 145°, and 160°. For the two molecules containing double-linear dihedrals, the two bond angles were constrained to all combinations of $\theta_{ABC}, \theta_{BCD} \in \{150°, 160°, 170°\}$. Dihedrals were constrained to a series of symmetry unique values to generate data for the full range $-180° < \phi \leq 180°$ at each constrained bond angle. Finally, the relative energy difference $(E_\mu^{el} - E_{opt}^{el})$ was computed for each geometry μ, where E_{opt}^{el} is the QM-computed electronic energy of the molecule's fully unconstrained ground-state geometry.

The Generalized Reduced Gradient (GRG) solver in Excel was used to optimize the force constants of the ADLD model potential plus the Manz angle-bending potential. These force constants were optimized to maximize the R-squared value (see eqn (17)) using the following definitions

$$\text{SSE} = \sum_{\mu \in \text{ training geoms}} \left(\left(E_\mu^{el} - E_{opt}^{el} \right) - \left(U_\mu^{model} - U_{opt}^{model} \right) \right)^2 \quad (243)$$

$$\text{SST} = \sum_{\mu \in \text{ training geoms}} \left(E_\mu^{el} - E_{opt}^{el} \right)^2 \quad (244)$$

while constraining each optimized force constant to be non-negative. For the single-linear dihedrals, the Manz angle-bending potential was included only for the bond angle that was constrained to a series of different values. For the double-linear dihedrals, the Manz angle-bending potential was included for both bond angles, because these were constrained to a series of different values. Because the unconstrained optimized geometry of each molecule contained mirror planes, the odd-function (*i.e.*, sine modes) in the ADLD model potential have zero contributions and were not included in the fitting

Table 11 Optimized ADLD force constants and *R*-squared values for five molecules having linear dihedrals. The number of unique geometries includes the optimized ground-state geometry plus displaced geometries

Formula	Number unique geometries in training set	Force constants optimized to nonzero values (kJ mol^{-1})	Force constants optimized to zero	*R*-squared
HCCH	61	$k_{LD5}^1 = 322.92$, (both) $k_{CCH}^{Manz_bend} = 105.25$	$k_{LD1}^1 = k_{LD2}^1 = k_{LD4}^1 = 0$	0.9988
HCNO	91	$k_{LD5}^1 = 473.10$, $k_{CNO}^{Manz_bend} = 329.99$	$k_{HCN}^{Manz_bend} = k_{LD1}^1 = k_{LD2}^1 = k_{LD4}^1 = 0$	0.9974
H_3CCN	22	$k_{LD5}^2 = 5.249$, $k_{CCN}^{Manz_bend} = 194.90$	$k_{LD4}^2 = 0$	0.9993
H_2BCN	22	$k_{LD2}^1 = 182.43$, $k_{BCN}^{Manz_bend} = 98.14$	$k_{LD1}^1 = 0$	0.9967
H_2BNC	22	$k_{LD2}^1 = 106.30$, $k_{BNC}^{Manz_bend} = 61.50$	$k_{LD1}^1 = 0$	0.9971

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

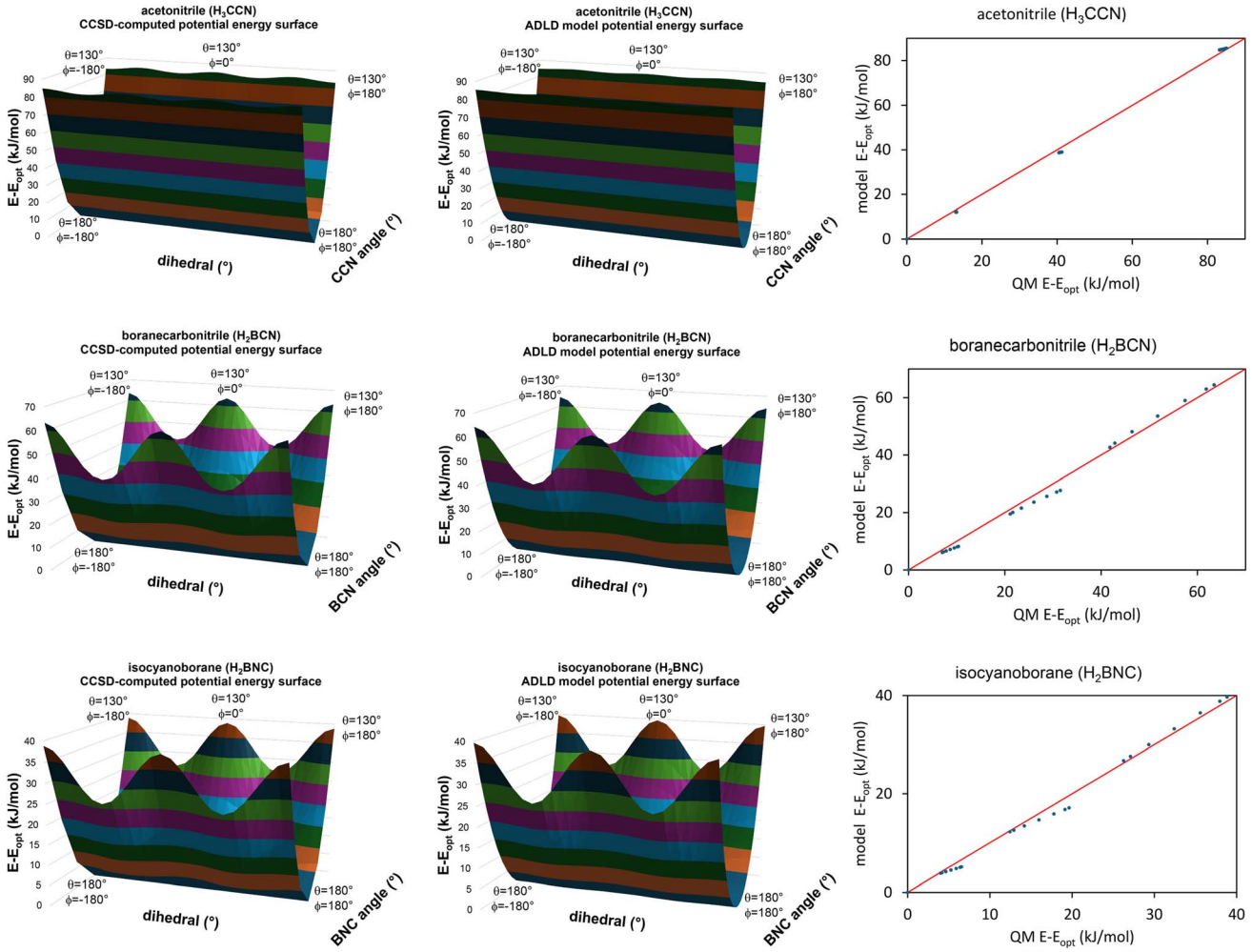

Fig. 13 Comparison of QM-computed CCSD potential energy surfaces (lefthand panels) to ADLD model potential energy surfaces (middle panels) for three molecules having single-linear dihedrals. The righthand panels are parity plots comparing QM-computed CCSD energies to ADLD model potential energies for the computed datapoints for these three molecules. Note: these surface plots used linear interpolation between computed datapoints to create the appearance of smooth surfaces.

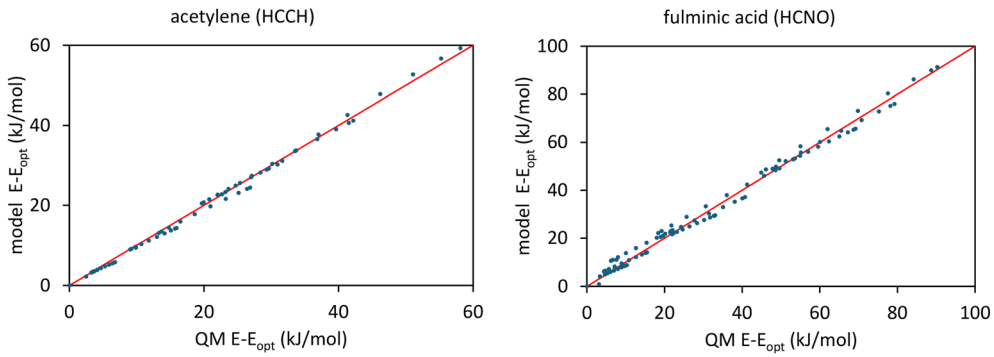

Fig. 14 Parity plots comparing QM-computed CCSD energies to ADLD model potential energies for two molecules having double-linear dihedrals. This figure contains all of the computed datapoints used to generate the contour plots in Fig. 15.

process. For the even-function (*i.e.*, cosine modes), the multiplicities (*i.e.*, n values) were chosen to match the n-fold rotational symmetry of the molecule's unconstrained optimized geometry. For example, H_3CCN has a 3-fold rotational symmetry, while H_2BCN and H_2BNC have 2-fold rotational symmetry. Table 11 summarizes the optimized force constants and R-squared values. The R-squared values were >0.99 indicating superb fits of the ADLD model potential.

