Geochimica et Cosmochimica Acta 396 (2025) 71–90

Contents lists available at ScienceDirect

Geochimica et Cosmochimica Acta

journal homepage: www.elsevier.com/locate/gca

Stable isotope equilibria in the dihydrogen-water-methane-ethane-propane system. Part 1: Path-integral calculations with CCSD(T) quality potentials

Roman Korol [a,b,*], Andrew C. Turner [c,d,e], Apurba Nandi [f,g], Joel M. Bowman [f], William A. Goddard III [a,h], Daniel A. Stolper [c,d]

[a] Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA
[b] Present Address: Department of Chemistry, University of Rochester, Rochester, NY 14627, USA
[c] Department of Earth and Planetary Science, University of California, Berkeley, CA 94720, USA
[d] Energy Geosciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[e] Present address: U.S. Geological Survey, Central Energy Resources Science Center, Denver, CO 80225, USA
[f] Department of Chemistry and Cherry L. Emerson Center for Scientific Computation, Emory University, Atlanta, GA 30322, USA
[g] Present address: Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg City, Luxembourg
[h] Materials and Process Simulation Center, California Institute of Technology, Pasadena, CA 91125, USA

ARTICLE INFO

Associate editor: Alexis Gilbert

Keywords:
Clumped isotope effects
Site-specific isotope effects
Alkanes
Stable isotope fractionations
Path integral calculations

ABSTRACT

Isotopic compositions of alkanes are typically assumed to be kinetically controlled, but recently it has been proposed that alkanes can isotopically equilibrate for both C and H isotopes during natural gas generation. Evaluation of this requires knowledge of the isotopic equilibrium between alkanes and other common hydrogen and carbon bearing species. Here we calculate isotopic equilibria within and between gaseous dihydrogen (H_2), water (H_2O), methane (CH_4), ethane (C_2H_6) and propane (C_3H_8), including isotope fractionation among molecules, clumped isotope effects, as well as among sites of propane (i.e., the site-specific isotope effects) from 0°C to 500°C using a path-integral method paired with high-level descriptions of molecular potentials and the diagonal correction to the Born–Oppenheimer approximation. While path-integral calculations with high-level CCSD(T) potentials are available for the isotopic equilibria involving methane, the path-integral calculations for ethane and propane have only been performed based on lower-level descriptions of the molecular potentials. We analyze the relative importance of various approximations that are commonly employed when isotopic equilibria are evaluated. We find that clumped isotope effects can be calculated to the same accuracy using computationally inexpensive combination of the Bigeleisen-Mayer-Urey model with the molecular potential from density functional theory. In contrast, fractionation and site preferences of both deuterium and carbon-13 benefit from the use of the higher level CCSD(T) potentials and accounting for anharmonic effects. Additionally, for fractionation and site preference of deuterium, corrections to Born–Oppenheimer approximation can also be important.

1. Introduction

The environmental sources and sinks of low-molecular weight alkane gases such as methane, ethane, and propane are commonly constrained based on their stable carbon and hydrogen isotopic compositions. The initial isotopic compositions of these gases are generally considered to be set by kinetic isotope effects associated with gas generation either microbially (see review in Turner et al., 2021) or via the thermally induced breakdown of larger hydrocarbons in sedimentary basins in which larger hydrocarbon molecules break down and convert to smaller products (see reviews in Hunt, 1996; Seewald 2003; Stolper et al., 2018). Support for this framework rests with the idea that hydrocarbon gas generation is controlled by irreversible reactions in which larger hydrocarbon molecules break down and convert to smaller products. Recently, it has been proposed that thermally induced cracking reactions could proceed via reversible radical reaction networks that allow both the hydrogen and carbon isotopic composition of alkanes to equilibrate during generation (Thiagarajan et al., 2020). Evidence for this idea is based in part on differences in stable isotopic compositions of alkanes being consistent with having attained isotopic equilibrium at temperatures expected for gas generation between different alkanes (Thiagarajan et al., 2020; Xie et al., 2021; Xie et al., 2024), between

* Corresponding author at: Department of Chemistry, University of Rochester, Rochester, NY 14627, USA.
E-mail address: roman.korol@rochester.edu (R. Korol).

https://doi.org/10.1016/j.gca.2025.02.028
Received 17 July 2024; Accepted 25 February 2025
Available online 28 February 2025

different sites in propane (Xie et al., 2018), and for clumped isotopic composition of methane vs. isotopic compositions of other alkanes (Dong et al., 2021; Xie et al., 2021; Eldridge et al., 2023).

Critical to evaluating this proposal are constraints on the temperature dependence of equilibrium fractionations that describe the isotope exchange reactions of interest. One approach is to determine this dependence experimentally. For reactions involving short-chain alkanes, this has been done at both low (3–200°C; Turner et al., 2021) and high temperatures (200–500°C) for the methane-H_2 system, at high temperature (>200°C) for methane-CO_2 (Horita, 2001; Kueter et al., 2019) and from 1–500°C for methane clumped isotopes (Eldridge et al., 2019). Hydrogen isotopes have been equilibrated between the center and terminal sites of propane from 30–200°C (Xie et al., 2018). Equilibrium isotopic compositions for carbon or hydrogen between ethane and propane or between them and any other species have only been experimentally estimated recently in Xie et al. (2024) at 100°C for hydrogen isotopic equilibrium for methane, ethane, and propane and at 200°C for propane, *i*-butane, *n*-butane, and *n*-propane.

An alternative approach to conducting experiments is to use theory to calculate equilibrium isotope effects. Since the 1930s, numerous gas and condensed phase systems have been examined using varying levels of theory ranging from those that incorporate harmonic effects only (e.g., Urey, 1947; Bigeleisen and Mayer, 1947) to more sophisticated models that incorporate various anharmonic effects (e.g., Richet et al., 1977; Bardo and Wolfsberg, 1978; Liu et al., 2010; Webb and Miller, 2014; Pinilla et al., 2014; Webb et al., 2017; Zhang and Liu, 2018; Gao et al., 2023). Historically, alkanes larger than methane have not been the focus of such theoretical studies. More recently, due in large part to analytical advances, a number of calculations have been conducted to elucidate clumping and site-specific isotope effects in ethane, propane, and butane, some of which have included anharmonic effects while others have made the harmonic approximation (Webb and Miller, 2014; Cheng and Ceriotti, 2014; Piasecki et al., 2016b; Webb et al., 2017; Liu et al., 2021; Yin et al., 2024). Calculations for carbon and/or hydrogen isotope effects between methane and short-chain alkanes have also been presented recently in Thiagarajan et al. (2020), Polyakov and Horita (2021), Xie et al. (2024), and Watts et al. (2024). These inter-alkane studies are all based on calculations that employ a series of approximations outlined in Urey (1947) and Bigeleisen and Mayer (1947) in which vibrations are treated harmonically, rotations are treated as classical, and other simplifying assumptions are invoked. Isotope effects between larger organic molecules have also been examined in a variety of studies (e.g., Rustad, 2009; Wang et al., 2009a, b; Iron and Gropp, 2019; Thiagarajan et al., 2020; Gropp et al., 2021; Polyakov and Horita, 2021; Watts et al., 2024; Xie et al., 2024). In these cases, harmonic approximations are generally used and, if anharmonicity is incorporated, it is based on a scaling correction to the zero-point energy term only (Iron and Gropp, 2019; Gropp et al., 2021).

To our knowledge, carbon and hydrogen equilibrium isotopic fractionation factors between short-chain alkanes have never been calculated with inclusion of anharmonicity at CCSD(T) level of theory. Use of high-quality potentials, incorporation of anharmonic corrections, and even corrections to the Born-Oppenheimer approximation have been shown in prior work to be necessary for achieving agreement between experiments and theory for equilibrium between H_2 gas and liquid water (Bardo and Wolfsberg, 1976; Turner et al., 2021) as well as between CH_4 and H_2 (Turner et al., 2021). In contrast, for clumped reactions, due to error cancellations, lower level theory and harmonic approximations are generally considered sufficient to yield agreement with experiments and higher level calculations (Wang et al., 2004; Webb and Miller, 2014; Liu and Liu, 2016; Eldridge et al., 2019).

Less is known about the need for inclusion of anharmonicity in calculations of site-specific isotope effects. Webb and Miller (2014) found that for propane, anharmonic corrections change results vs. harmonic calculations by <5% relative. However, this was done using a crude molecular potential, so it is not clear whether this extends to high-quality potentials. Recent work by Yin et al. (2024) using a high-quality molecular potential for propane confirms the earlier finding of Webb and Miller (2014) that anharmonic corrections are of second-order (<10% relative) importance. In contrast, for isomers of butane with similar quality of potential, Liu et al. (2021) found anharmonic corrections to the harmonic approximation can change site-specific hydrogen isotope effects by as much as ~30% relative.

We consider that a better understanding of the requirements necessary for accurate calculations of equilibrium isotope effects in alkanes is a pressing need given the various emergent measurement techniques that are beginning to allow for new measurements of clumped isotope effects and site-specific isotope effects in a range of molecules (e.g., Julien et al., 2020; Lloyd et al., 2021, 2023; Gilbert, 2021; Wilkes et al., 2022; Neubauer et al., 2023; Weiss et al., 2023; Csernica et al., 2023, Zeichner et al., 2023). Many of these molecules are sufficiently large that computationally intensive theoretical calculations will likely not be possible in the near future, nor will equilibrium necessarily be achievable experimentally. As such, knowledge of what level of theory is needed to establish equilibrium fractionation factors at necessary accuracy for larger molecules will be an ongoing and critical need for the study of isotope effects.

In two linked studies, one theoretical and the other experimental, we present results on isotope effects in the systems H_2-H_2O-CH_4-C_2H_6-C_3H_8 expanding our prior work (Eldridge et al., 2019; Turner et al., 2021) to include the two larger alkanes (ethane and propane). In this contribution, we focus on theoretical calculations of short-chain alkanes and examine the importance of (i) the quality of molecular potential (both the level of theory and the basis set size); (ii) incorporation of anharmonic effects, and (iii) corrections to the Born-Oppenheimer effects for the calculation of isotope effects between molecules, sites in a molecule, and clumped isotope effects. Our goal in doing this is to develop an understanding of the importance, or lack thereof, of incorporating these effects into calculations involving organic molecules more complex than methane. In the second contribution, we provide results on experimental equilibrations of ethane and propane with H_2 for hydrogen isotopes and compare these results to the theoretical results presented here.

2. Nomenclature

In this section, we provide the essential nomenclature used to describe relative differences in the isotopic composition between different species, specific sites in a molecule, and the clumped-isotope composition of molecule. Although much of the notation used is standard in geochemistry, we have found that notational differences do exist between chemical and geochemical studies that can lead to confusion in regard to the meaning of what is actually being calculated or measured. As such, we provide a complete explanation of the nomenclature used here, the approximations employed (and justifications that these do not introduce inaccuracies) and note where prior work has used different notations in section S1 of the supplementary material.

Differences in the isotopic composition between phases, species, or sites within a molecule are given using alpha notation, where alpha is termed the fractionation factor:

$$ {}^{i}\alpha_{A-B} = \frac{{}^{i}R_A}{{}^{i}R_B} \tag{1} $$

${}^{i}R = [{}^{heavy}N]/[{}^{light}N]$). The excess or deficit of a clumped isotopologue vs. a random distribution is given using Δ notation:

$$ \Delta_i = 1000 \times \left(\frac{X_{i,\,sample}}{X_{i,\,random}} - 1 \right) \tag{2} $$

Here X = [multiple substituted isotopologue]/[unsubstituted isotopologue] and 'sample' refers to the measured relative concentrations of the isotopologues whereas the 'random' refers to the composition that would be calculated if all isotopes are randomly distributed amongst all

isotopologues.

Both the α (Eq. (1)) and Δ values (Eq. (2)) can be related to expressions involving equilibrium constants for isotope exchange reactions. We use isotope exchange reactions with the following generic form between phases or species A or B where isotopic substitution is given by *:

$$^{*}A + B \rightleftharpoons A + {}^{*}B \tag{3}$$

An example would be $CH_3D + H_2O \rightleftharpoons CH_4 + HDO$. For these reactions we can relate the equilibrium constant K_{eq} to α as follows:

$$\alpha^{eq} = \frac{K}{K^{random}} \tag{4}$$

K^{random} is the equilibrium constant that would be found at the infinitely high temperature limit. This normalization is done in order to remove any dependence of equilibrium constants on symmetry numbers. Equivalent forms can be written for isotope exchange reactions between sites as described in section S1.2.

For clumped isotope exchange reactions, we write:

$$\Delta_X^{eq} = 1000 \times \ln \frac{K_X}{K_X^{random}} \tag{5}$$

As described in section S1.3, reactions describing clumping require additional consideration when site-specific effects are present as well.