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

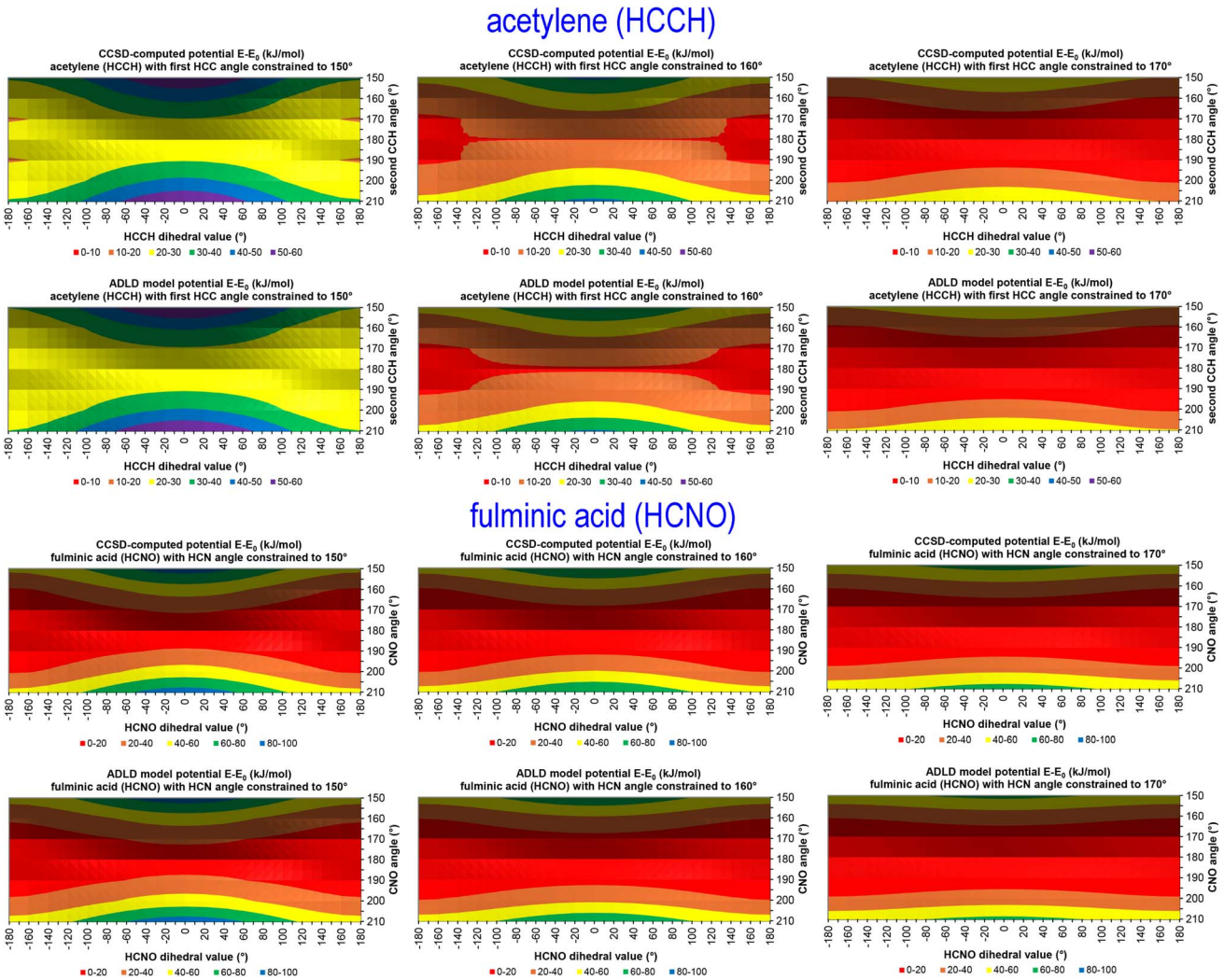

Fig. 15 Comparison of QM-computed CCSD potential energy contour plots (first and third rows) to ADLD model potential energy contour plots (second and fourth rows) for two molecules having double-linear dihedrals. Note: these contour plots use interpolation between computed datapoints.

As shown in Fig. 13, the ADLD model potential energy surfaces closely reproduced the QM-computed CCSD potential energy surfaces for the three molecules with single-linear dihedrals. Fig. 13 also contains parity plots showing excellent agreement between the QM-computed CCSD energies and the ADLD model predicted energies for the computed datapoints. As shown in Fig. 14 and 15, the ADLD model potential closely reproduced the CCSD-computed energies for the two molecules with double-linear dihedrals. Overall, these results showed that my ADLD model potential (when including the Manz angle-bending potential) does an excellent job of describing potential energy surfaces for linear dihedrals.

In Fig. 15, each midline corresponds to one of the bond angles being constrained to 180° while the other bond angle was constrained to 150°, 160°, or 170°. The corresponding CCSD energy for each midline was computed in Gaussian16 (ref. 33) using the keyword opt = (ModRedundant,GIC) where GIC stands for Generalized Internal Coordinates. Since one of the angles was linear (*i.e.*, 180°) along this midline, its dihedral value was undefined. Using GIC allowed the constrained midline geometry during the CCSD calculation to be specified using constrained linear bend internal coordinates plus unconstrained bond lengths and constrained nonlinear bond angle. (These midlines were not included in the training dataset used to optimize the force constants.) For the ADLD model panels in Fig. 15, the energy along the corresponding midline was computed directly from the parameterized ADLD model potential.

10.6 Brief recap of results for metal–organic frameworks containing linear dihedrals

In a companion article, 5 out of 116 MOFs studied contained after-pruning linear dihedrals.[18,19] As shown in Fig. 16, these included: (a) accidental single-linear dihedrals (Case #3) in the MOFs BEPVID, GIWMOP, and XAHROQ, (b) a symmetry-induced single-linear dihedral (Case #1a) in HECQUB, and (c) both a symmetry-induced double-linear dihedral (Case # 2, which locally looks like Case # 2a) and a symmetry-induced

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

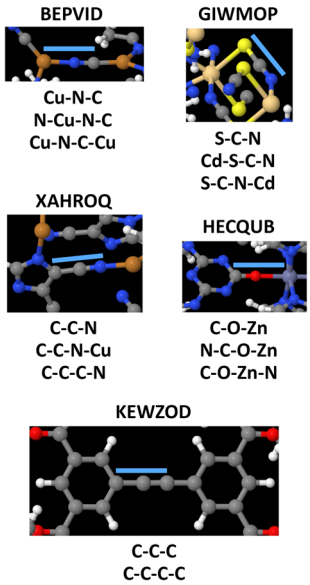

Fig. 16 Some linear dihedrals in MOFs studied in ref. 18 and 19: BEPVID, GIWMOP, HECQUB, KEWZOD, and XAHROQ. In each panel, the blue line is adjacent to the linear bond angle. Below each panel are listed the atomic symbols comprising the linear bond angle and the associated after-pruning linear dihedrals.

single-linear dihedral (Case # 1, which locally looks like Case # 1a) in KEWZOD. In the companion article, the $n = 2j$ but not the $n = (2j - 1)$ modes shown in eqn (223) were derived using part of the theoretical analysis described in Section 9.1 above.[18] The optimized flexibility models for those 5 MOFs included flexibility terms for the single-linear dihedrals and used the modes corresponding to the force constants k^1_{LD1} and k^1_{LD2} in eqn (223).[18,19] For KEWZOD, the single-linear dihedral was included in the parameterized flexibility model, but the double-linear dihedral was not included in the parameterized flexibility model.

A linear dihedral has no torsion barrier at its equilibrium bond angle. For small deviations of the bond angle from its equilibrium value, the torsion barrier will be small in magnitude. Consequently, it is often not required to include torsion potentials for the linear dihedrals when constructing classical forcefields for complex materials. For the 5 MOFs having after-pruning linear dihedrals, flexibility models parameterized with and without including torsion potentials for the linear dihedrals produced essentially the same validation *R*-squared values.[18,19] These *R*-squared values are listed in Table 12.

10.7 Torsions in chiral chain-like molecules

The rigid (or relaxed) torsion scans used in ADDT, CADT, ADCO, or CACO torsion mode smart selection (see Section 7.2 above and ref. 18) are intended to be used when the rotatable group on one side of the rotatable middle bond is small enough that it does not sterically collide with the surrounding structure when rigidly rotated. Typical examples of small rotatable groups include methyl ($-CH_3$), hydroxyl (–OH), amino ($-NH_2$), nitro ($-NO_2$), carboxyl (–C(O)OH), carboxylate ($-CO_2^-$), *etc.* substituent groups.

Sometimes rotatable dihedrals are contained in long chains comprising polymers, large biomolecules (*e.g.*, proteins, DNA, RNA, enzymes, fatty acids, polysaccharides, phospholipids, *etc.*), hydrocarbon chains (*e.g.*, petroleum), *etc.* A chain-like molecule often has many different conformers that are local ground-state structures in which all atom-in-material forces are zero. (Each conformer is a local energy minimum on the molecule's potential energy surface.) For example, a specific protein chain can often fold in different ways to form many different conformers. For these chain-like molecules, a rigid torsion scan of the entire chain may not be appropriate, because it may cause one part of the molecule's chain to sterically collide with another part of the molecule's chain. In these cases, a different strategy besides torsion scans of the full chain should be used to construct the training dataset for optimizing the chain's torsion force constants.

Two basic strategies could be suggested for optimizing the torsion force constants of chain-like molecules. The first strategy uses short pieces of the chain to separately optimize the torsion force constants for each type of rotatable middle bond. Each snippet of the chain should be capped with appropriate atoms; for example, a cut C–C single bond could be turned into a C–H termination. This strategy uses model molecules that are small enough so that rigid torsion scans can be performed without encountering any steric clashes of the rotating group. The reference energy for each torsion scan geometry would be computed using a high-level quantum chemistry method. If the molecule is small enough, a QM-computed relaxed torsion scan could be performed (instead of a rigid torsion scan) without causing large changes (*i.e.*, large relaxations) in the unconstrained internal coordinate values (*i.e.*, bond lengths, angle values, and unconstrained other dihedral values). This rigid (or

Table 12 Comparison of validation R-squared values for 5 MOFs containing after-pruning linear dihedrals. Values are listed with using individual [average] equilibrium values of internal coordinates. This data is taken from calculations (after pruning and without bond–bond cross terms) described in ref. 19. For KEWZOD, the torsion potential for linear dihedrals included the single-linear dihedral but not the double-linear dihedral

MOF refcode	Validation *R*-squared with torsion potential for linear dihedrals	Validation *R*-squared using no torsion potential for linear dihedrals
BEPVID	0.9392 [0.9356]	0.9392 [0.9356]
GIWMOP	0.9209 [0.9044]	0.9210 [0.9044]
HECQUB	0.9002 [0.9002]	0.9002 [0.9002]
KEWZOD	0.9208 [0.9194]	0.9208 [0.9194]
XAHROQ	0.9385 [0.9377]	0.9386 [0.9377]

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

relaxed) torsion scan allows smart mode selection (see Section 7.2) to be used successfully. The model molecule should be large enough to capture the relevant chemical environment around the corresponding target middle bond.