We calculate equilibrium constants using appropriate reduced partition function ratios (RPFRs), where for the example given in Eq. (3):

$$\alpha_{A-B}^{eq} = \frac{RPFR\left(\frac{^{*}B}{B}\right)}{RPFR\left(\frac{^{*}A}{A}\right)} \tag{6}$$

The RPFRs in Eq. (6) are defined as follows (Bigeleisen and Mayer, 1947):

$$RPFR\left(\frac{A}{B}\right) = \frac{Q_A}{Q_B}\frac{\sigma_A}{\sigma_B}\prod_{atoms}\left(\frac{m_B}{m_A}\right)^{\frac{3}{2}} \tag{7}$$

where Q represents the partition function for isotopologues A and B. The product of the ratio of masses raised to the power of 3/2 runs over all isotopes that are different between A and B, and σ are the corresponding rotational symmetry numbers. Eq. (7) inherently normalizes out both the mass terms and the symmetry numbers in Q such that at infinite temperature (i.e., in the classical limit), an RPFR is unity.

3. Methods

3.1. Reduced partition function ratio (RPFR) calculations

We calculate RPFRs using two distinct approaches: (*i*) harmonically using the approach originally outlined in Bigeleisen and Mayer (1947) and Urey (1947) and (*ii*) using the path integral Monte Carlo (PIMC) method following the approach developed by Tuckerman et al., (1993) and Cheng and Ceriotti (2014). Harmonic calculations require fewer computational resources compared to the PIMC method but make a number of approximations (as detailed in the following section), that can lead to systematic errors (i.e., inaccuracies) in calculated RPFRs (e.g., Wang et al., 2004; Liu et al., 2010; Webb and Miller, 2014; Webb et al., 2017; Eldridge et al., 2019; Turner et al., 2021). In some cases, these errors mostly cancel out when harmonic RPFRs are used to calculate equilibrium isotope effects (Webb and Miller, 2014; Eldridge et al., 2019), but this is not always the case (Liu et al., 2010; Webb and Miller, 2014; Turner et al., 2021). For all calculations of RPFRs (harmonic and PIMC) we neglect the effects of: (1) intermolecular interactions, (2) electronic excitations, (3) internal structure of the nuclei, and (4) relativistic effects. The justifications for these assumptions are given in prior studies (Webb and Miller, 2014; Webb et al., 2017; Eldridge et al., 2019; Turner et al., 2021).

Finally, all of our RPFR calculations require knowledge of the potential energy of the molecule as a function of geometry. As such, we also describe our methods to find the potential energy used to calculate the RPFR, as discussed below.

Below we provide brief descriptions of methods used for these calculations and provide a more detailed description in section S3 as our approaches follow previously published methods.

3.2. Harmonic calculation

The harmonic method assumes that the total partition function can be written as the product of vibrational, rotational and translational components; then the vibrations are approximated as harmonic, rotations as rigid, and both rotations and translations are treated as if they are classical. With these approximations, the RPFR of an isotopologue pair (starred over unstarred) can be written as (Urey, 1947; Bigeleisen and Mayer, 1947; Richet et al., 1977; O'Neil, 1986; Blanchard et al., 2017):

$$RPFR_{harmonic} = e^{-\frac{E_0^{*}-E_0}{k_B T}}\prod_{j=1}^{a}\frac{\omega_j^{*}}{\omega_j} \times \frac{1-e^{-\frac{hc\omega_j}{k_B T}}}{1-e^{-\frac{hc\omega_j^{*}}{k_B T}}} \tag{8}$$

In Eq. (8), ω_j is the harmonic frequency in wavenumbers (cm^{-1}) of the j^{th} normal mode, a is the total number of vibrational modes ($a = 3N - 5$ for linear molecules, $a = 3N - 6$ for non-linear molecules and N is the number of atoms). The leading term (E_0) is the combined harmonic zero-point energy of all modes. We note that in some presentations of Eq. (8), the contribution of the zero-point energy of each mode is included instead in the product over the modes.

We calculate the harmonic frequencies (given in table S7) using the free (for academic use) ORCA 5.2 program (Neese and Wiley, 2012; Neese, 2018; Neese et al., 2020), except for CCSD(T)-F12A-based frequencies (which are not available within ORCA 5.2). For these only, we use the MOLPRO program as detailed in section 3.3.1 An example input file and the brief description of the calculation algorithm are provided in section S3.1.

We calculate harmonic frequencies using a variety of common methods employed in prior isotope geochemistry studies to assess the relative importance of the approach used to describe of the potential energy of the molecule. These include density functional theory (DFT) with the B3LYP functional (Lee et al., 1988; Becke, 1993) and four increasingly large Pople's (Ditchfield et al., 1971) basis sets (6–31G, 6–311G, 6–311G**, 6–311++G**) as well as second order Møller-Plesset perturbation theory (MP2), coupled-cluster with single and double excitations (CCSD), and the traditional "gold standard" CCSD(T) where triple excitations are included perturbatively. We use the correlation-consistent basis sets (Dunning, 1989) that are designed for post-Hartree-Fock calculations. We use four basis sets of increasing size: cc-pVDZ, aug-cc-pVDZ, cc-pVTZ, and aug-cc-pVTZ. For completeness, we also include calculations using the restricted Hartree-Fock (RHF) method. This mean-field theory does not take into account the electron–electron correlations and, as such, is generally not used in modern calculations. Finally, we also calculate the harmonic frequencies using two empirical force fields (CHARMM and AIREBO) because these appeared in recent studies of equilibrium isotope effects (Webb and Miller, 2014; Cheng and Ceriotti, 2014), including for some molecules examined here.

3.3. PIMC calculations

Unlike the harmonic approach, path-integral methods like PIMC or path integral molecular dynamics (PIMD) fully account for the

anharmonicity of vibrations and rovibrational couplings. This more rigorous description comes at a computational cost that can become prohibitive for large molecules or where highly precise RPFR values are required (e.g., <0.1‰ precision). It is also important to note that the accuracy of path integral-based calculations also depends on the quality of the description of the potential energy of the molecule. Thus, a PIMC calculation will not necessarily lead to a more accurate RPFR if the potential energy of the molecule is itself not sufficiently accurate. Put another way, just because PIMC or PIMD is used does not necessarily mean it is more accurate than a harmonic approximation as the accuracy also depends on the quality of the potential energy used.

Path-integral-based methods utilize the imaginary-time path-integral formalism (Feynman and Hibbs, 1965) to sample the quantum Boltzmann statistics of the molecule. The quantum statistics are sampled by mapping each of the N (quantum) atoms onto P (classical) copies of that atom (Schweizer et al., 1981), yielding NP classical particles that interact via the modified potential:

$$U_P\left(\left\{\boldsymbol{r}_j^{(k)}\right\}\right) = \sum_{k=1}^{P} U\left(\boldsymbol{r}_1^{(k)}, \cdots, \boldsymbol{r}_N^{(k)}\right) + \sum_{j=1}^{N}\sum_{k=1}^{P} \frac{m_j \omega_P^2}{2}\left(\boldsymbol{r}_j^{(k)} - \boldsymbol{r}_j^{(k-1)}\right)^2 \tag{9}$$

where $\boldsymbol{r}_j^{(k)}$ indicates the position of the jth atom in the kth ring-polymer bead. Atoms inside each of the P copies of the system interact via the original potential energy surface (PES) $U(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_N)$ as described by the first term. The second term connects every atom in each of the P copies to the same atom of the previous copy via a harmonic spring with frequency $\omega_P = \frac{Pk_BT}{\hbar}$; note that $\boldsymbol{r}_j^{(0)} = \boldsymbol{r}_j^{(P)}$, which means that the last copy of the system is linked to the first copy, closing the ring. Thus, each quantum atom of the original molecule creates a ring polymer of P beads, which are connected via harmonic springs.

In the limit of an infinite number of beads P, the partition function of the real quantum system is equal to the partition function of the fictitious P-times larger classical system. In practice, we use approximations to evaluate the potential energy, and a sufficiently large but finite P is chosen (table S8). Then Monte Carlo (MC) or molecular dynamics (MD) is used to sample the classical system — here we employed the MC importance sampling method (Frenkel and Smit, 2002) and averaged over tens of millions of samples (table S9).

We compute partition function ratios with PIMC (table S10) using the same methodology as that employed by Webb et al. (2017), Eldridge et al. (2019), and Turner et al. (2021) as described in detail in the supplementary section S3. In brief, we use the direct scaled-coordinate estimator (Cheng and Ceriotti, 2014) to calculate the RPFR for every single heavy substitution relative to the lighter isotopologue and for every double substitution relative to the corresponding singly substituted isotopologue. We choose the number of beads P to balance the accuracy (for a given calculated potential energy) that improves as P increases vs. the statistical uncertainty of sampling that increases with P.

3.3.1. Potential energy surfaces used in the PIMC calculations

Our PIMC calculations of RPFRs require the potential energy to be calculated for $>10^{10}$ different molecular geometries for each isotopologue. In contrast, harmonic calculations described in section 3.2 require about 10^3 potential energy evaluations to compute the harmonic frequencies of propane and fewer for the smaller molecules. As a result, we use a different approach for the PIMC potential energy evaluations in order achieve better computational efficiency compared to the direct (i.e., geometry in, energy out) calculations done for the harmonic calculations.

To do this, we created and used a fit of the potential energy surface (PES) of all molecules for our PIMC calculations in an identical manner via permutationally invariant Δ-learning (Nandi et al., 2021; see section S3.2 for details) using CCSD(T)-F12A (explicitly correlated coupled-cluster with single, double, and perturbative triple excitations) method (Adler et al., 2007) with the aug-cc-pVTZ basis set (Dunning, 1989) calculated in MOLPRO 2021.3 package (Werner et al., 2012, 2020). We refer to this method using a shorthand "F12/ATZ" throughout the paper.

We note, this approach differs to that of our prior studies (Eldridge et al., 2019; Turner et al., 2021), where we used what we considered to be the best published potential energy surfaces available for $H_2O_{(g)}$, $CO_{2(g)}$, and $CH_{4(g)}$ (but not H_2, which can be calculated to high accuracy). A possible issue with this approach is that the methods used to calculate the PES of different molecules generally rely on different approximations in order to obtain the molecular potential possibly limiting favorability of error cancellation when RPFRs between different molecules are calculated. We observed this previously for the harmonic RPFRs in Turner et al. (2021): different methods and basis set sizes can yield RPFRs that differ by as much as 40‰. However, when the relevant ratios of RPFRs (Eq. (7) are calculated using the same methods and basis set sizes to obtain a fractionation factor, most of these differences cancel out. As such, here we chose to calculate all PES in the same manner.

3.4. Diagonal Born-Oppenheimer correction

All potential energy calculations in this work are done within the Born-Oppenheimer (BO) approximation (Born and Oppenheimer, 1927; Handy et al., 1986), which is typically employed in calculations of equilibrium isotope effects. However, it has been shown that corrections to the BO approximation can substantially change calculated fraction factors for isotopes of hydrogen. For example, application of the diagonal Born-Oppenheimer (DBO) correction changes the fractionation of deuterium between $H_{2(g)}$ and $H_2O_{(g)}$ by ~20‰ at room temperature (Bardo and Wolfsberg, 1978).

We explored the importance of the DBO corrections on the molecules studied here using the CFOUR package to calculate the DBO correction at the CCSD level of theory and aug-cc-pVTZ basis set (table S11). The exact DBO correction depends on both the molecular geometry and the level of theory. We verify that these dependencies are weak in section S3.3, as expected from equivalent calculations reported in Zhang and Liu (2018). Therefore, we compute these corrections only at the geometries minimized by the F12/ATZ method at CCSD/ATZ level of theory and assume it stays the same for all other geometries (see section S3.3 for justification). With this assumption, a correction can be applied to any RPFR calculated based on the Born-Oppenheimer approximation as an after-the-fact multiplicative correction:

$$RPFR_{DBOC} = RPFR \times \exp\left(-\frac{\Delta C}{k_B T}\right) \tag{10}$$

This is only applied to fractionations between molecules or sites but not clumped isotopes as this correction has no effect on calculations of clumped isotope equilibria (including site-specific clumping; Bardo and Wolfsberg, 1975).

4. Results

Here we present results for the five molecules studied organized based on the factors that we wish to understand better and discuss: (*i*) the differences in equilibrium isotope effects associated with choice of PES for harmonic calculations (tables S12-S14); (*ii*) the size of PIMC corrections on harmonic calculations; and (*iii*) the importance of DBO corrections.

4.1. Differences in harmonic calculations using different PESs

4.1.1. Harmonic calculations for alkanes

We compare alkanes first to explore how well errors cancel when RPFRs based on the same method are combined to calculate a fractionation factor in similar bond types (C–C and C–H) between different

molecules. We present theses as differences (absolute left axis, relative right axis) in the fractionation factor ($1000 \times \ln\alpha^{eq}$) vs. 1000/T relative to the F12/ATZ method. Intra-alkane fractionation factors with methane vs. ethane are shown in Fig. 1a and 1b and methane vs. propane in Fig. 1c and 1d. All are from 0 to 500°C. Here and throughout, we only present results for the larger basis set sizes (triple-ζ and augmented triple-ζ) and the higher-quality electronic structure method (omitting the Hartree-Fock and force field-based results). The lower-quality calculations are given in supplementary figures and table S12.

For both C and H isotopes, we see up to about 10% relative difference (up to 10‰ absolute for H and up to 1.5‰ absolute for C) in $1000 \times \ln\alpha^{eq}$ between the studied methods. Convergence to similar values does not occur as method quality improves from lowest quality of MP2 through DFT (B3LYP) to CCSD, to CCSD(T), nor use of triple-ζ vs augmented triple-ζ. Rather, we see that different methods largely scatter in the ±10% relative range and are similar to results obtained with smaller double-ζ basis sets (DZ and ADZ; Fig. S2). In contrast, computationally less expensive descriptions of the potential energy of the molecules based on restricted Hartree-Fock (RHF) and the empirical force fields that we tested (AIREBO and CHARMM) typically yield much larger relative difference (see Fig. S3; up to 200% relative for H and 100% for C).

4.1.2. Harmonic calculations of hydrogen isotope fractionations in alkanes, H_2, and H_2O

We next turn to hydrogen isotope fractionation factors involving molecules other than alkanes, specifically H_2 and H_2O. We provide a similar plot as Fig. 1 for all species relative to H_2 in Fig. 2a-b and of water relative to other molecules in Fig. 2c-d, all from 0 to 500°C — for the latter we do not include H_2 as the size of $1000 \times \ln\alpha^{eq}_{H_2O-H_2}$ is ~ 10 times larger than the others. Unlike for Fig. 1, temperatures are not shown on the secondary x axis as they differ at a given value of $1000 \times \ln\alpha^{eq}$ depending on the molecule pair.

We observe that fractionation factors relative to dihydrogen show absolute differences of up to ±50‰ with respect to the F12/ATZ result. However, because the absolute values of $1000 \times \ln\alpha^{eq}_{molecule-H_2}$ are ~10x larger than for the alkanes, relative differences are less than ±3% for the larger basis sets and less than ±5% for the smaller basis sets (Fig. S4).

Fractionations relative to water yield relative differences larger than 25% (±40‰) and thus are the most sensitive relatively compared to alkane-alkane or alkane-H_2 hydrogen isotope fractionations. Double- and triple-ζ (both augmented and not) basis sets yield similar differences. In contrast, differences are significant (−50‰ to −180‰, which changes the sign of $1000 \times \ln\alpha^{eq}$ with respect to the reference at low temperatures) for the smaller 6–31G and 6–311G basis sets at the DFT (B3LYP) level of theory (see Fig. S4, table S12). The lower level RHF method shown in Fig. S5 differs significantly from the results presented in Fig. 2, i.e., up to +200‰ (+15% relative) for fractionations with dihydrogen and up to +100‰ (+25% to +200% relative) for fractionations between water and alkanes. Note, we did not use the CHARMM force field for water or dihydrogen and could not use the AIREBO force field as it is not defined for these molecules.

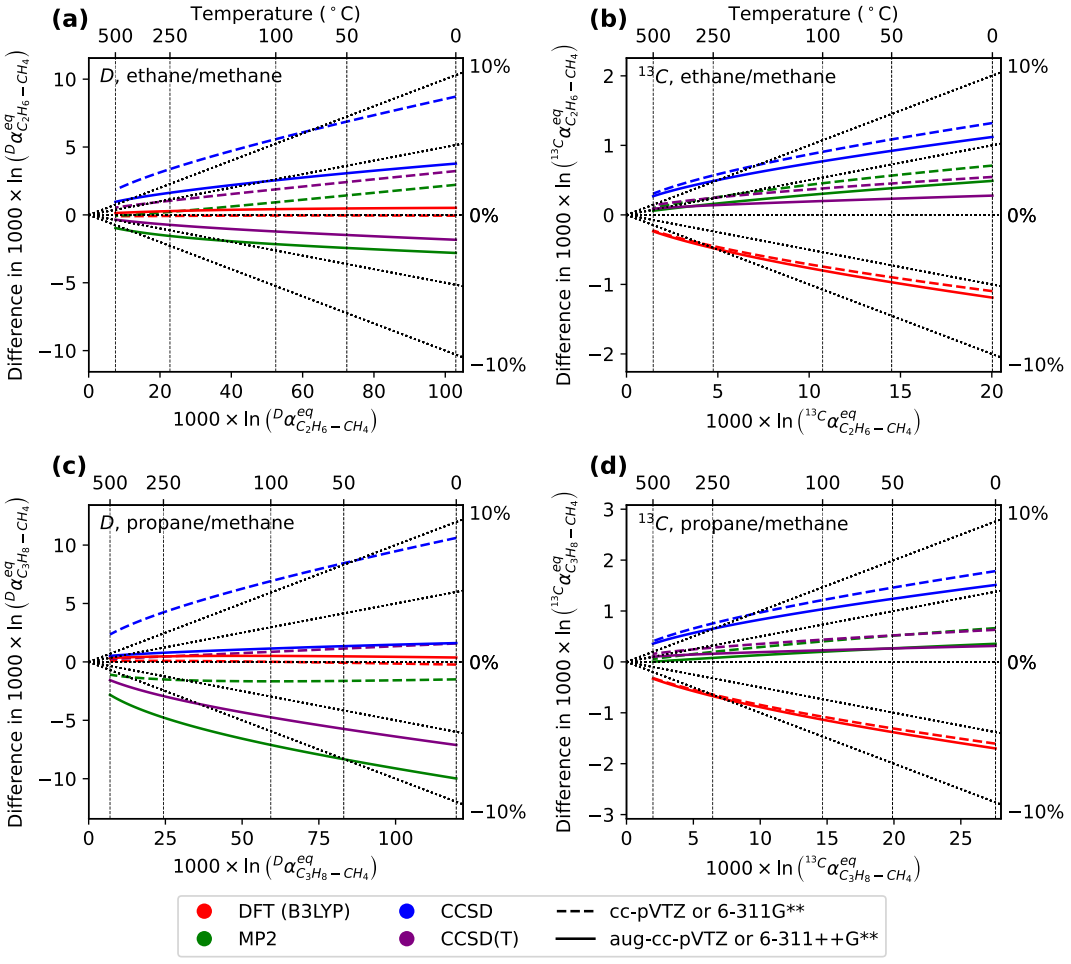

Fig. 1. Difference in harmonic fractionation of deuterium (a,c) and carbon-13 (b,d) vs. methane for ethane (a-b) and propane (c-d) computed with 4 commonly used electronic structure methods and triple-ζ basis sets. Dotted vertical lines indicate temperatures (from 0°C to 500°C) and slanted dotted black lines denote relative difference (0, 5, and 10%). Note that the basis set is specified in all figure labels as Dunning-style (e.g., cc-pVDZ) or Pople-style (e.g., 6–311G). The former refers to all non-DFT calculations, whereas Pople-style label specifies the basis set used in DFT (B3LYP) calculations only.

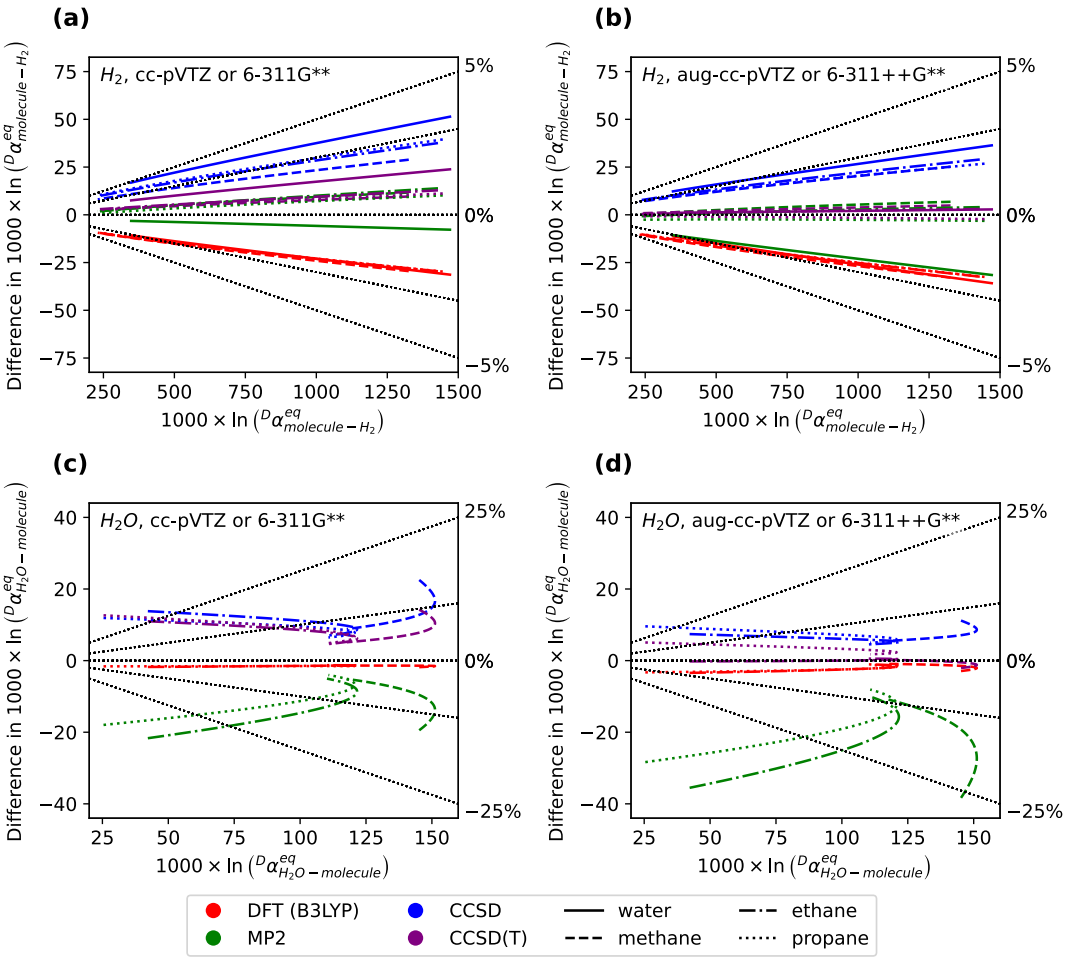

Fig. 2. Difference in harmonic fractionation of deuterium vs. dihydrogen (a-b) and vs. water (c-d) computed with 4 commonly used methods relative to the F12/ATZ method. The smaller 6–311G** basis set was used for the DFT calculations and cc-pVTZ basis set for all other methods in panels (a,c); the larger 6–311++G** basis set was used for the DFT calculations and aug-cc-pVTZ basis sets for all other methods in panels (b,d). Dotted vertical lines indicate temperatures (from 0°C to 500°C). The slanted dotted black lines denote relative difference (0, 3, 5, 10, and 25%).

4.1.3. *Harmonic site-specific isotope effects in propane*

We next present results for site-specific isotope effects in propane (Fig. 3; table S13). Relative differences in 1000 × $\ln\alpha^{eq}_{\text{center-terminal}}$ of up to ±10% are observed for both D and ^{13}C, which corresponds to absolute differences up to 12‰ for D and up to 1.2‰ for ^{13}C. Use of smaller basis set size does not change these general observations (see Fig. S6), while computationally less expensive methods (RHF, AIREBO, CHARMM) yield results that differ significantly from the F12/ATZ reference for D of

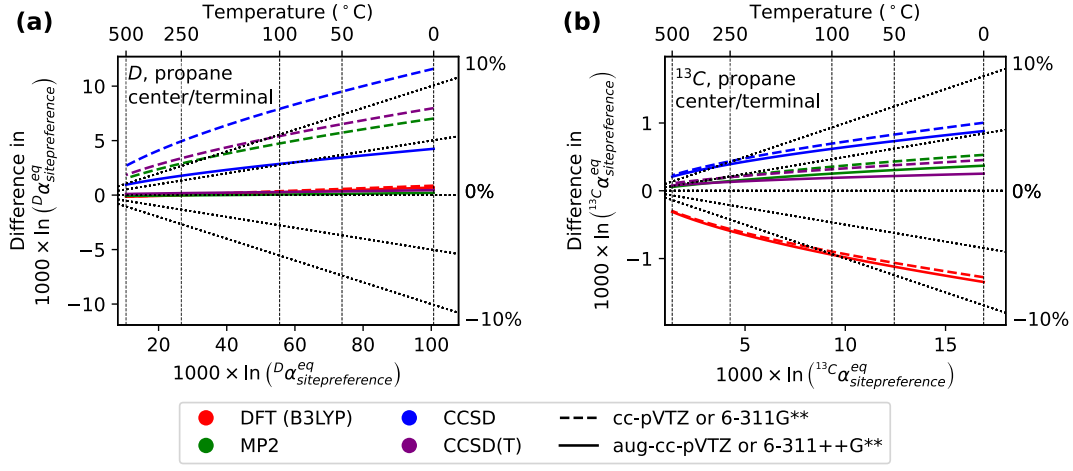

Fig. 3. Differences in propane position-specific isotope effect for (a) deuterium and (b) carbon-13 computed with four methods at the triple-ζ basis set size. All computations are harmonic and are relative to the harmonic F12/ATZ method. The 6–311G** and 6–311++G** basis sets were used for the DFT calculations, while cc-pVTZ and aug-cc-pVTZ basis sets were used for all other methods in panels. Dotted vertical lines indicate temperatures (from 0°C to 500°C). The slanted dotted black lines denote relative difference (0, 5 and 10%).

up to 80% relative (Fig. S7a). Site preferences for carbon-13 in propane described by the computationally less expensive methods (RHF, AIREBO, CHARMM) yield similar (up to ±1.6‰) deviations as observed for the more sophisticated electronic structure calculations (see Fig. S7b).