The second strategy employs the entire full chain without invoking model molecules. One could generate a training set containing many chain configurations. For example, molecular dynamics or Monte Carlo simulations could be used to generate an ensemble containing many chain configurations. The energy of each chain configuration would then be computed using a high-level quantum chemistry method. These quantum-mechanically-computed energies would then be used in the training set used to optimize the ADDT, CADT, ADCO, CACO, and/or ADLD torsion force constants. Symmetry properties could be used to eliminate some of the unimportant torsion modes prior to force-constant optimization. For example, the 3-fold rotation symmetry of methyl groups implies corresponding torsion modes of the form $\cos[3j(\phi - \phi_{eq})]$ and $\sin[3j(\phi - \phi_{eq})]$ for ADDT or CADT model potentials or $\cos[3j\phi]$ for ADCO or CACO model potentials, where j is a whole number. Similarly, the 2-fold rotation symmetry of amino ($-NH_2$), nitro ($-NO_2$), and carboxylate ($-CO_2^-$) groups implies corresponding torsion modes of the form $\cos[2j(\phi - \phi_{eq})]$ and $\sin[2j(\phi - \phi_{eq})]$ for ADDT or CADT model potentials or $\cos[2j\phi]$ for ADCO or CACO model potentials, where j is a whole number.

Most biomolecules contain many chiral centers. Chiral centers lack mirror-image (reflection) symmetry.

Fig. 11 shows a dramatic failure of the cosine-only model potential for the relaxed and rigid torsion scan curves of the FCOH dihedral in the *S* enantiomer of the chiral molecule C(OH)ClFH. As shown in Fig. 8, my CADT and ADDT model potentials gave *R*-squared values of 0.9995 (relaxed torsion scan) and 0.9999 (rigid torsion scan) for this same dihedral.

As a more complex example, here we study the 2*S*-2-amino-propanal molecule. This molecule contains a chiral center. '2*S*' indicates the specific enantiomer studied here. (The 2 in '2*S*' indicates the chain position of the S chiral center.) Although this molecule only contains three rotatable middle bonds, it is still large enough to form many different conformers. The various conformers shown in Fig. 17 were generated by the following procedure. First, the three rotatable dihedrals (*i.e.*, OCCN, HNCC, and HCCH) were systematically rotated to a series of different values to generate a set of 24 starting structures. Beginning with each starting structure, geometry optimization was then performed in Gaussian 16 to compute equilibrium structures in which all atom-in-material forces are zero. I computed these using the B3LYP + D3BJ/def2-TZVPD (ref. 34 and 36–39) level of theory. Next, a frequency calculation was performed on each distinct equilibrium structure to identify whether it is a local ground state (*i.e.*, local energy minimum) or a saddle point (*e.g.*, transition state). Fig. 17 shows the resulting conformers that had all real-valued non-negative frequencies.

For comparison, Fig. 17 also shows the lowest energy conformer of the 2*R* enantiomer. The 2*R* enantiomer has the same number of conformers as the 2*S* enantiomer, such that each 2*R* conformer is a mirror image of and has the same energy as the corresponding 2*S* conformer.

As shown in Fig. 18, the CADT model potential accurately reproduces the distinct mirror-image torsion scan curves for the *R* and *S* enantiomers. As described in Sections 3–5 above, both the ADDT and CADT model potentials achieve this in a way that allows mirror-image dihedral instances to be included within the same dihedral type. When using the ADDT and CADT model potentials, both the *R* and *S* enantiomers are described by the same dihedral types and torsion force constant values. In stark contrast, the cosine-only torsion model potential (*i.e.*, eqn (202) and (150)) cannot provide distinct mirror-image torsion scan curves for the *R* and *S* enantiomers. Depending on the specific situation, this limitation of the cosine-only torsion model potential (*i.e.*, eqn (202) and (150)) may result in any of the following: (a) negligible change in the *R*-squared value (*e.g.*,

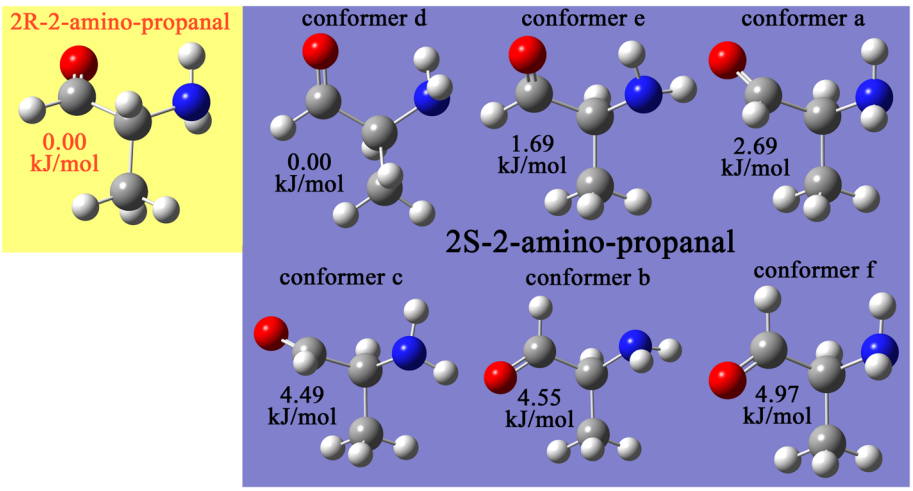

Fig. 17 Various conformers of the 2S-2-amino-propanal molecule as computed using the B3LYP + D3BJ/def2-TZVPD level of theory. The computed relative energy (in kJ mol^{-1}) is listed next to each conformer. For comparison, the lowest energy conformer of the 2*R* enantiomer is also shown.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

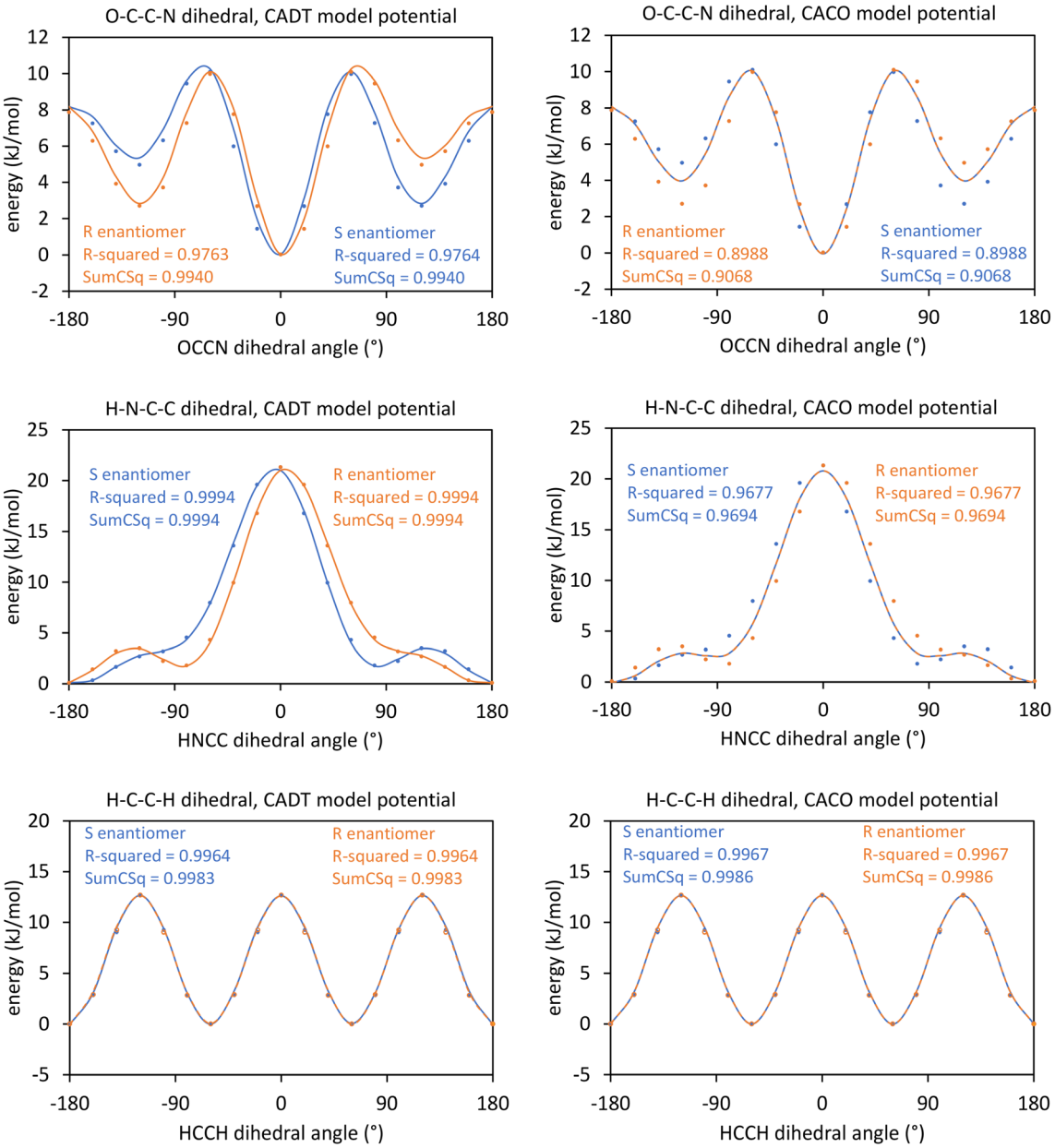

Fig. 18 Better performance of the CADT model potential (left panels) compared to the CACO torsion model potential (right panels). Results are shown here for three rotatable dihedrals in the *R* (orange points and orange curves) and S (blue points and blue curves) enantiomers of the 2-amino-propanal molecule. The filled circles show the QM-computed (B3LYP + D3BJ/def2-TZVPD) energies for the dihedral scan curve in which all geometric parameters were relaxed except the constrained dihedral. The solid lines show the fitted model potentials including CADT modes 1–7 (left panels) and CACO modes 1–4 (right panels).

rotation of a methyl group as shown in the lower panels of Fig. 18), (b) a moderate reduction in the *R*-squared value (*e.g.*, rotation of the HNCC or OCCN dihedrals in 2-amino-propanal as shown in the middle and top panels of Fig. 18), or (c) a huge catastrophic reduction in the *R*-squared value (*e.g.*, rotation of the FCOH dihedral in the C(OH)ClFH molecule as shown in Fig. 11).