4.1.4. Harmonic clumped isotope effects

We now present clumped isotope effects (table S14) within the harmonic approximation for heavy isotope substitutions that are as close as possible in terms of distance: either bound to each other or, if that is not possible (e.g., D+D clumping in methane), bound to the same site. These types of bonds generally produce the strongest clumping effect (i.e., largest value of Δ^{eq}) vs. heavy substitutions at larger distances (Wang et al., 2004). The D+D clumped isotope effect calculations for the different molecules studied here are shown in each of the six panels of

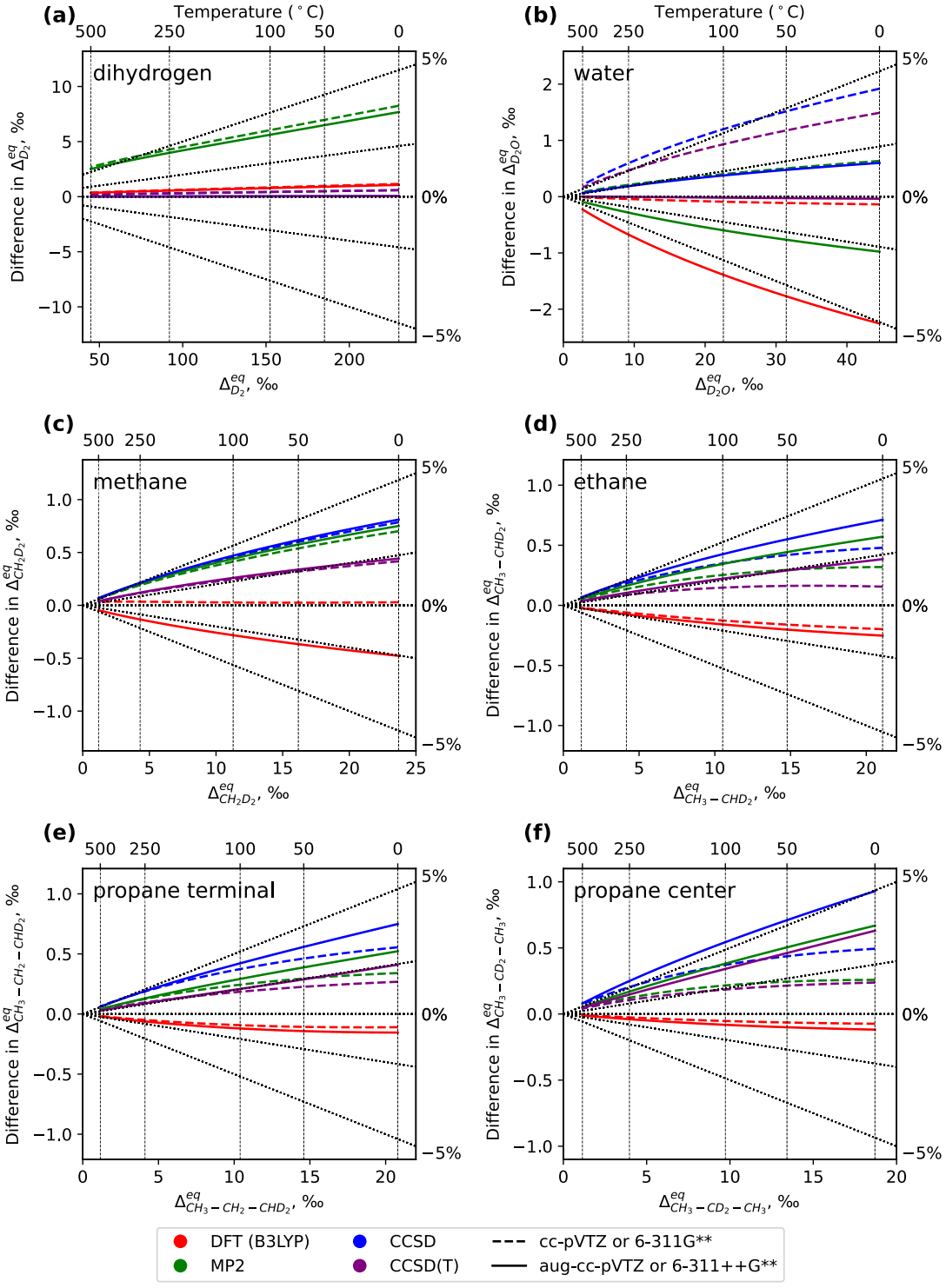

Fig. 4. Differences in D+D clumped isotope effects for (a) dihydrogen, (b) water, (c) methane, (d) ethane and (e-f) propane sites computed with four methods at the triple-ζ basis set size. All computations are harmonic and are relative to the harmonic F12/ATZ method. For ethane and propane only the strongest clumped effects (i.e., when both deuterium atoms are bonded to the same atom) are shown. Dotted vertical lines indicate temperatures (from 0°C to 500°C). The slanted dotted black lines denote relative difference (0, 2, and 5%).

Fig. 4. We find that D+D clumping as calculated by Δ^{eq} deviate by less than ±5% relative as compared to the F12/ATZ reference method. The absolute deviations are up to +7‰ for dihydrogen, up to ±2‰ for water, and less than ±1‰ for alkanes. The range of deviation for the relative differences in ^{17}O+D and ^{18}O+D effects in water is smaller (i.e., ±2% relative or up to ±0.2‰ absolute) with the same observed for ^{13}C+D effects, which are within ±3% of the reference value (i.e., ±0.2‰ absolute, see Fig. 5). Finally, although the ^{13}C+^{13}C effects in ethane and propane (shown on Fig. 6) have larger deviations of up to −8% relative to the reference for the DFT calculations, these are small in absolute terms (i.e., −0.02‰ absolute). MP2, CCSD and CCSD(T) differ by less than +5% relative to the reference, which is less than +0.01‰ absolute.

Clumped isotope effects computed with the triple-ζ basis set are only a few (1–3%) percent closer to the reference than that calculated with the double-ζ basis sets for the post-Hartree-Fock methods (Figs. S8, S10 and S12). In contrast, we observe the basis set size can be more

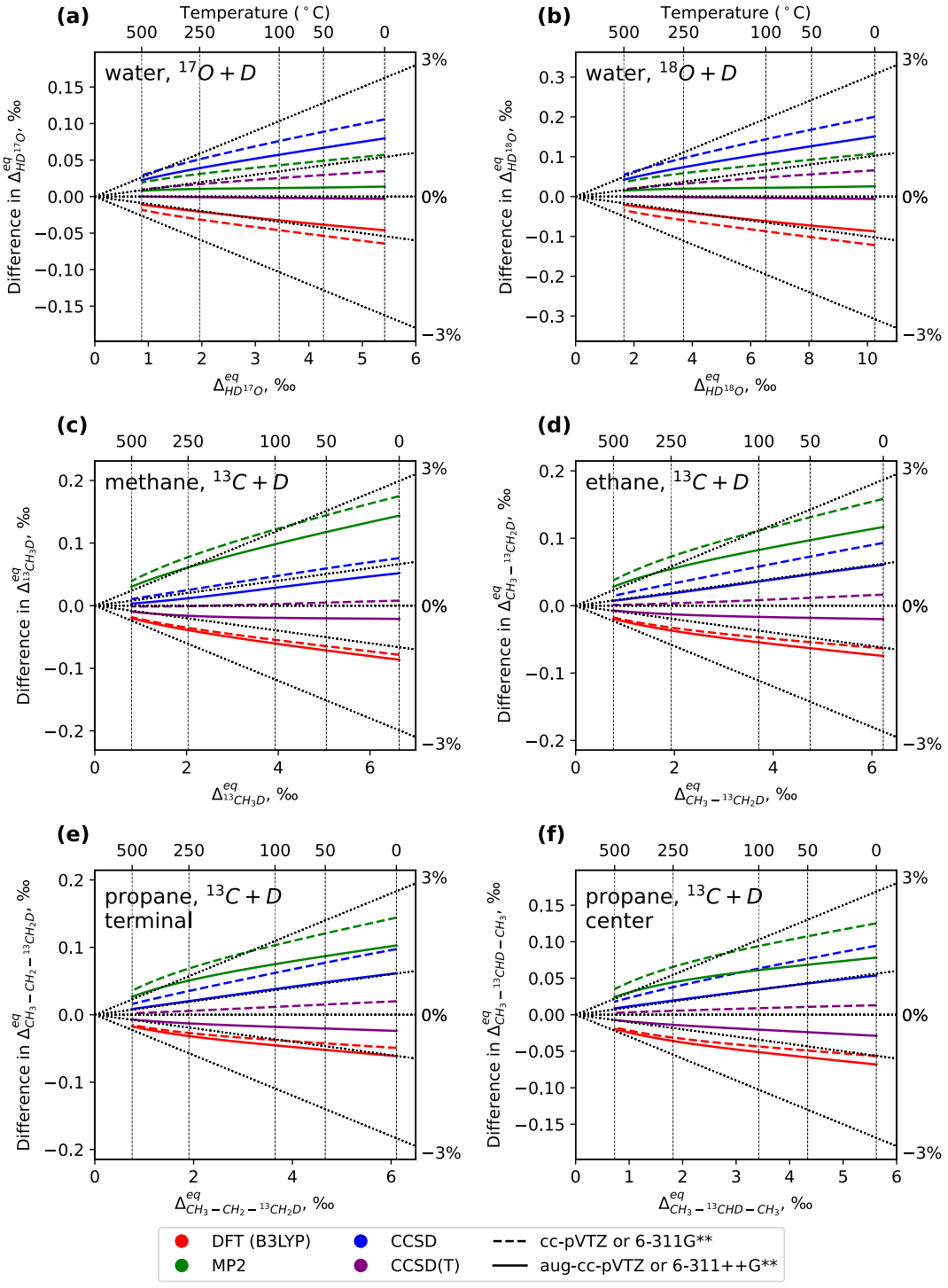

Fig. 5. Differences in clumped isotope effect involving deuterium directly bonded to a heavy atom (oxygen or carbon) computed with four methods at the triple-ζ basis set size. The panels are as follows: ^{17}O+D (a) and ^{18}O+D (b) in water, ^{13}C+D in methane (c), ethane (d) and propane (e-f). All computations are harmonic relative to the harmonic F12/ATZ method. Dotted vertical lines indicate temperatures (from 0°C to 500°C). The slanted dotted black lines denote relative difference (0, 1, and 3%).

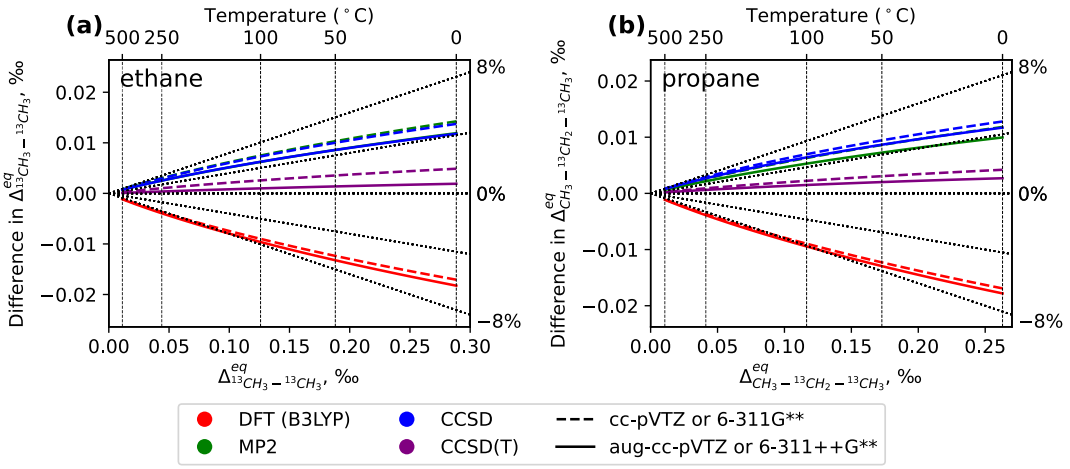

Fig. 6. Differences in $^{13}C+^{13}C$ clumped isotope effect for heavy isotopes directly bound to each other in ethane (a) and propane (b) computed with four methods at the triple-ζ basis set size. All computations are harmonic relative to the harmonic F12/ATZ method. Dotted vertical lines indicate temperatures (from 0°C to 500°C). The slanted dotted black lines denote relative difference (0, 4, and 8%).

important for DFT calculations with deviations up to 5–10% relative for D+D clumping in alkanes and O-D clumping in water (Fig. S8).

The empirical force fields (CHARMM, AIREBO) are inconsistent in terms of agreement with the reference value. They sometimes yield good agreement (<10% relative difference) with the reference (see Figs. S11c-f, S12), and sometimes do not (Fig. S9c-f).