The computed *sym*_value (see eqn (159)) for these three dihedrals was 0.299 for the OCCN dihedral, 0.174 for the HNCC dihedral, and 0.014 for the HCCH dihedral. Using these *sym*_values, the equilibrium bond angle values, and the flowchart shown in Fig. 6, the CADT model potential is recommended for all three of these dihedrals. For the CADT model, the following modes were smart selected using the cutoff values recommended in Section 7.2: modes 3, 5, and 7 for the HCCH dihedral; modes 1, 2, 3, 5, 6, and 7 for the HNCC dihedral, and modes 1, 2, 3, 4, 5, and 6 for the OCCN dihedral. For the CACO model, the following modes were smart selected using the cutoff value of 0.001 recommended in Section 7.2: mode 3 for the HCCH dihedral; modes 1, 2, 3, and 4 for the HNCC dihedral, and modes 1, 2, 3, and 4 for the OCCN dihedral.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

For this molecule, different sets of torsion force constants were optimized by fitting models to a training dataset. The training dataset contained the following QM-computed (B3LYP + D3BJ/def2-TZVPD) geometries and energies for the *S* enantiomer: (a) the six conformers shown in Fig. 17, (b) the symmetry-unique geometries for the relaxed torsion scan curves shown in Fig. 18 (*i.e.*, 18 geometries for OCCN dihedral, 18 geometries for HNCC dihedral, and 6 geometries for HCCH dihedral), and (c) 30 geometries in which the 3 dihedral values were chosen using a uniform random number generator. For (c), the geometries were relaxed in Gaussian 16 (using the B3LYP + D3BJ/def2-TZVPD level of theory) keeping the three randomized dihedral values constrained but allowing the bond lengths and angles to relax. This training dataset contained a total of 78 geometries.

Torsion force constants were optimized in Matlab by minimizing the following loss function:

$$L = \sum_{\mu \in \text{ training geoms}} \left(\left(E_\mu^{\text{el}} - E_{\text{opt}}^{\text{el}} \right) - \left(U_\mu^{\text{FF}} - U_{\text{opt}}^{\text{FF}} \right) \right)^2 \quad (245)$$

$$U_\mu^{\text{FF}} = U_\mu^{\text{bonded}} + U_\mu^{\text{nonbonded}} \quad (246)$$

$$U_\mu^{\text{nonbonded}} = U_\mu^{\text{intracluster_nonbonded}} + U_\mu^{\text{intercluster_nonbonded}} \quad (247)$$

where μ is the geometry number in the training dataset. E_μ^{el} and E_{opt}^{el} are the QM-computed electronic energies of training geometry μ and the fully optimized ground-state geometry, respectively. U_μ^{FF} and $U_{\text{opt}}^{\text{FF}}$ are the forcefield model's potential energies of training geometry μ and the fully optimized ground-state geometry, respectively. Since the forcefield was trained on an isolated bonded cluster (*i.e.*, an isolated molecule), there were no intercluster interactions in the training dataset geometries: $U_\mu^{\text{intercluster_nonbonded}} = 0$.

Six variations of the forcefield model were constructed from two different torsion models and three different nonbonded interaction models: 2 (torsion models) × 3 (nonbonded interaction models) = 6 forcefield models. The two torsion models were: (a) the CADT model potential with smart selected modes and (b) the CACO torsion model potential using smart selected modes. Since the bond angles and bond lengths were relaxed in all training and validation geometries, dihedral torsions were the only bonded interactions included in the six forcefield models. This was chosen to directly compare the CADT model to the CACO torsion model.

The three nonbonded interaction models were: (i) no nonbonded interactions (*i.e.*, all nonbonded interactions set to zero), (ii) electrostatic and Lennard-Jones interactions between all atoms except first-, second-, and third-neighbors (*i.e.*, excluding self, 1-2, 1-3, and 1-4 nonbonded interactions), and (iv) electrostatic and Lennard-Jones interactions between all atoms except first- and second-neighbors (*i.e.*, excluding self, 1-2, and 1-3 nonbonded interactions). Nonbonded electrostatic interactions were modeled with an atom-centered point charge model computed using the DDEC6 (ref. 40 and 41) method. Lennard-Jones parameters and combining rules were taken from the Universal Force Field (UFF).[42] These were used as inputs for Manz's interaction separation ansatz following the case in which a cutoff distance is not used for the nonbonded interactions.[20] Following the notation of ref. 20, this gives:

$$U_\mu^{\text{intracluster_nonbonded}} = \sum_{A \in \text{ cluster_j}} \sum_{B \in (\text{cluster_j} - \{\text{excluded}_A\})} \Phi_{ABx}^{\text{intracluster}} \quad (248)$$

$$\Phi_{ABx}^{\text{intracluster}} = \tanh^2\left[\frac{d_{AB}^{\text{eq_j}}}{d_{AB}} - \frac{d_{AB}}{d_{AB}^{\text{eq_j}}}\right]\left(U_{ABx,\text{intracluster}}^{\text{nonbonded}}\left[\left\{\vec{R}_C\right\}\right] - U_{ABx,\text{intracluster}}^{\text{nonbonded}}\left[\left\{\vec{R}_C^{\text{eq_j}}\right\}\right]\right) \quad (249)$$

$$U_{ABx,\text{intracluster}}^{\text{nonbonded}}\left[\left\{\vec{R}_C\right\}\right] \approx U_{AB}^{(q+\text{LJ})} \quad (250)$$

$$U_{AB}^{(q+\text{LJ})} = \overbrace{\frac{q_A q_B}{4\pi\varepsilon_0 d_{AB}}}^{\text{atomic charges}} + \underbrace{\varepsilon_{AB}^{LJ}\left(\left(\frac{d_{AB}^{\text{LJ}}}{d_{AB}}\right)^{12} - 2\left(\frac{d_{AB}^{\text{LJ}}}{d_{AB}}\right)^{6}\right)}_{\substack{\text{repulsion+dispersion} \\ (e.g.,\ \text{Lennard-Jones})}} \quad (251)$$

where $\vec{R}_C$ is the nuclear position of atom C, cluster_j is the specific molecule, $d_{AB}^{\text{eq_j}}$ is the equilibrium distance between atoms A and B in the optimized ground state of isolated cluster_*j*, {excluded$_A$} is the set of excluded nonbonded interactions (*e.g.*, self, 1–2, 1–3, and optionally 1–4 interactions), $\varepsilon_{AB}^{\text{LJ}}$ is the Lennard-Jones well-depth, and d_{AB}^{LJ} is the Lennard-Jones reference distance.

The validation dataset contained 30 geometries in which the 3 dihedral values were chosen using a uniform random number generator. These geometries were relaxed in Gaussian 16 (using the B3LYP + D3BJ/def2-TZVPD level of theory) keeping the three randomized dihedral values constrained but allowing the bond lengths and angles to relax. These 30 validation geometries were generated independently (*i.e.*, using different random numbers) from the 30 randomized geometries used in the training dataset.

R-squared values were computed for the training and validation datasets using eqn (17) with the following definitions from ref. 20:

$$\text{SST} = \sum_\mu \left(E_\mu^{\text{el}} - E_{\text{opt}}^{\text{el}} \right)^2 \quad (252)$$

$$\text{SSE} = \sum_\mu \left(\left(E_\mu^{\text{el}} - E_{\text{opt}}^{\text{el}} \right) - \left(U_\mu^{\text{FF}} - U_{\text{opt}}^{\text{FF}} \right) \right)^2 \quad (253)$$

Minimizing the loss function in eqn (245) gives the same optimized force constant values as maximizing *R*-squared training. The root-mean-squared-error (RMSE) was defined as

$$\text{RMSE} = \sqrt{\text{SSE}/N_{\text{obs}}} \quad (254)$$

where N_{obs} is the total number of observation datapoints (*i.e.*, number of '*y* values') in the (training or validation) dataset used to compute SSE. (In cases where the training or validation dataset contains atom-in-material forces and/or total energies, a total energy or atom-in-material force component counts as a '*y* value'.[18])

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 13 Performance of six different forcefield models for describing dihedral torsions in the chiral 2-amino-propanal molecule

Torsion model	Nonbonded interactions included	*R*-squared training	RMSE training (eV)	*R*-squared validation	RMSE validation (eV)
CADT	None	(0.9424)[a]	(0.0326)[a]	(0.9227)[a]	(0.0651)[a]
CADT	Except 1-2, 1-3, 1-4	0.9317	0.0355	0.9444	0.0552
CADT	Except 1–2, 1–3	0.9259	0.0370	0.9427	0.0560
CACO	None	(0.9034)[a]	(0.0423)[a]	(0.9451)[a]	(0.0548)[a]
CACO	Except 1–2, 1–3, 1–4	0.7803	0.0637	0.8806	0.0809
CACO	Except 1–2, 1–3	0.8542	0.0519	0.9170	0.0674

[a] The results with all nonbonded interactions excluded are listed here for comparison purposes only. Since this molecule contains several 1–5 pairs (*i.e.*, fourth neighbors), it would actually not be appropriate to use a forcefield for this molecule that neglected all nonbonded interactions. Please see the main text for a discussion of this issue.

Tables S3 and S4 of the ESI† list the optimized torsion force constant values. The Matlab codes and output results are found in the supporting data zip archive of the ESI.† As shown in Table 13, the CADT model potential performed better overall than the CACO torsion model potential for this molecule.

Although the two forcefield models that omitted all nonbonded interactions also appeared to perform well for this small molecule, the following point should be kept in mind. For long flexible chains, including nonbonded interactions is mandatory during classical molecular dynamics and Monte Carlo simulations. If all nonbonded interactions are excluded during such atomistic simulations, then different parts of a long flexible chain could occupy the same spatial local volume. This behavior would be physically unreasonable, because it violates the Pauli exclusion principle. To prevent such unphysical behavior, the forcefield must include short-range repulsion that causes nonbonded atoms to repel each other at short distances. For example, the Lennard-Jones potential includes a short-range repulsion that prevents two non-bonded atoms from occupying the same spatial local volume. Examining Table 13, for the forcefields including some nonbonded interactions, the CADT model performed significantly better than the CACO model for this molecule.