4.2. Anharmonicity and the diagonal Born-Oppenheimer correction

In this section we provide results of calculations that incorporate anharmonicity and the diagonal correction to the Born-Oppenheimer approximation (Figs. 7-8) by comparing harmonic calculations (on the x-axis) vs. either DBO-corrected harmonic calculations or PIMC calculations. We control for the role of the PES calculation method in this section by using the same PES for all calculations based on F12/ATZ (i.e., our reference potential, see section 3.3.1).

4.2.1. Impact of the diagonal Born-Oppenheimer correction

The diagonal Born-Oppenheimer correction (DBOC) increases the fractionation of D between propane and methane ($1000\times\ln{}^{D}\alpha^{eq}_{C_3H_8-CH_4}$), ethane and methane ($1000\times\ln{}^{D}\alpha^{eq}_{C_2H_6-CH_4}$) as well as between alkanes and dihydrogen ($1000\times\ln{}^{D}\alpha^{eq}_{alkane-H_2}$) by up to +5‰ absolute, which corresponds to about +5% relative between alkanes (i.e.,; Fig. 7a) and up to +0.5% relative with dihydrogen (Fig. S13). Fractionations of D involving water are more substantially affected by the DBO correction, as exemplified by methane-water fractionation $1000\times\ln{}^{D}\alpha^{eq}_{CH_4-H_2O}$ (Fig. 7a) that is decreased by 10–30% relative or −10 to −35‰ absolute due to this correction (orange squares) over the range of temperatures studied. We see similar absolute effects for $1000\times\ln{}^{D}\alpha^{eq}_{H_2O-H_2}$ (Fig. S13). These calculated effects of DBOC on fractionation of D are consistent with our prior calculations in Turner et al. (2021). In contrast, the DBO correction changes ^{13}C fractionations between alkanes by less than −1% relative (less than −0.05‰ absolute; Fig. 7b and table S15).

Site-specific isotope effects in propane (table S16) are affected by the DBO correction to a similar extent as for fractionations between molecules. For C site preference, the DBO correction decreases $1000\times\ln{}^{13C}\alpha^{eq}_{center-terminal}$ by a little over −1% relative or −0.1 to −0.2‰ absolute (see Fig. 8b). For the deuterium site-specific effect, the impact of the DBO correction is more pronounced with $1000\times\ln{}^{D}\alpha^{eq}_{center-terminal}$ increasing by about 5% relative (or 5–10‰ absolute; see Fig. 8a). No results are presented for clumped isotope effects as they are not influenced by the DBO correction (Bardo and Wolfsberg, 1975).

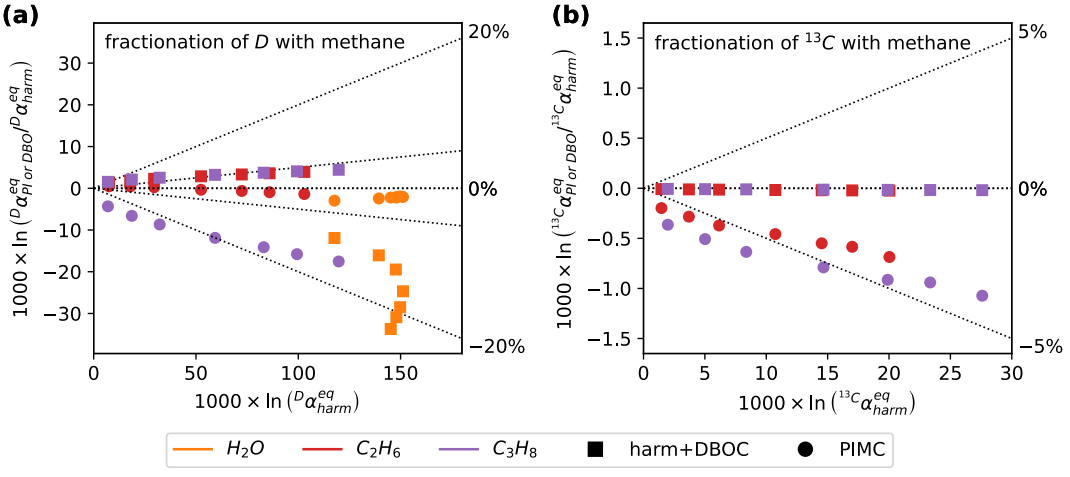

Fig. 7. Isolated effects of the diagonal Born-Oppenheimer correction (harm+DBOC, squares) and anharmonic contributions (PIMC, circles) vs. harmonic only calculations on the fractionation of deuterium (a) and (b) carbon-13 and with methane for water ($\alpha^{eq}_{H2O\text{-}CH4}$), ethane ($\alpha^{eq}_{C2H6\text{-}CH4}$) and propane ($\alpha^{eq}_{C3H8\text{-}CH4}$). The following temperatures are shown: 500°C, 300°C, 200°C, 100°C, 50°C, 25°C, and 0°C (left to right). The slanted black lines denote relative difference (0, 5 and 20%).

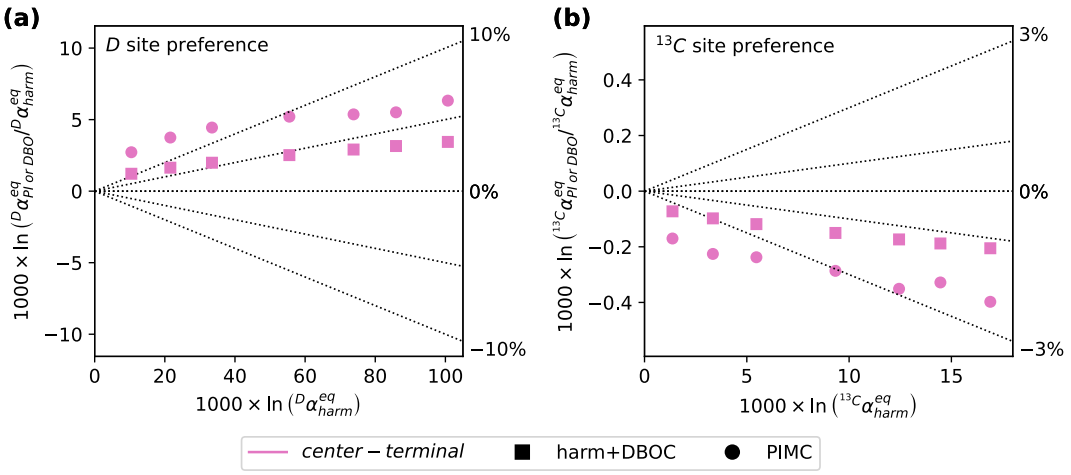

Fig. 8. Isolated effects of the diagonal Born-Oppenheimer correction (harm+DBOC, squares) and anharmonic contributions (PIMC, circles) vs. harmonic only calculations on the site-specific isotope effect of deuterium (a) and carbon-13 (b) in propane. The following temperatures are shown: 500°C, 300°C, 200°C, 100°C, 50°C, 25°C, and 0°C (left to right). The slanted black lines denote relative difference (0, 1, 3, 5 and 10%).

4.2.2. Impact of the PIMC calculations

We finally report the effects of including anharmonic terms by comparing the PIMC vs. harmonic calculations using the F12/ATZ potential energy surface (see tables S17-S19). We begin with fractionations between molecules relative to methane. We observe that the anharmonic effects decrease the fractionation of ^{13}C for $1000 \times \ln {}^{13C}\alpha^{eq}_{C_3H_8-CH_4}$ and $1000 \times \ln {}^{13C}\alpha^{eq}_{C_2H_6-CH_4}$ by about −5% relative (or up to −1‰ absolute, see Fig. 7b). Fractionation of D between methane and propane ($1000 \times \ln {}^{D}\alpha^{eq}_{C_3H_8-CH_4}$) is decreased significantly (by about −20% relative or up to −20‰), while other fractionations ($1000 \times \ln {}^{D}\alpha^{eq}_{H_2O-CH_4}$ and $1000 \times \ln {}^{D}\alpha^{eq}_{C_2H_6-CH_4}$) are affected less (up to −5% relative difference and less than −5‰ in absolute terms; Fig. 7a). Fractionations with dihydrogen show similar absolute trends: larger deviations for propane ($1000 \times \ln {}^{D}\alpha^{eq}_{C_3H_8-H_2}$) of up to −45‰ (about 3% relative) and smaller for methane ($1000 \times \ln {}^{D}\alpha^{eq}_{CH_4-H_2}$) and ethane ($1000 \times \ln {}^{D}\alpha^{eq}_{C_2H_6-H_2}$) of −10‰ to −20‰ (2% relative; see Fig. S13 and table S17). Note that for D fractionations between alkanes, the DBO and anharmonic effects have the opposite sign and therefore they partially cancel out when considered together.

For site-specific isotope effects in propane, the anharmonic effects (table S18) result in a +5% relative change (or +5 to +10‰) for $1000 \times \ln {}^{D}\alpha^{eq}_{center-terminal}$ (Fig. 8a) and −3% relative (−0.2 to −0.4‰ absolute) decrease in $1000 \times \ln {}^{13C}\alpha^{eq}_{center-terminal}$ (Fig. 8b). The anharmonic effects have the same sign as the DBO correction for both carbon and hydrogen and thus the two reinforce each other relative to the harmonic result.

For clumped isotope reactions, we observe anharmonicity effects change Δ^{eq} values most significantly for D+D in dihydrogen and water (Fig. 9 and table S19). The harmonic treatment underestimates both D+D clumped effect in dihydrogen up to −3.6‰ and in water by up to −1.3‰ with largest deviations at the lowest temperature (0°C). This corresponds to only a −1% relative difference for the harmonic result relative to the PIMC result for dihydrogen and −2% (at the lowest temperature) to −22% (at the highest temperature) for water. The full anharmonic treatment of all other clumped isotope effects is comparable to the harmonic result within statistical uncertainty of the path-integral calculations (see Fig. 10 and S14-S17) as described in the remainder of this paragraph. Based on our calculations and the sampling of the PIMC calculations, we observe that differences between anharmonic and harmonic treatments of clumping over the temperature range examined

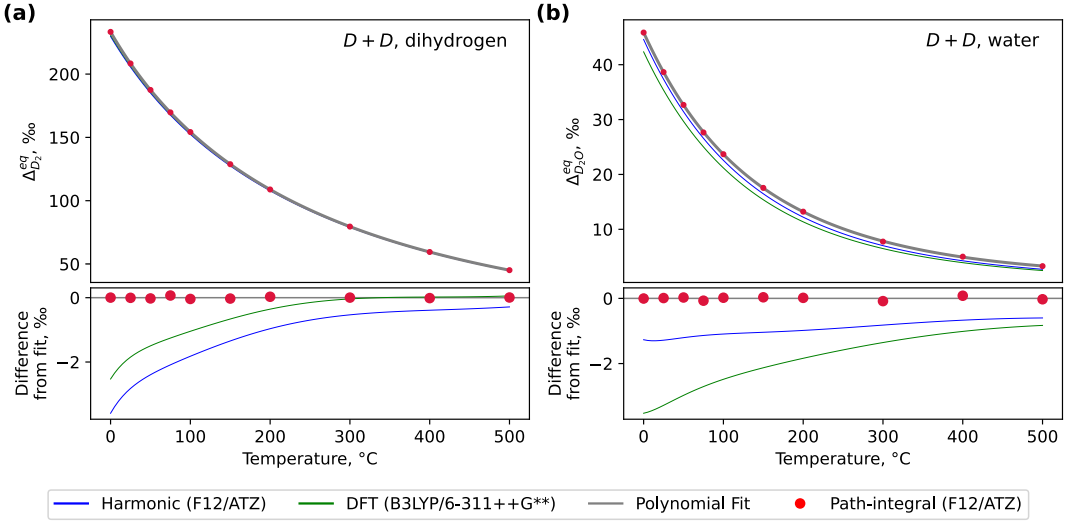

Fig. 9. Fifth order least squares polynomial fit (in gray) to the clumped D+D isotope effect in dihydrogen (a) and water (b) calculated with path-integral (red circles). The harmonic result with the same electronic structure method is shown in blue and the best DFT result obtained in this study in green. The bottom panels show the same data, but the polynomial fit is subtracted off to show differences from the PIMC calculation.

(0 to 500°C) are less than ±0.13‰ (i.e., up to ±2.2% relative) for ^{17}O+D and ^{18}O+D clumping effects in water (Fig. S13). Similarly, for the ^{13}C+D clumping in methane (Fig. 10a) and ethane (Fig. 10b) the effect is less than ±0.3‰ absolute (or less than ±6% relative). For the two ^{13}C+D clumping effects in propane, we see larger (up to ±1‰ absolute and up to ±20% relative; see Fig. S15) differences between the harmonic and PIMC-based results. Finally, D+D clumped effects in ethane and propane deviate by as much as ±1.2‰ in absolute sense (and up to ±20% relative, see Figs. S16 and S17). We discuss the implications of this below.

Finally, we note that all other clumped isotope effects in ethane and propane have small absolute magnitudes (i.e., less than 0.6‰ based on harmonic calculations) including ^{13}C+^{13}C effects as well as D+D effects in which the two deuterium atoms are not bound to the same carbon and ^{13}C+D effects where the heavy atoms are not directly bound to each other. The small size of these clumped effects means that the number of samples we have taken in our calculations does not yield useful differences in values of Δ^{eq} over the examined temperature range as the statistical uncertainty in the PIMC result is larger than the target of the calculation (i.e., Δ^{eq} values, see Fig. S18 and table S20).

5. Discussion of importance of different approximations in calculations of equilibrium isotope effects

A focus of this work is to quantify how various approximations do or do not influence final calculations of fractionation factors and clumped isotopic compositions in alkanes and other simple hydrogen bearing molecules (H_2 and H_2O). The key approximations that we explore and evaluate are the quality of the molecular potential used and the inclusion of anharmonic effects as well as the diagonal Born-Oppenheimer correction. In doing this evaluation, it is worth having a general sense of the current precision for measurements of fractionations discussed here as these set the minimum current capacity to both differentiate samples and evaluate if any calculated differences based on theory are measurable. This is a complex issue due to a range of different measurement techniques (many of which are evolving and improving in real time), and the issues of standardization and reproducibility of measurements. As such, a full discussion of this is beyond the scope of this paper and the following should be taken as a guide for an order of magnitude sense of the issue. Measured uncertainties for bulk carbon and hydrogen isotopes of alkanes and H_2 are generally <0.1‰ for ^{13}C and <2‰ for D (1 standard deviation; Dai et al., 2012). Water hydrogen isotopic measurements are typically made to <1‰ (1 standard deviation, e.g., Wassenaar et al., 2014). Site specific carbon and hydrogen isotopes can be measured to <1‰ and 10‰ (1 standard deviation) respectively (e.g., Gilbert et al., 2016; Piasecki et al., 2016a; Liu et al., 2018; Xie et al., 2018). Current Methane ^{13}C+D precisions are typically between 0.1 to 0.3‰ and D+D between 0.5 to 2‰ (Stolper et al., 2014; Ono et al., 2014; Young et al., 2016; Eldridge et al., 2019; Gonzalez et al., 2019). ^{13}C+^{13}C clumping in ethane has been measured to ~0.1‰ (1 standard deviation; Clog et al., 2018; Taguchi et al., 2020) while D+D and ^{13}C+D clumping in ethane and any clumping in propane have yet to be reported outside of conference proceedings (Anadu et al., 2024). D+D clumping in molecular hydrogen has been measured to precisions of ~2–7‰ (1 standard deviation; Popa et al., 2019; Mangenot et al., 2023). To our knowledge, neither clumped measurements of propane nor water have been reported.

5.1. Importance of the molecular potential

We begin by discussing the effects associated with approximations of the potential energy of the molecule as a function of molecular geometry. In Figs. 1-6 we presented how values of $1000 \times \ln\alpha^{eq}$ and Δ^{eq} differ using different levels of theory and basis set sizes relative to the reference (F12/ATZ) method within the harmonic approximation. Importantly, this comparison can only identify differences based on the harmonic approximation, i.e., those that arise due to the different curvature of the potentials at the minimal energy (equilibrium) geometry.

A key observation we take up and discuss here is that for fractionations calculated using the harmonic approximation between molecules, between sites, and clumping reactions, calculated results do not converge as the quality of the description of the molecular potential is increased from DFT (B3LYP functional) or the post-Hartree-Fock methods (MP2, CCSD, CCSD(T), (F12/ATZ). This is not because these potentials predict the same RPFRs for a given molecule, but rather, errors correlate such that these differences cancel out. This is important from a practical point of view as it allows for more accurate calculations of isotope effects at lower levels of theories. However, it also limits the ability to predict what level of theory is needed for a given calculation as the degree of error cancellation is not, to our knowledge, predictable at the outset.

Fig. 11 graphically illustrates these trends by showing relative error for fractionations (a-b), site-specific effect (c-d) and clumping (e-f) in alkanes involving deuterium (on the left) and carbon-13 (on the right).

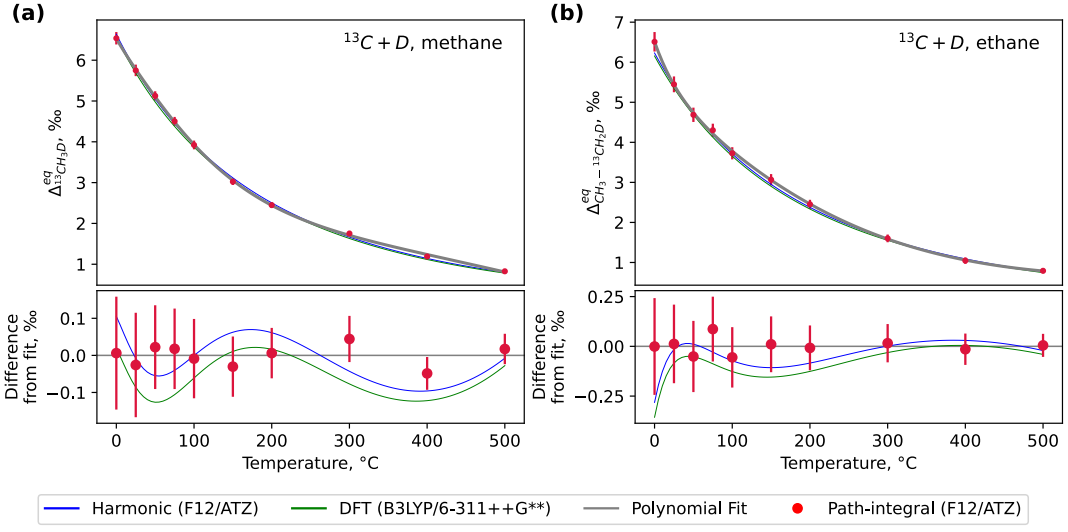

Fig. 10. Fifth order least squares polynomial fit (in gray) to the clumped ^{13}C+D isotope effect in methane (a) and ethane (b) calculated with path-integral (red circles). The harmonic result with the same electronic structure method is shown in blue and the best DFT result obtained in this study in green. The bottom panels show the same data, but the polynomial fit is subtracted off to show differences from the PIMC calculation.

Here, we plotted the frequency of the relative differences in fractionation factors and clumped isotope effects of the post Hartree-Fock methods used. For our purposes, we do not consider the distribution shape to be important, but rather the overall spread. Specifically, we observe that the relative differences span a range of ±40% for fractionation factors, including for site-specific isotope effects vs. the narrower range for clumped isotope effects of ±10% relative for D+D clumping (panel e) and ±5% relative for the ^{13}C+D clumping (panel f).

We first visually demonstrate the degree of error cancellation in Fig. 12 between molecules and for site preferences by comparing differences in the RPFR of a singly substituted species (on the x axes) to the corresponding differences in isotope fractionations and site preferences (the y axes). In Figs. 12 and 13, perfect cancellation of differences between two RPFR's corresponds to the horizontal line y = 0. Data plotting between the x-axis and a slanted line labelled as 'X%' has at least X% of the RPFR difference (relative to the reference) cancelled out.

We observe that for heavy isotope fractionation between molecules (Fig. 12a-d) and sites of propane (Fig. 12e-f) error cancellation varies from virtually no error cancellation between H_2 and methane calculated with CCSD to 98% of the difference in RPFRs between alkanes calculated with DFT cancelling out (Fig. 12c). We interpret the lack of error cancellation for fractionations with H_2 as being due to the high accuracy of the calculated RPFR for H_2 for all methods examined. Specifically, unlike all other molecules, dihydrogen only has two electrons and thus it is described much more accurately by both CCSD (same as CCSD(T) for H_2) and DFT (compare panel a to others in Fig. 4). As such, and perhaps counterintuitively, the simplicity of dihydrogen means that calculations of isotopic equilibria with other molecules using DFT, CCSD or CCSD(T) methods show virtually no error cancellation, as there is little to no error in the H_2 RPFR left to cancel.

Following typical chemical intuition, we observed that for hydrogen isotopes, error cancellations tend to be better when molecules are more alike (e.g., alkane vs. alkane), and worse when molecules have different bond types and properties (e.g., dipole moments) such as alkanes vs. water. This is indeed what we observe with alkane-water D fractionation differences of up to 60‰; see Fig. 2c-d). The differences in fractionation of D between alkanes are smaller (i.e., less than 20‰, Fig. 1a,c). As we show in Fig. 12, only about 20% of the differences in RPFRs cancel out for methane-water fractionation as described by both CCSD and MP2 (panel b). In contrast, for fractionation between alkanes (panel c) about 80% of the differences in RPFRs cancel out. Interestingly, error cancellation for fractionation of ^{13}C as well as site preference in propane is worse than for hydrogen isotopes, indicating perhaps that carbon atoms in methane vs ethane vs. propane are less alike than the corresponding

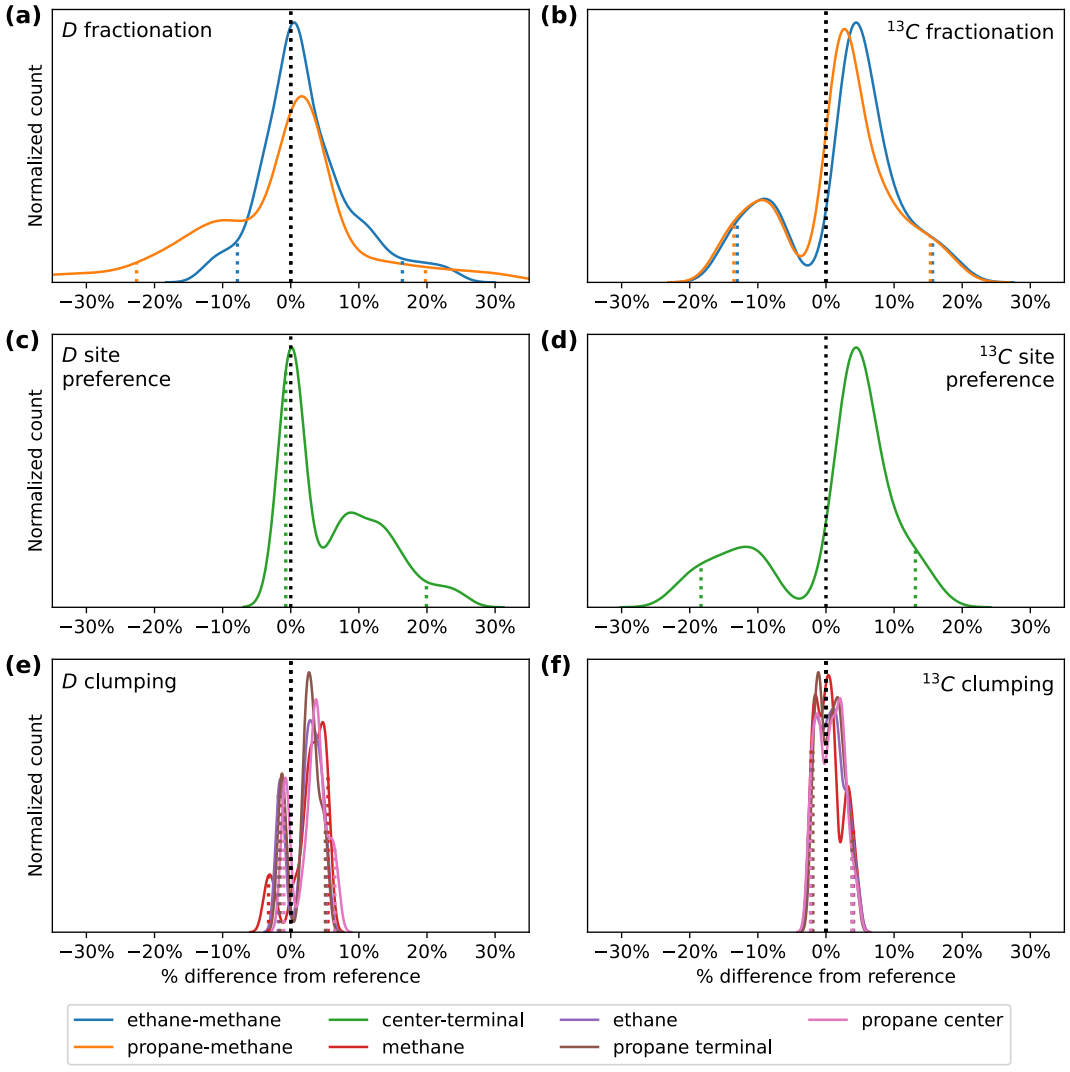

Fig. 11. Relative deviations from the reference value (dotted black line) for fractionations (a-b), position-specific (c-d) and clumped (e-f) isotope effects in alkanes involving deuterium (left panels) and carbon-13 (right panels). Harmonic calculations over the temperature range of 0–500°C are done using DFT (B3LYP) with 6–311G** and 6–311++G** basis sets as well as MP2, CCSD and CCSD(T) with cc-pVTZ and aug-cc-pVTZ basis sets each. Colored dotted lines are 5th and 95th percentile of the data on each distribution.