The same force constant values, *R*-squared values, and RMSE values would have resulted if some or all of the geometries in either the training or validation datasets were swapped for their mirror images. (The ADDT, ADCO, CADT, CACO, and ADLD model potentials are constructed to achieve this property.) Accordingly, changing *S* to *R* enantiomers by taking mirror image geometries would not have altered any of the results. Of course, for those mirror-image geometries, one is required to use the correct dihedral equilibrium value (ϕ_{eq}) for each corresponding dihedral instance in that enantiomer. Taking a mirror image (*e.g.*, changing *S* to *R* enantiomer) has the following effect:

$$\phi^{eq}_{ABCD}[R\text{ enantiomer}] = \pi - \mathrm{mod}[(\phi^{eq}_{ABCD}[S\text{ enantiomer}] + \pi), 2\pi] \quad (255)$$

$$\phi_{ABCD}[R\text{ enantiomer}] = \pi - \mathrm{mod}[(\phi_{ABCD}[S\text{ enantiomer}] + \pi), 2\pi] \quad (256)$$

In eqn (255) and (256), the mod function ensures that $-\pi < \phi_{ABCD}[R\text{ enantiomer}] \leq \pi$.

10.8 Comparison of vibrational frequencies computed from forcefield models to experimental data

Forcefields were optimized for the isocyanic acid (HNCO), hydrogen peroxide (HOOH), and acetylene (HCCH) molecules. These molecules were chosen, because they allow each of the five dihedral torsion model potentials (*i.e.*, ADCO, ADDT, CACO, CADT, and ADLD) to be tested, and the computed vibrational frequencies were compared to experimental data. No cross terms were included in these forcefields, because computational tests showed they were not required to achieve good performance. The small size of these molecules greatly simplified the forcefield parameterization and testing. Since these molecules contained no further than third-neighbors, all intracluster nonbonded interactions were excluded.

For simplicity, the harmonic bond stretch potential was used:

$$U^{harmonic_stretch}_{AB} = \tfrac{1}{2}k_{AB}(d_{AB} - d^{eq}_{AB})^2 \quad (257)$$

Prior work showed the Manz stretch potential better models bond stretch anharmonicity.[20] However, this involves an additional non-empirical parameter (*i.e.*, the quantum-mechanically-computed exponent γ°_{AB}).[20] Since the normal mode frequencies were computed here within the harmonic oscillator approximation which neglects bond stretch anharmonicities, the simpler harmonic bond stretch potential was used in these optimized forcefields.

Vibrational frequencies were computed using the procedure described in a prior publication: "For each flexibility model, normal vibrational mode analysis within the harmonic oscillator approximation was performed by diagonalizing the mass-weighted Hessian (MWH) matrix expressed in Cartesian coordinates:

$$\mathrm{MWH}_{(3(A-1)+i),(3(B-1)+j)} = \frac{1}{\sqrt{m_A m_B}} \frac{\partial^2 U}{\partial\left(\vec{R}_A\right)_i \partial\left(\vec{R}_B\right)_j} \quad (258)$$

where m_A is the mass of atom A. Here, $(\vec{R}_A)_i$ for $i \in \{1,2,3\}$ denotes the *X*, *Y*, or *Z* component of the nuclear position $\vec{R}_A$. The second derivatives can be computed either analytically or

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

numerically; here, they were computed numerically using the central finite difference approximation. The eigenvalues $\{\lambda_i\}$ of the MWH matrix are related to the normal mode frequencies $\{\text{freq}_i\}$ *via*:[43]

$$\text{freq}_i = \sqrt{\lambda_i}\Big/(2\pi) \tag{259}$$

Each normal mode frequency was converted to wavenumber by dividing by the speed of light, c. Each eigenvector $(\vec{V}^{\text{normal_mode}})$ of the MWH matrix is the corresponding normal mode's mass-weighted differential displacement vector:

$$\varepsilon\vec{V}^{\text{normal_mode}} = \sum_{\text{A}}\left(\text{d}\vec{R}_{\text{A}}^{\text{normal_mode}}\sqrt{m_{\text{A}}}\right) \tag{260}$$

for infinitesimal $|\varepsilon|$.

For linear molecules, five of the MWH eigenvalues are zero; these correspond to molecular rotation (2 modes) and center-of-mass translation (3 modes). For nonlinear molecules, six of the MWH eigenvalues are zero; these correspond to molecular rotation (3 modes) and center-of-mass translation (3 modes)."[20]

For HNCO, the training dataset included 145 geometries: (a) the optimized ground-state geometry (1 geometry), (b) a rigid torsion scan in 20° increments over the range $180° < \phi \leq 180°$ (17 geometries outside the optimized geometry), (c) a partly relaxed torsion scan (in which bond angles were constrained but bond lengths were relaxed) in 20° increments over the range $180° < \phi \leq 180°$ (17 geometries outside the optimized geometry), (d) all possible combinations of bond lengths being displaced by −0.05, 0.00, and/or +0.05 Å (26 geometries outside the optimized geometry), (e) a rigid HNC angle scan in 5° increments over the range $30° \leq \Delta\theta_{\text{HNC}} \leq 30°$ (12 geometries outside the optimized geometry), (f) a relaxed HNC angle scan in 5° increments over the range $30° \leq \Delta\theta_{\text{HNC}} \leq 30°$ (12 geometries outside the optimized geometry), (g) a rigid NCO angle scan in 2° increments over the range $150° \leq \theta_{\text{NCO}} \leq 180°$ (16 geometries), (h) a relaxed NCO angle scan in 2° increments over the range $150° \leq \theta_{\text{NCO}} \leq 164°$ (8 geometries), (i) a torsion scan in 20° increments over the range $180° < \phi \leq 180°$ in which θ_{NCO} was constrained to 155° while all other geometric parameters were relaxed (18 geometries), and (j) a torsion scan in 20° increments over the range $180° < \phi \leq 180°$ in which θ_{NCO} was constrained to 165° while all other geometric parameters were relaxed (18 geometries).

In this subsection, the force constants were optimized using least-squares fitting in Matlab to minimize the cost function shown in eqn (245), where the forcefield contained no intra-cluster nonbonded interactions as explained above. The training R-squared and RMSE values were computed using eqn (17) and (252)–(254).

Table 14 summarizes the optimized force constants and training dataset statistics for the HNCO molecule using forcefields with ADCO or ADDT dihedral torsion model potentials. A bend potential for the NCO angle is not needed and was not used, because this angle scan curve is completely described by the HNCO torsion offset potential as clearly shown in Fig. 9. For this molecule, the QM-computed dihedral torsion potential energy surface contains no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$). For ADCO and ADDT, the smart mode selection criteria described in Section 7 above were used. As shown in Table 14, the forcefields using the ADCO and ADDT model potentials performed extremely well and yielded nearly identical results for this molecule.

Table 15 summarizes the validation dataset results for this molecule. The vibrational frequencies computed using each forcefield were in good agreement with the experimentally-measured frequencies.

For HOOH, the training dataset included 132 geometries: (a) the optimized ground-state geometry and its mirror image (2 geometries), (b) a rigid torsion scan in 20° increments over the range $180° < \phi \leq 180°$ (18 geometries), (c) a relaxed torsion scan in 20° increments over the range $180° < \phi \leq 180°$ (18 geometries), (d) a torsion scan in 20° increments over the range $180° < \phi \leq 180°$ in which $\Delta\theta_{\text{HOO}}$ was constrained to −10° while all other geometric parameters were relaxed (18 geometries), (e) a torsion scan in 20° increments over the range $180° < \phi \leq 180°$ in which $\Delta\theta_{\text{HOO}}$ was constrained to +10° while all other geometric parameters were relaxed (18 geometries), (f) all symmetry-unique combinations of bond lengths being displaced by −0.14, 0.00, and/or +0.14 Å (17 geometries outside the optimized geometry), (g) mirror images of the structures from (*f*) (17 geometries), (h) a rigid HOO angle scan in 10° increments over the range $80° \leq \theta_{\text{NCO}} \leq 130°$ (6 geometries), (i) a relaxed HOO angle scan in 10° increments over the range $80° \leq \theta_{\text{NCO}} \leq 130°$ (6 geometries), and (j) mirror images of the structures from (*h*) and (i) (12 geometries).

Table 16 summarizes the optimized force constants and training dataset statistics for the HOOH molecule using forcefields with CACO or CADT dihedral torsion model potentials. For this molecule, the QM-computed dihedral torsion potential

Table 14 Training dataset performance for isocyanic acid (HNCO) forcefields using harmonic bond stretch, Manz angle-bend, and ADCO or ADDT dihedral torsion model potentials. The QM-computed (CCSD/def2-TZVPD) equilibrium values were 1.00578 Å (HN), 1.21144 Å (NC), 1.16029 Å (CO), 123.57915° (HNC), 172.98777° (NCO), and 180.000° (HNCO)

	k_{stretch} (eV Å^{-2})	k_{bend} (eV)	k_{torsion} (eV)	Training RMSE (eV)	Training R-squared	ϕ_{opt} (°)
Forcefield *w.* ADCO	45.10 (HN), 56.60 (NC), 104.46 (CO)	1.696 (HNC)	0.05818, (c_i = [0.999999, 0.001117, 0.000000, 0.000000])	0.0283	0.9880	180.000°
Forcefield *w.* ADDT	45.10 (HN), 56.61 (NC), 104.46 (CO)	1.696 (HNC)	0.05817 (mode 1)	0.0283	0.9881	180.000°