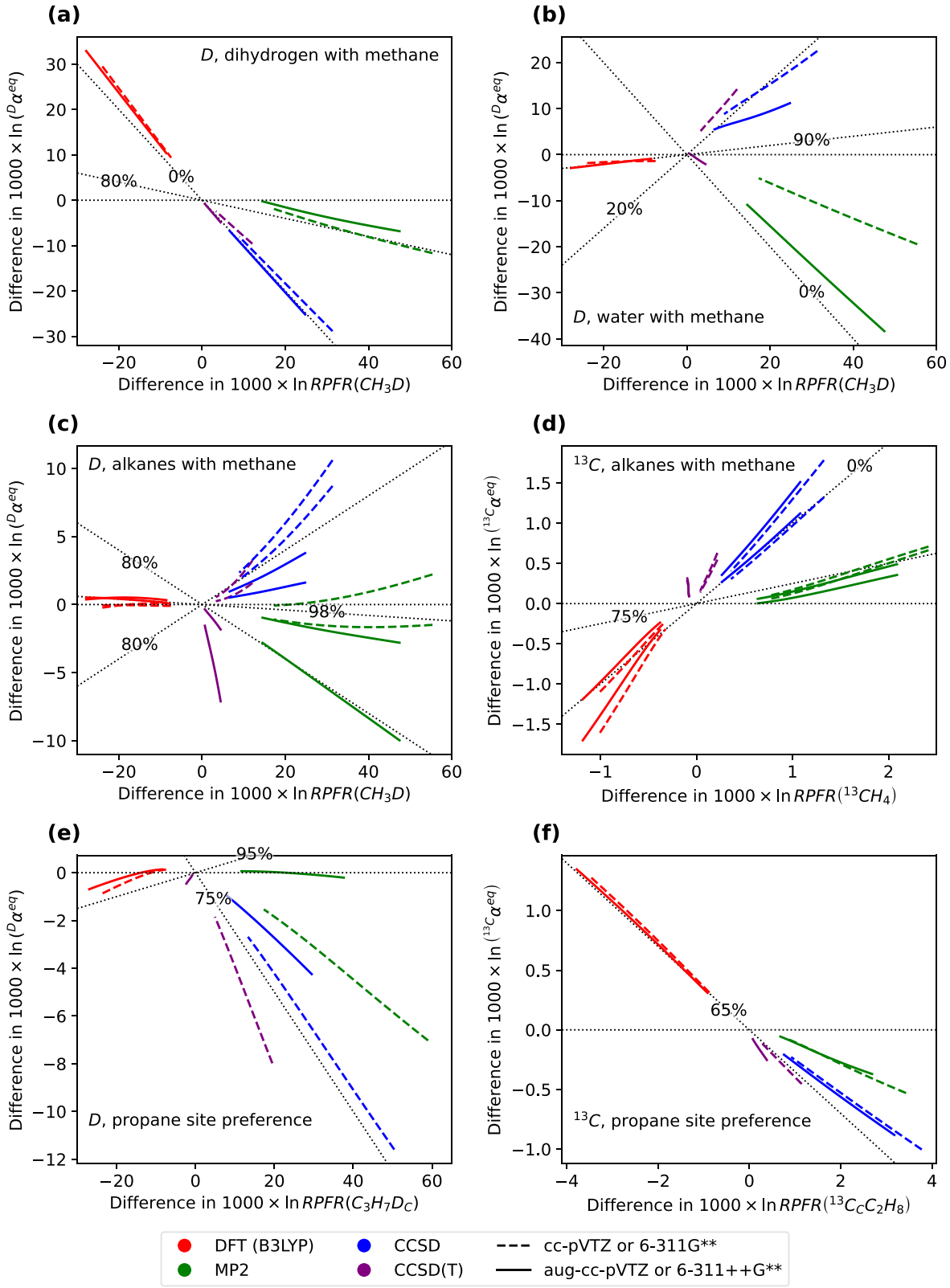

Fig. 12. Comparison of differences in RPFRs of singly substituted isotopologues (x-axis) to the differences in the corresponding fractionations (a-d) and site-specific isotope effects in propane (e-f) shown on the y-axis over the temperature range of 0 to 500°C. The differences are computed with four commonly used electronic structure methods and triple-ζ basis sets relative to the F12/ATZ method. Slanted dotted lines label the % of error that cancels out between RPFR's of singly and doubly substituted species when $1000 \times ln\alpha^{eq}$ is calculated via Eqs. (6) and (S38) with the horizontall line denoting 100% of error cancellation.

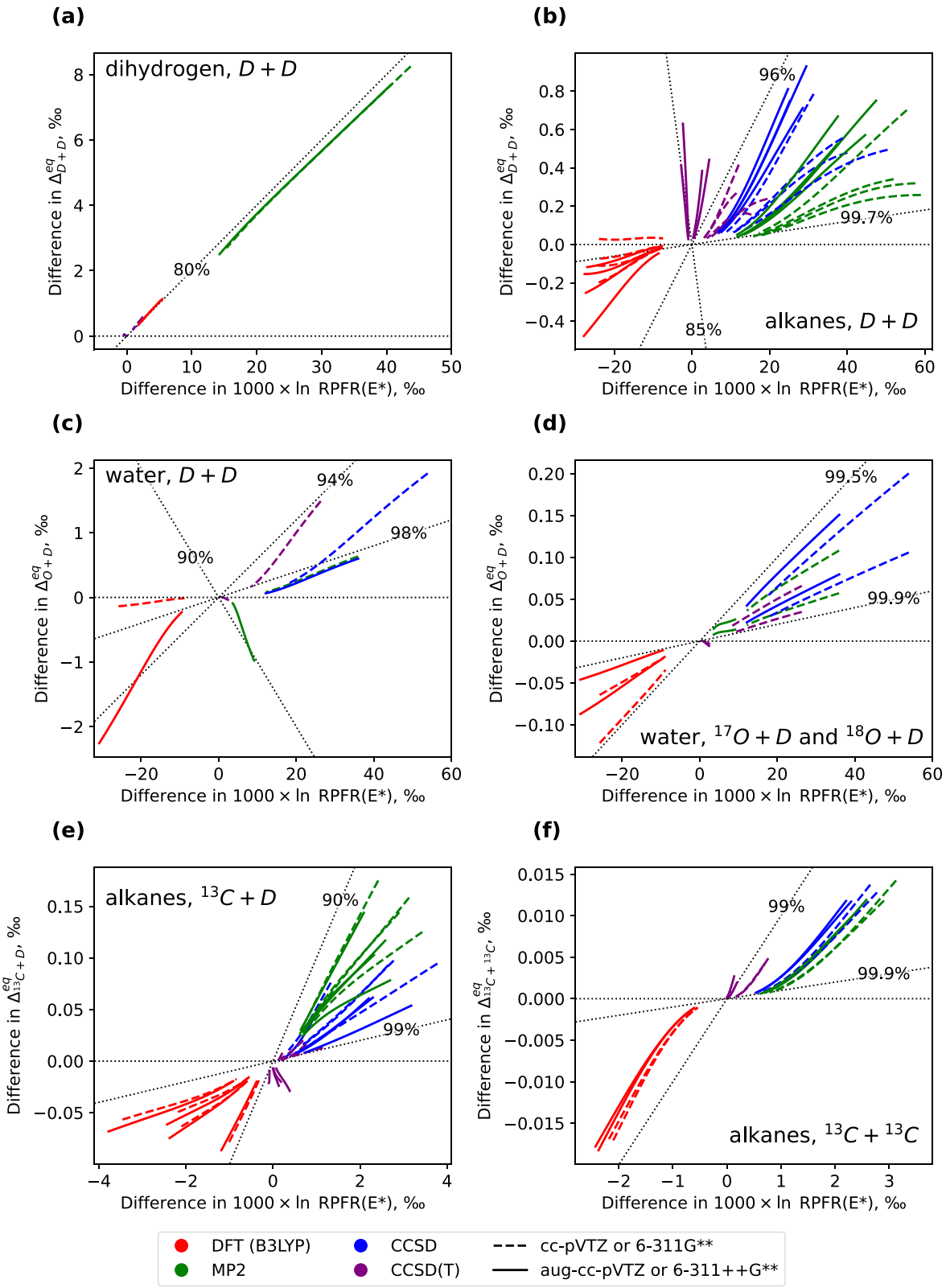

Fig. 13. Comparison of differences in RPFR of singly substituted isotopologue (x-axis) to the differences in the corresponding clumped isotope effects (y-axis) of dihydrogen (a), alkanes (b, e, f) and water (c-d), over the temperature range of 0 to 500°C. The differences are computed with four commonly used electronic structure methods and triple-ζ basis sets relative to the F12/ATZ method. Slanted dotted lines label the % of error that cancels out between RPFR's of singly and doubly substituted species when Δ^{eq} is calculated via Eqs. (5) and (S39) with the horizontall line denoting 100% of error cancellation. The RPFRs on x-axes are for the following singly substituted species: (a) *HD*, (b) CH_3D, C_2H_5D, $CH_3CH_2CH_2D$, CH_3CHDCH_3, (c-d) *HOD*, (e) $^{13}CH_4$, $^{13}CH_3CH_3$, $^{13}CH_3CH_2CH_3$, $CH_3^{13}CH_2CH_3$ (f) $^{13}CH_3CH_3$, $^{13}CH_3CH_2CH_3$.

hydrogen isotopes.

We now turn to clumped isotope effects shown in Fig. 13. We observe that clumped isotopes yield error cancellations of ~80% for H_2 and 90–99.9% for H_2O and alkanes. The more limited error cancellation for H_2 is likely again due to the fact that less error remains to cancel between methods as they are all accurate for this molecule. For the other molecules, clumped-isotope effects enjoy high error cancellation regardless of molecules type. In section S4 we demonstrate that this error cancellation can be illustrated through an analysis of differences in calculated zero-point energies of various isotopologues.

As such, we propose that it can be considered typical that most of the difference in RPFRs calculated at a given theoretical value will cancel out when clumped isotope effects are computed (at least for gases), provided the same method is used to describe both molecules. We interpret this to also be the reason why we do not observe consistently better agreement with the reference method when the basis set size is increased or when more sophisticated level of theory is used beyond MP2.

Consistent error cancellation in clumped-isotope calculations provides important practical considerations for future computational studies because it means that lower-cost calculations can potentially be used with a limited loss of accuracy, at least for the molecules studied here under the harmonic approximation (the validity of which we analyze in more detail in the next section).

5.2. Importance of the DBO correction

We now turn the importance of corrections to the Born-Oppenheimer approximation. Based on the results presented above, we propose that DBO corrections are of secondary importance when calculating isotope effects that do not involve deuterium. For example, the DBOC affects C fractionations between alkanes by less than −0.03‰ (−0.1% to −0.2% relative), Fig. 7b and table S15) while propane site-specific ^{13}C effects are affected by −0.1‰ to −0.2‰ (−1% to −2% relative; see Fig. 8b and table S16), all significantly less than current measurement precisions. Given these small differences and the computational cost of adding the DBO corrections, we do not consider them important for carbon isotope studies, especially given the observation that different molecular potentials yield much larger relative differences (up to ±10%) in the harmonic calculation.

In contrast, DBO corrections can matter for hydrogen isotope fractionations with effects on calculated fractionations with water changing by up to 35‰ (30% relative) (Fig. 7a and table S15). For other molecule pairs and site-specific effects, the DBO corrections are commonly ~ 5‰ (5% relative) and thus similar to or smaller than the changes associated with the choice of molecular potential or the inclusion of the anharmonic effects, provided the molecules have similar polarity (i.e., alkanes and dihydrogen). These observations are largely consistent with our earlier predictions on the importance of the DBO correction (Turner et al., 2021). Based on this, we propose that DBO corrections for hydrogen isotopes are of sufficient magnitude to necessitate consideration in the calculation of accurate fractionation factors between molecules and for site-specific effects as they are on the same order or larger than current measurement precision.

5.3. Importance of the anharmonic effects and consideration of when they should be included in calculations

Anharmonic effects are not always included in calculations of isotopic equilibrium, presumably because they are more difficult to incorporate vs. the harmonic treatment (especially for larger molecules), and accounting for them does not necessarily increase accuracy due to the need for additional considerations in the calculation (Webb and Miller, 2014). The question we wish to explore here is whether inclusion of anharmonic effects makes a sufficient difference for a given molecular potential that it should be included where possible.

We observe that inclusion of anharmonicity is generally of second order importance for calculating clumped isotope compositions. For methane it has already been shown that the path-integral-based calculations of clumping are in good agreement with the harmonic approximation (Webb and Miller, 2014; Eldridge et al., 2019). Fig. 10, S14-S17 and tables S19-S20 show that within the statistical error of the PIMC calculations, this is also true for ethane, and propane. For example, anharmonicity changes ^{13}C+D clumping by <0.002‰ on average over the studied temperature range and <0.03‰ for D+D clumping, which are both 10x smaller than current measurement precision for methane (ethane, propane, and water do not have published measurements yet to compare to). Although it is possible to reduce the statistical uncertainty in clumped isotope effects computed with the path-integral method with additional sampling, we do not attempt to do it here as these results do not indicate significant changes in clumped isotope effects due to the anharmonic effects. On this basis, we propose that for clumped isotope effects, one can in general use values based on the harmonic frequencies. However, for the two most pronounced (i.e., largest numerical value of Δ^{eq}) clumped isotope effects, which are D+D in dihydrogen and water, we do recommend using the path-integral based data, as the anharmonic effects lead to small but verified difference of up to 2‰ for water and up to 4‰ for molecular hydrogen (see Fig. 9, tables S19-S20), which are on the same order as current measurement precisions for dihydrogen of 2 to 7‰.