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 15 Validation dataset performance for isocyanic acid (HNCO) forcefields using harmonic bond stretch, Manz angle-bend, and ADCO or ADDT dihedral torsion model potentials. This validation dataset compares the normal-mode vibrational frequencies computed from the forcefield to experimentally-measured frequencies (in wavenumber, cm^{-1}). The percent error relative to experiment is shown in parentheses. Modes 4 and 6 are two different in-plane bends, and each of these modes involves changes in both the HNC and CNO angles

freq →	1 (HN stretch)	2 (asymmetric NCO stretch)	3 (symmetric NCO stretch)	4 (in-plane bend)	5 (torsion)	6 (in-plane bend)
Experiment (ref. 44 and 45)	3538	2269	1322	777	656	577
Forcefield *w.* ADCO	3616 (+2%)	2279 (0%)	1160 (−12%)	789 (+2%)	574 (−13%)	542 (−6%)
Forcefield *w.* ADDT	3616 (+2%)	2279 (0%)	1161 (−12%)	789 (+2%)	574 (−13%)	542 (−6%)

Table 16 Training dataset performance for hydrogen peroxide (HOOH) forcefields using harmonic bond stretch, Manz angle-bend, and CACO or CADT dihedral torsion model potentials. The QM-computed (CCSD/def2-TZVPD) equilibrium values were 0.9666 Å (HO), 1.4378 Å (OO), 100.8215° (HOO), and 111.0568° (HOOH)

	$k_{stretch}$ (eV Å^{-2})	k_{bend} (eV)	$k_{torsion}$ (eV)	Training RMSE (eV)	Training *R*-squared	ϕ_{opt} (°)
Forcefield *w.* CACO	50.71 (HO), 31.81 (OO)	5.683 (HOO)	0.1552, (c_i = [0.833919, 0.549530, 0.050017, 0.009368])	0.1194	0.9530	±111.47
Forcefield *w.* CADT	50.96 (HO), 31.85 (OO)	5.760 (HOO)	0.04466 (mode 1), 0.06460 (mode 2), −0.00888 (mode 3), −0.11657 (mode 5), 0.04935 (mode 6), −0.05040 (mode 7)	0.1192	0.9531	−104.71, +111.0568

energy surface contains no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$). For CACO and CADT, the smart mode selection criteria described in Section 7 above were used. Except for one key distinction, the forcefields using CACO and CADT model potentials yielded similar results. The key distinction is that CACO strictly yielded $U[\phi] = U[-\phi]$ giving $\phi_{opt} = \pm 111.47°$ which is close to but not identical to the QM-computed value of ±111.0568°. On the other hand, CADT reproduced ϕ_{opt} = +111.0568° on the positive side but yielded a slightly asymmetric potential $U[\phi] \cong U[-\phi]$ so that on the negative side ϕ_{opt} was −104.71° instead of −111.0568°. Both forcefields yielded training R-squared > 0.95.

Table 17 summarizes the validation dataset results for this molecule. The vibrational frequencies computed using each forcefield were in good agreement with the experimentally-measured frequencies. For this molecule, the experimental reference frequency for each vibrational mode was taken as the median of the experimental values compiled from various sources as listed in the NIST Chemistry Webbook.[46] Please see the ESI† for calculation of these median experimental values.

For HCCH, the training dataset included 78 geometries: (a) the optimized ground-state geometry (1 geometry), (b) all symmetry-unique combinations of bond lengths being displaced by −0.14, 0.00, and/or +0.14 Å (17 geometries outside the optimized geometry), (c) torsion scans in 20° increments over the range $0 \leq \phi \leq 180°$ in which each HCC angle was constrained to 150, 160, or 170° while all other geometric parameters were relaxed (60 geometries). For (c), the range $-180° < \phi < 0$ was not required, because due to symmetry $U[\phi] = U[-\phi]$. For (c), due to the molecular symmetry, there were six pairs of symmetry-unique constrained bond angles: (θ_1, θ_2) = (150°, 150°), (160°, 150°), (170°, 150°), (160°, 160°), (170°, 160°), and (170°, 170°).

Table 18 summarizes the optimized force constants and training dataset statistics for the acetylene molecule using a forcefield with the ADLD model potential. Since this molecule contains a double-linear dihedral, the $n = 2(j = 1) - 1 = 1$ modes form the dominant contribution. Due to the mirror plane, the sine torsion modes do not contribute, and thus $k^1_{LD3} = k^1_{LD6} = 0$. As shown in Table 11, for this molecule k^1_{LD5} >

Table 17 Validation dataset performance for hydrogen peroxide (HOOH) forcefields using harmonic bond stretch, Manz angle-bend, and CACO or CADT dihedral torsion model potentials. This validation dataset compares the normal-mode vibrational frequencies computed from the forcefield to experimentally-measured frequencies (in wavenumber, cm^{-1}). The percent error relative to experiment is shown in parentheses

freq →	1 (asymmetric OH stretch)	2 (symmetric OH stretch)	3 (symmetric HOO bend)	4 (asymmetric HOO bend)	5 (OO stretch)	6 (torsion)
Experiment	3613	3587	1389	1271	869	371
Forcefield *w.* CACO	3815 (+6%)	3814 (+6%)	1421 (+2%)	1349 (+6%)	981 (+13%)	388 (+5%)
Forcefield *w.* CADT	3824 (+6%)	3823 (+7%)	1430 (+3%)	1358 (+7%)	982 (+13%)	378 (+2%)

Table 18 Training dataset performance for acetylene forcefields using harmonic bond stretch, Manz angle-bend, and ADLD torsion model potentials. The QM-computed (CCSD/def2-TZVPD) equilibrium values were 1.06358 Å (HC), 1.20183 Å (CC), and 180.000° (HCC)

	k_{stretch} (eV Å^{-2})	k_{bend} (eV)	k_{torsion} (eV)	Training RMSE (eV)	Training R-squared
Forcefield *w.* ADLD	41.91 (HC), 111.34 (CC)	1.045 (HCC)	3.400 (k^1_{LD5})	0.1307	0.9681

Table 19 Validation dataset performance for acetylene (HCCH) and deuterated acetylene (DCCD) forcefields using harmonic bond stretch, Manz angle-bend, and ADLD torsion model potentials. This validation dataset compares the normal-mode vibrational frequencies computed from the forcefield to experimentally-measured frequencies (in wavenumber, cm^{-1}). The percent error relative to experiment is shown in parentheses. Due to the molecular symmetry, mode 5 has the same frequency as mode 4, and mode 7 has the same frequency as mode 6

Molecule	freq →	1 (symmetric CH stretch)	2 (asymmetric CH stretch)	3 (CC stretch)	4, 5 (bowl)	6, 7 (wave)
HCCH	Experiment (ref. 47)	3372	3295	1974	730	599
HCCH	Forcefield *w.* ADLD	3588 (+6%)	3501 (+6%)	2106 (+7%)	701 (−4%)	641 (+7%)
DCCD	Experiment (ref. 47)	2704	2431	1766	538	504
DCCD	Forcefield *w.* ADLD	2855 (+6%)	2571 (+6%)	1872 (+6%)	534 (−1%)	515 (+2%)

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

0 is significant while $k^1_{\text{LD1}} = k^1_{\text{LD2}} = k^1_{\text{LD4}} = 0$. Consequently, k^1_{LD5} was the only ADLD mode included in the parameterized forcefield. This forcefield yielded training R-squared > 0.95.

Table 19 summarizes the validation dataset results for this molecule. For both HCCH and DCCD, the vibrational frequencies computed using the forcefield were in excellent agreement with the experimentally-measured frequencies. Both HCCH and DCCD used the same forcefield parameters from Table 18. The change in frequencies from HCCH to DCCD arose solely due to the different masses of D *versus* H atoms.

11. Conclusions

Dihedral torsion model potentials can be categorized into five classes. Class A (aka 'dihedral-only') torsion potentials depend exclusively on the dihedral value (*e.g.*, ϕ_{ABCD}) with no explicit dependence on the bond lengths or bond angles. Class B (aka 'angle-damped') torsion potentials depend exclusively on the dihedral value (*e.g.*, ϕ_{ABCD}) and the two contained bond angle values (*i.e.*, θ_{ABC} and θ_{BCD}) with no explicit dependence on the bond lengths. Class C (aka 'distance-damped') torsion potentials depend exclusively on the dihedral value (*e.g.*, ϕ_{ABCD}) and the three contained bond lengths (*i.e.*, R_{AB}, R_{BC}, and R_{CD}) with no explicit dependence on the bond angles. Class D (aka 'fully-damped') torsion potentials depend exclusively on the dihedral value (*e.g.*, ϕ_{ABCD}), the two contained bond angle values (*i.e.*, θ_{ABC} and θ_{BCD}), and the three contained bond lengths (*i.e.*, R_{AB}, R_{BC}, and R_{CD}). Class E contains all of the miscellaneous torsion potentials that do not fit into any of the first four classes.

The most important new developments in this article pertain to the Class B ('angle-damped') dihedral torsion potentials. First, the combined angle-dihedral coordinate branch equivalency conditions and mathematical constraints were derived and used to construct a model of the angle-damping factors. Second, using the concept of 'completing the squares' these angle-damping factors were used to construct series expansions defining the ADDT, ADCO, and ADLD model potentials. These new dihedral torsion model potentials require only a small number of terms to achieve excellent accuracy, high computational efficiency, and continuous derivatives of all orders with respect to atom-in-material displacements. They capture correct dynamics across a wide range of bond angles including the limiting value of $\theta = \pi$. In contrast, most previously used dihedral–torsion model potentials have either a derivative discontinuity or incorrect dynamics when the bond angle reaches $\theta = \pi$.

To properly resolve the derivative discontinuity at $\theta = \pi$, I showed the torsion term must depend on both the bond angles and dihedral value. Of particular interest, this gives rise to a new torsion-derived angle-bending energy term called the torsion offset potential (TOP). I showed the TOP gives rise in some materials to the unusual physical phenomenon of slip torsion.

The CADT and CACO torsion model potentials apply only to torsions for which it is energetically inaccessible for any contained bond angle to reach $\theta = \pi$. Because neither contained bond angle can energetically reach $\theta = \pi$, this allows the derivative continuity condition for $\theta = \pi$ to be relaxed for such torsions. These two Class A torsion model potentials approximate the torsion barrier height as constant as the contained bond angles change values. As a less computationally expensive and simpler option compared to the ADDT and ADCO model potentials, the CADT and CACO model potentials can be used when the equilibrium values of both contained bond angles are <130°.

I derived a new orthonormal representation of the independent rotatable torsion modes that facilitates automated identification of which particular torsion modes contribute significantly for each dihedral type. This smart selection enables insignificant torsion modes to be excluded from the subsequent forcefield parameterization process. This makes the parameterized forcefields both compact and accurate.

These torsion model potentials cover the following situations:

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

(1) The new angle-damped dihedral torsion (ADDT) model potential is preferred when neither contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $\neq$ 180°), at least one of the contained equilibrium bond angles is $\geq$ 130° (*i.e.*, (θ^{eq}_{ABC} or θ^{eq}_{BCD}) $\geq$ 130°), and the dihedral torsion potential contains some odd-function contributions (*i.e.*, $U[\phi] \neq U[-\phi]$).

(2) The new angle-damped cosine only (ADCO) model potential is preferred when neither contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $\neq$ 180°), at least one of the contained equilibrium bond angles is $\geq$ 130° (*i.e.*, (θ^{eq}_{ABC} or θ^{eq}_{BCD}) $\geq$ 130°), and the dihedral torsion potential contains no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$).

(3) The new constant amplitude dihedral torsion (CADT) model potential is preferred when neither contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $\neq$ 180°), both contained equilibrium bond angles are < 130° (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) < 130°), and the dihedral torsion potential contains some odd-function contributions (*i.e.*, $U[\phi] \neq U[-\phi]$).

(4) The constant amplitude cosine only (CACO) model potential is preferred when neither contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) $\neq$ 180°), both contained equilibrium bond angles are <130° (*i.e.*, (θ^{eq}_{ABC} and θ^{eq}_{BCD}) < 130°), and the dihedral torsion potential contains no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$).

(5) The new angle-damped linear dihedral (ADLD) model potential is preferred when at least one contained equilibrium bond angle is linear (*i.e.*, (θ^{eq}_{ABC} or θ^{eq}_{BCD}) = 180°).

The combination of dihedral pruning;[18,19] classifying each dihedral type as non-rotatable, rotatable, hindered, or linear;[18,19] selecting a ADDT, CADT, ADCO, CACO, or ADLD model potential for each dihedral type; and torsion mode smart selection results in an extremely computationally efficient, accurate, numerically stable, and versatile treatment of dihedral torsion in classical forcefields.

Analytic first derivatives and forces associated with the ADDT, CADT, ADCO, CACO, and ADLD potentials are derived in the ESI.† These analytic formulas are useful for computing forces during classical molecular dynamics simulations. As summarized in Section S11 of the ESI,† these analytic derivative and force formulas were rigorously checked *via* extensive comparisons to values computed numerically using the central finite-difference approximation. The ESI† zip archive contains software code that implements these analytic derivative and force formulas and compares them to values computed numerically using the central finite-difference approximation.

In Section 10, these five torsion model potentials were extensively tested for selected molecules and compared to high-level quantum chemistry calculations or experimental data. In Section 10.2, detailed torsion modal analysis using the ADDT and CADT model potentials was performed for 13 molecules that have two nonlinear equilibrium bond angles: $(CHFCl)_2$, C(OH)ClFH, ethane, FSSF, glyoxal, H_2O_2, HNCO, HNCS, HONC, HSNC, IF_3ClOH, N_2O_2, and PF_4OH. Torsion modal analysis was performed for all of these molecules using rigid bonds and angles, and for some of these molecules also using relaxed bonds and angles. As shown in Table 3 and Fig. 8, for a torsion model containing only the first seven ADDT/CADT orthonormal modes, the model showed superb fit (R-squared > 0.93) to the QM-computed CCSD/def2-TZVPD results.

For C(OH)ClFH, FSSF, HOOH, HNCO, HNCS, HONC, HSNC, IF_3ClOH, ONNO, and PF_4OH, angle-bending curves were QM-computed and compared in Fig. 9 to the sum of the Manz angle-bending potential plus the new torsion offset potential and found to be in excellent agreement. For HNCO, HNCS, HONC, HSNC, excellent agreement was achieved without using any adjustable parameters. For the other six molecules, excellent agreement was achieved using only one adjustable parameter, k_{angle}, which is the angle-bending force constant. The other parameters, which were not adjusted when preparing these angle-bending curves, were taken directly from the torsion modal analysis (see Table 3) for each molecule.

For HNCO, HNCS, HONC, HSNC, C(OH)ClFH, FSSF, H_2O_2, IF_3ClOH, N_2O_2, and PF_4OH, torsion modal analysis was performed at different constrained bond angles and compared to predictions of the ADDT and CADT model potentials. As summarized in Table 7, the ADDT model potential is a more accurate predictor of the torsion barrier height than the CADT model potential when the equilibrium value of at least one of the contained bond angles is ≥130°. In addition to predicting changes in the torsion barrier height, ADDT also predicts changes in the modal coefficients (see Fig. 10) and ϕ_{min} as the constrained bond angle varies. When the equilibrium values of both contained bond angles are <130°, the ADDT model has similar overall accuracy to the CADT model; however, the CADT model is preferred in this case due to its greater simplicity.

Section 10.3 performs an analogous torsion modal analysis using the ADCO and CACO model potentials for those molecules containing no odd-function contributions (*i.e.*, $U[\phi] = U[-\phi]$ which makes *sym*_value = 0). The ADCO and CACO model potentials provided results similar to (but not necessarily identical to) the ADDT and CADT model potentials for these systems. As summarized in Table 10, the ADCO model potential is a more accurate predictor of the torsion barrier height than the CACO model potential when the equilibrium value of at least one of the contained bond angles is ≥130°. Section 10.3 also demonstrated (see Fig. 11) that the cosine-only model torsion potential fails when *sym*_value is substantially larger than zero.

Section 10.4 proves the torsion offset potential is indispensable. Specifically, the CCSD-computed potential energy surface for the HNCO molecule can only be reproduced when the torsion offset potential is included. This molecule exhibits slip torsion.

Section 10.5 analyzed in detail three molecules (*i.e.*, H_3CCN, H_2BCN, and H_2BNC) having single-linear dihedrals plus two molecules (*i.e.*, HCCH and HCNO) having double-linear dihedrals. For each of these five molecules, the Manz bend and ADLD force constants were optimized to a training dataset containing the optimized geometry plus numerous geometries having constrained bond angle(s) and/or constrained dihedral values. In all cases, the training *R*-squared was >0.995, and this

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

indicates superb model performance. The CCSD-computed and ADLD model potential energy surfaces were nearly identical.

Section 10.6 briefly recaps flexibility model performance for five MOFs containing linear dihedrals as described in a companion article.[18,19] Since the torsion barrier is zero when the bond angle is linear, excellent performance was obtained even when the linear dihedral torsions were omitted from the flexibility models.

Section 10.7 studies a chiral molecule containing multiple adjacent rotatable dihedrals leading to many local ground-state conformations. This type of example was included to examine coupling between multiple adjacent rotatable dihedrals. Also, this molecule is large enough to examine the effects of different intracluster nonbonded interaction models. Using Manz's bonded/nonbonded interaction separation ansatz,[20] the training and validation R-squared values depended only weakly on the specific choice of intracluster nonbonded interaction model. Comparisons were also made between the CADT and a cosine-only torsion model potential. Some of the prior literature used a cosine-only torsion model potential even for molecules where it does not strictly apply because $U[\phi] \neq U[-\phi]$ which makes *sym*_value > 0. Although this approximation sometimes yields reasonable results, I recommend that cosine-only torsion potentials such as ADCO and CACO be used only when $U[\phi] = U[-\phi]$ which makes *sym*_value = 0. The ADDT and CADT model potentials are preferred when $U[\phi] \neq U[-\phi]$. A key advantage of the ADDT and CADT model potentials is that a single set of force constant values simultaneously describes both mirror images (*e.g.*, enantiomers) of a chiral center. This was clearly demonstrated for the *S* and *R* enantiomers of the chiral 2-amino-propanal molecule.

Section 10.8 compared vibrational frequencies computed from forcefield models to experimental data for the HNCO, HOOH, and HCCH molecules. These molecules were chosen to provide examples including the ADCO, ADDT, CACO, CADT, and ADLD torsion potentials. Full flexibility models including bond stretches, angle bends, and dihedral torsions were parameterized. For each molecule, the training dataset *R*-squared was >0.95. Vibrational frequencies computed from the parameterized forcefields using the harmonic oscillator approximation were in excellent agreement with the previously published experimentally-measured vibrational frequencies.

In summary, all of these computational tests showed that the new theory of angle-damped dihedral torsion introduced in this article is a remarkable success.

Data availability

Optimized geometries of molecules, selected data analysis spreadsheets, and selected Matlab codes and results are included as part of the ESI.†

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

The author gratefully acknowledges financial support from NSF Career Award DMR-1555376. This work used the Expanse cluster at the San Diego Supercomputing Center (SDSC) through allocation CTS100027 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS[48]) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296. Special thanks to the anonymous reviewers for suggesting that I include examples covering linear dihedrals, computed vibrational frequencies, and a chain-like molecule having multiple conformers. The author thanks Reza Ghanavati for helpful discussions relating to the CACO model potential.

References

1 W. C. Swope and D. M. Ferguson, Alternative expressions for energies and forces due to angle bending and torsional energy, *J. Comput. Chem.*, 1992, **13**, 585–594, DOI: **10.1002/jcc.540130508**.

2 W. Pauli, On the connection of the arrangement of electron groups in atoms with the complex structure of spectra, *Z. Phys.*, 1925, **31**, 765–783, DOI: **10.1007/BF02980631**.

3 L. Radom, W. J. Hehre and J. A. Pople, Molecular orbital theory of electronic structure of organic compounds 13. Fourier component analysis of internal rotation potential functions in saturated molecules, *J. Am. Chem. Soc.*, 1972, **94**, 2371–2381, DOI: **10.1021/ja00762a030**.

4 R. E. Tuzun, D. W. Noid and B. G. Sumpter, Efficient treatment of out-of-plane bend and improper torsion interactions in MM2, MM3, and MM4 molecular mechanics calculations, *J. Comput. Chem.*, 1997, **18**, 1804–1811.

5 D. Dubbeldam, K. S. Walton, T. J. H. Vlugt and S. Calero, Design, parameterization, and implementation of atomic force fields for adsorption in nanoporous materials, *Adv. Theory Simul.*, 2019, **2**, 1900135, DOI: **10.1002/adts.201900135**.

6 S. K. Burger, P. W. Ayers and J. Schofield, Efficient parameterization of torsional terms for force fields, *J. Comput. Chem.*, 2014, **35**, 1438–1445, DOI: **10.1002/jcc.23636**.

7 D. Steele, An ab initio investigation of the torsional potential function of n-butane, *J. Chem. Soc., Faraday Trans. 2*, 1985, **81**, 1077–1083, DOI: **10.1039/f29858101077**.

8 K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**, 1950–1958, DOI: **10.1002/prot.22711**.

9 P. Dauber-Osguthorpe and A. T. Hagler, Biomolecular force fields: where have we been, where are we now, where do we need to go and how do we get there?, *J. Comput. Aided Mol. Des.*, 2019, **33**, 133–203, DOI: **10.1007/s10822-018-0111-4**.

10 J. R. Maple, M. J. Hwang, T. P. Stockfisch, U. Dinur, M. Waldman, C. S. Ewig and A. T. Hagler, Derivation of class II force fields. I. Methodology and quantum force

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

field for the alkyl functional group and alkane molecules, *J. Comput. Chem.*, 1994, **15**, 162–182, DOI: 10.1002/jcc.540150207.

11 C. Tan, J. Jung, C. Kobayashi and Y. Sugita, A singularity-free torsion angle potential for coarse-grained molecular dynamics simulations, *J. Chem. Phys.*, 2020, **153**, 044110, DOI: 10.1063/5.0013089.

12 M. P. Howard, A. Statt and A. Z. Panagiotopoulos, Smooth torsional potentials for degenerate dihedral angles, *J. Chem. Phys.*, 2017, **146**, 226101, DOI: 10.1063/1.4985251.

13 R. E. Tuzun, D. W. Noid and B. G. Sumpter, Molecular-dynamics treatment of torsional interactions accompanied by dissociation, *Macromol. Theory Simul.*, 1995, **4**, 909–920, DOI: 10.1002/mats.1995.040040504.

14 M. Bulacu and E. van der Giessen, Effect of bending and torsion rigidity on self-diffusion in polymer melts: a molecular-dynamics study, *J. Chem. Phys.*, 2005, **123**, 114901, DOI: 10.1063/1.2035086.

15 S. Grimme, A general quantum mechanically derived force field (QMDFF) for molecules and condensed phase simulations, *J. Chem. Theory Comput.*, 2014, **10**, 4497–4514, DOI: 10.1021/ct500573f.

16 M. Bulacu and E. van der Giessen, Molecular-dynamics simulation study of the glass transition in amorphous polymers with controlled chain stiffness, *Phys. Rev. E*, 2007, **76**, 011807, DOI: 10.1103/PhysRevE.76.011807.

17 R. A. Scott and H. A. Scheraga, Conformational analysis of macromolecules. 2. Rotational isomeric states of normal hydrocarbons, *J. Chem. Phys.*, 1966, **44**, 3054–3069, DOI: 10.1063/1.1727180.

18 R. Ghanavati, A. C. Escobosa and T. A. Manz, An automated protocol to construct flexibility parameters for classical forcefields: applications to metal-organic frameworks, *RSC Adv.*, 2024, **14**, 22714–22762, DOI: 10.1039/d4ra01859a.

19 R. Ghanavati, A. C. Escobosa and T. A. Manz, Correction to "An automated protocol to construct flexibility parameters for classical forcefields: applications to metal-organic frameworks", *RSC Adv.*, 2025, under review.

20 T. A. Manz, A formally exact theory to construct nonreactive forcefields using linear regression to optimize bonded parameters, *RSC Adv.*, 2024, **14**, 33345–33383, DOI: 10.1039/d4ra01861c.

21 R. J. Tibshirani, The LASSO problem and uniqueness, *Electron. J. Stat.*, 2013, **7**, 1456–1490, DOI: 10.1214/13-EJS815.

22 R. Tibshirani, Regression shrinkage and selection via the LASSO, *J. Roy. Stat. Soc. B*, 1996, **58**, 267–288, DOI: 10.1111/j.2517-6161.1996.tb02080.x.

23 B. J. Balakumar, J. Friedman, T. Hastie, R. Tibshirani and N. Simon, *Glmnet Python Version 1.0*, https://github.com/bbalasub1/glmnet_python, accessed November 1, 2023.

24 T. Chen and T. A. Manz, A collection of forcefield precursors for metal-organic frameworks, *RSC Adv.*, 2019, **9**, 36492–36507, DOI: 10.1039/c9ra07327b.

25 C. W. Hopkins and A. E. Roitberg, Fitting of dihedral terms in classical force fields as an analytic linear least-squares problem, *J. Chem. Inf. Model.*, 2014, **54**, 1978–1986, DOI: 10.1021/ci500112w.

26 D. Brown and J. H. R. Clarke, Molecular dynamics simulation of an amorphous polymer under tension. 1. Phenomenology, *Macromol.*, 1991, **24**, 2075–2082, DOI: 10.1021/ma00008a056.

27 J. P. Ryckaert and A. Bellemans, Molecular dynamics of liquid alkanes, *Faraday Discuss.*, 1978, **66**, 95–106, DOI: 10.1039/dc9786600095.

28 M. J. Robertson, J. Tirado-Rives and W. L. Jorgensen, Improved peptide and protein torsional energetics with the OPLS-AA force field, *J. Chem. Theory Comput.*, 2015, **11**, 3499–3509, DOI: 10.1021/acs.jctc.5b00356.

29 I. Yildirim, S. D. Kennedy, H. A. Stern, J. M. Hart, R. Kierzek and D. H. Turner, Revision of AMBER torsional parameters for RNA improves free energy predictions for tetramer duplexes with GC and iGiC base pairs, *J. Chem. Theory Comput.*, 2012, **8**, 172–181, DOI: 10.1021/ct200557r.

30 N. Foloppe and A. D. MacKerell, All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data, *J. Comput. Chem.*, 2000, **21**, 86–104.

31 A. Blondel and M. Karplus, New formulation for derivatives of torsion angles and improper torsion angles in molecular mechanics: elimination of singularities, *J. Comput. Chem.*, 1996, **17**, 1132–1141.

32 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. MacKerell, CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *J. Comput. Chem.*, 2010, **31**, 671–690, DOI: 10.1002/jcc.21367.

33 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16, Revision B.01*, Gaussian, Inc., Wallingford CT, 2016.

34 D. Rappoport and F. Furche, Property-optimized Gaussian basis sets for molecular response calculations, *J. Chem. Phys.*, 2010, **133**, 134105, DOI: 10.1063/1.3484283.

35 A. L. L. East, The 16 valence electronic states of nitric oxide dimer $(NO)_2$, *J. Chem. Phys.*, 1998, **109**, 2185–2193, DOI: 10.1063/1.476786.

36 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, Ab-initio calculation of vibrational absorption

Open Access Article. Published on 07 March 2025. Downloaded on 2/12/2026 6:46:00 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

and circular-dichroism spectra using density-functional force-fields, *J. Phys. Chem.*, 1994, **98**, 11623–11627, DOI: **10.1021/j100096a001**.

37 P. J. Stephens, F. J. Devlin, C. S. Ashvar, C. F. Chabalowski and M. J. Frisch, Theoretical calculation of vibrational circular-dichroism spectra, *Faraday Discuss.*, 1994, **99**, 103–119, DOI: **10.1039/fd9949900103**.

38 A. D. Becke, Density-Functional Thermochemistry .3. The role of exact exchange, *J. Chem. Phys.*, 1993, **98**, 5648–5652, DOI: **10.1063/1.464913**.

39 S. Grimme, S. Ehrlich and L. Goerigk, Effect of the damping function in dispersion corrected density functional theory, *J. Comput. Chem.*, 2011, **32**, 1456–1465, DOI: **10.1002/jcc.21759**.

40 T. A. Manz and N. Gabaldon Limas, Introducing DDEC6 atomic population analysis: part 1. Charge partitioning theory and methodology, *RSC Adv.*, 2016, **6**, 47771–47801, DOI: **10.1039/c6ra04656h**.

41 N. Gabaldon Limas and T. A. Manz, Introducing DDEC6 atomic population analysis: part 4. Efficient parallel computation of net atomic charges, atomic spin moments, bond orders, and more, *RSC Adv.*, 2018, **8**, 2678–2707, DOI: **10.1039/c7ra11829e**.

42 A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, UFF, a full periodic-table force-field for molecular mechanics and molecular-dynamics simulations, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035, DOI: **10.1021/ja00051a040**.

43 P. Atkins and R. Friedman, *Molecular Quantum Mechanics*, Oxford University Press, Oxford, United Kingdom, 5th edn, 2011.

44 M. Mladenovic, Six-dimensional potential energy surface and rotation-vibration energy levels of HNCO in the ground electronic state, *J. Serb. Chem. Soc.*, 2019, **84**, 845–859, DOI: **10.2298/JSC190427039M**.

45 M. J. Coffey, H. L. Berghout, E. Woods and F. F. Crim, Vibrational spectroscopy and intramolecular energy transfer in isocyanic acid (HNCO), *J. Chem. Phys.*, 1999, **110**, 10850–10862, DOI: **10.1063/1.479026**.

46 NIST Chemistry Webbook, **https://webbook.nist.gov/chemistry/**, accessed October 2024.

47 C. Andrew, E. E. Etim, O. A. Ushie and G. P. Khanal, Vibrational-rotational spectra of normal acetylene and doubly deuterated acetylene: experimental and computational studies, *Chem. Sci. Trans.*, 2018, **7**, 77–82, DOI: **10.7598/cst2018.1432**.

48 T. J. Boerner, S. Deems, T. R. Furlani, S. L. Knuth and J. Towns, Advancing Innovation: NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support, in, *Proceedings of the Practice and Experience in Advanced Research Computing (PEARC'23), Portland, Oregon*, 2023, 173–176, DOI: **10.1145/3569951.3597559**.

© 2025 The Author(s). Published by the Royal Society of Chemistry