In summary, we find that clumped-isotope calculations achieve high error cancellation for DFT and post Hartree-Fock methods as well as when anharmonicity is included vs. is excluded. On this basis we propose that, except in situations in which clumped effects are especially large (e.g., hundreds of per mil for D_2), harmonic calculations using DFT or post Hartree-Fock methods are likely sufficient to obtain high-accuracy results. We especially recommend using harmonic calculations over PIMC for clumped isotope effects that are small (i.e., involving either nonhydrogen isotopes not directly bound to each other or hydrogen isotopes separated by more than one other atom), as the statistical uncertainty of the PIMC calculation approaches the size of the effect unless significant computational resources are directed to the calculation (see Fig. S18 and table S20). A possible route towards reducing the statistical uncertainty using higher order action in path integral calculations was suggested and tested by Zhang et al. (2024).

In contrast, we observe that the use of the harmonic approximation vs. including anharmonic effects can result in significant changes in calculated equilibrium fractionations between different molecules, especially for deuterium, with differences up to 20‰ from 0 to 500°C for the molecules studied here. These differences are sufficiently large that the sign of enrichment changes for ethane vs. propane over the studied temperature range. Specifically, our harmonic calculation predicts that propane is enriched in D with respect to ethane below the crossover temperature of 412°C (i.e., $1000\ln(\alpha) = 0$; see Stern et al., 1968 for more detail on crossovers). In contrast, PIMC prediction for the crossover temperature is 400 degrees lower, i.e. $1000\ln(\alpha) = 0$ at 12°C such that above this crossover temperature propane becomes lighter than ethane (see table S17). Including anharmonic effects for fractionation of ^{13}C is moderately impactful as well, changing the result by about 5% relative (up to 1‰).

The inclusion of full anharmonic treatment changes the site-specific isotope effect in propane by about +5% relative (about +5‰ absolute) for deuterium and −1% relative (i.e., −0.3‰ absolute) for carbon-13, both with respect to the harmonic result (see Fig. 8 and table S18). These effects are larger than the respective DBO corrections, but smaller than the differences due to the molecular potential (Fig. 3).

On this basis, we consider inclusion of PIMC to be important for calculations of fractionation factors between alkanes and other hydrogen bearing molecules when ±5 to 20% relative accuracy is needed, especially when molecules have multiple sites (e.g., propane and larger molecules).

6. Discussion of recommended values to use for calculated isotope effects based on this study

6.1. Recommended calculated isotope effects for use in geochemical applications

Here we present our recommended values for stable isotope fractionations of alkanes, water, and dihydrogen as a function of temperature as informed by the discussion given above. For all calculations, we recommend using results based on F12/ATZ PES as they are typically considered to be the most accurate among the potentials we used in this study (Ten-no and Noga, 2012).

We recommend use of calculations for isotope fractionations between molecules and site preference in propane that incorporate PIMC (section 3.3) based on the F12/ATZ potential energy surfaces (section 3.3.1) and adjusted by the DBOC using Eq. (10). We provide $1000 \times \ln\alpha^{eq}$ values in tables S21-S22. The 5th order interpolative polynomial fit coefficients for $1000 \times \ln RPFR(Z^*)$, Eq. (11) which are needed to calculate ^{13}C and D fractionations as well as the site-specific isotope effect in propane are given in Tables 1 and 2 and S23. We fit RPFRs over fractionation factors to reduce the number of equations needed. The best fit polynomials produce fits with residuals of <0.03‰ vs. the calculated data. A fifth order fit is used in all cases as a compromise between faithful representation of data and avoidance of the risk of overfitting (i.e., fitting noise) with the following form:

$$1000 \times \ln RPFR(Z^*) \approx A + B\left(\frac{1000}{T}\right) + C\left(\frac{1000}{T}\right)^2 + D\left(\frac{1000}{T}\right)^3 + E\left(\frac{1000}{T}\right)^4 + F\left(\frac{1000}{T}\right)^5 \quad (11)$$

For convenience, we also provide calculators for each pair of molecules in table S23.

For the clumped isotope effect calculations, we recommend the use of the F12/ATZ potential energy surfaces within the harmonic approximation (Eq. (8)) except for D+D clumping in H_2 and H_2O. For these, we recommend use of clumped isotope effects based on PIMC as the anharmonic effects are sufficiently large as to modify the final results relative to the harmonic calculations. We provide the recommended values in table S24 and the 5th order fits and calculators of Δ^{eq} for all clumped effects based on these recommendations in tables 3 and S25.

6.2. Discussion of our recommended results

We now discuss a few specific aspects of the recommended values based on this study and compare them to other theoretical studies where applicable.

6.2.1. Trends in heavy isotope fractionation among alkanes

Recent theoretical calculations for equilibrium isotope effects based on a harmonic approximation and DFT-level theory are presented in Thiagarajan et al. (2020) and Xie et al. (2024) (with both following the same framework). They predict that for hydrogen isotopes, methane will always be lighter than ethane, which in turn will be lighter than propane with no cross overs (i.e., with a monotonic approach to no fractionation at infinite temperature).

In contrast, our calculations predict a difference from typical expectations in terms of the isotopic ordering between small alkanes for hydrogen isotopic equilibrium. As highlighted on Fig. 14(a) and noted above, based on the path integral calculations we predict a crossover in bulk fractionation between ethane and propane at 12°C; it is only at temperatures lower than this that ethane is lower in D/H vs. propane. This crossover is not observed for carbon-13 (Fig. 14b), as in this case both sites of propane are heavier than ethane. As such, we propose that instances in which propane is lighter than ethane by up to ~5‰ may be a fingerprint for hydrogen isotopic equilibrium being reached between these molecules at temperatures relevant to natural gas generation and storage (e.g., ~50–300°C). Harmonic calculations with the F12/ATZ potential and the DBO correction yield a much higher crossover temperature of 412°C. This shows the importance of including rigorous methods like PIMC for interpreting equilibrium isotope effects involving hydrogen. We discuss this in more detail in the part 2 of this contribution (Turner et al., 2025) in which experimental calibrations of methane-ethane-propane hydrogen isotopic equilibrium are presented as well as comparisons to environmental data.

6.2.2. The site-specific effect in propane

Equilibrium site-specific isotope effects for propane, both for carbon and hydrogen. have been calculated previously using PIMC (Webb and Miller, 2014; Cheng and Ceriotti, 2014) and based on the harmonic approximation by Piasecki et al. (2016b). Xie et al. (2018) summarized these theoretical results and compared them to their experimental data (right-hand panel of Fig. 6 from Xie et al., 2018). These comparisons were all done in the manner we employed here, i.e., with the isotopic difference between sites given. There are two clear disagreements between different theoretical calculations and between some theoretical calculations and data, which we now address given our results presented above. First, Xie et al. (2018) observed that the site preference for deuterium in propane calculated in Cheng and Ceriotti (2014) using PIMC based on AIREBO force field yielded a different sign of fractionation (depleting the center site of deuterium) compared to both their experiments, as well as harmonic calculations given in Piasecki, et al. (2016b), and PIMC calculations given in Webb and Miller (2014). Second, Xie et al. (2018) observed that their experimental results agreed best with the PIMC results of Webb and Miller (2014) and that these theoretical results disagree with those of Piasecki et al. (2016b) by up to 30‰. Xie et al. (2018) interpreted this difference as due to differences in computational methodologies employed. These differences were also presented in Fig. 5 of Piasecki et al. (2016b).

Given this past disagreement, we specifically looked to compare our PIMC calculations to prior work (Fig. 15, table S26). In compiling the past data, we found the results from Webb and Miller (2014) differ by up to 25‰ from that plotted in Xie et al. (2018). We observed that PIMC and harmonic results of Webb and Miller (2014), harmonic results of Piasecki et al. (2016b), and our results here, agree within 5‰ over the temperature range covered here (0–500°C), consistent with the sorts of differences we observed above for use of different molecular potentials, DBOC, and inclusion of anharmonicity. This specific issue as identified

Table 1
5th order interpolative polynomial fit coefficients for 1000xln(RPFR[single deuterated/unsubstituted]) in the form of Eq. (11) as a function of 1000/T (K^{-1}) based on PIMC calculations at CCSD(T)-F12A/aug-cc-pVTZ level.

molecule (site)	A	B	C	D	E	F
dihydrogen	−119.623851	353.6872971	40.35399998	−15.80555474	3.034618147	−0.230471732
water	−200.912214	501.4586186	231.3391294	−74.46786999	12.20932606	−0.807987429
methane	−80.8960615	157.8891624	420.5002459	−127.9279883	20.30959092	−1.32296931
ethane	−68.912306	126.3321409	452.3711583	−137.1242268	22.24062909	−1.495958204
propane (terminal)	−78.7778752	148.2027496	423.7966543	−124.610809	19.65663571	−1.290163833
propane (center)	−92.3678522	182.6487517	396.3041832	−105.7814834	15.06961491	−0.884931568
propane (both sites)	−82.7113601	158.1243055	415.6881108	−119.3481254	18.40024142	−1.17987188

Table 2
5th order interpolative polynomial fit coefficients for 1000 ln(RPFR[single carbon-13/unsubstituted]) in the form of Eq. (11) as a function of 1000/T (K^{-1}) based on PIMC calculations at CCSD(T)-F12A/aug-cc-pVTZ level.

molecule (site)	A	B	C	D	E	F
methane	−7.17911082	16.79338501	10.32542366	−2.050921327	0.144117639	0.002915318
ethane	−4.1342579	8.57018664	17.40371529	−3.744213772	0.413830438	−0.018107605
propane (terminal)	−0.77945316	0.739860462	24.20712143	−6.58583549	1.009210131	−0.066384283
propane (center)	2.22019309	−7.05398733	30.76188509	−8.045780654	1.192090798	−0.077072045
propane (both sites)	0.21916055	−1.85392827	26.38768208	−7.071112118	1.070355102	−0.069985935

Table 3
5th order interpolative polynomial fit coefficients for Δ as a function of 1000/T (K^{-1}) in the form of equation (11). For dihydrogen and water these are based on PIMC calculations at CCSD(T)-F12A/aug-cc-pVTZ level. All other molecules are harmonic at the same level of theory. The number of pluses signifies the number of atoms between the atoms involved in clumping — no plus is a direct bond, one plus indicates one atom between, etc.

Molecule (site)	Clumped effect	A	B	C	D	E	F
dihydrogen	DD	−22.9171158	19.32480879	37.5314004	−10.85221923	1.412236928	−0.058889858
water	D+D	11.86191226	−26.308564	21.3337278	−6.008993433	1.120101888	−0.094244668
methane	D+D	−0.155391326	1.361281879	−2.980173286	2.818265911	−0.603647015	0.044044301
ethane	D+D	0.236094112	0.423727258	−2.176668255	2.563110487	−0.600686905	0.046537571
propane (terminal)	D+D	0.213699495	0.470577794	−2.203264175	2.556184616	−0.598117011	0.046241949
propane (center)	D+D	0.555184862	−0.38109047	−1.423413905	2.267025297	−0.575582395	0.046799932
ethane	D++D	0.003557495	0.023986048	−0.073235605	0.072583417	−0.017551959	0.001513357
propane	D++D	0.00164519	0.025878752	−0.070855346	0.067368695	−0.015961092	0.001329286
propane	D+++D	0.002534542	−0.00691268	0.007130186	−0.003069095	0.000826635	−6.71887E-05
water	^{17}OD	0.066018009	−0.9505354	1.84127978	−0.6001325	0.100976408	−0.006902001
water	^{18}OD	0.131988667	−1.81334586	3.491302169	−1.137185544	0.191249688	−0.013067962
methane	^{13}CD	0.604847919	−2.12164262	2.517947289	−0.721981412	0.113670169	−0.007514097
ethane	^{13}CD	0.654626318	−2.2482802	2.641893634	−0.801859145	0.130982986	−0.008799347
propane (terminal)	^{13}CD	0.656159561	−2.24315923	2.627127203	−0.798343197	0.130056463	−0.008726055
propane (center)	^{13}CD	0.701233276	−2.37222671	2.772194225	−0.88543448	0.146869611	−0.009865762
ethane	^{13}C+D	−0.024933202	0.079453154	−0.102617049	0.06698258	−0.012574611	0.00083102
propane	^{13}C++D	−0.017337035	0.060149759	−0.084522948	0.060059573	−0.012018076	0.000834932
propane (terminal D)	^{13}C+D	0.000335572	−0.00089942	0.000918107	−0.000422472	0.000245968	−2.27338E-05
propane (center D)	^{13}C+D	−0.020798513	0.069618026	−0.0944391	0.064694286	−0.012449697	0.000841209
ethane	^{13}C^{13}C	−0.01573669	0.049155245	−0.062259626	0.039758916	−0.007203778	0.000459061
propane	^{13}C^{13}C	−0.012629472	0.041218358	−0.054554522	0.036333208	−0.006796486	0.000447018
propane	^{13}C+^{13}C	−0.000377749	0.001009683	−0.001052631	0.000525466	4.02878E-05	−8.55275E-06

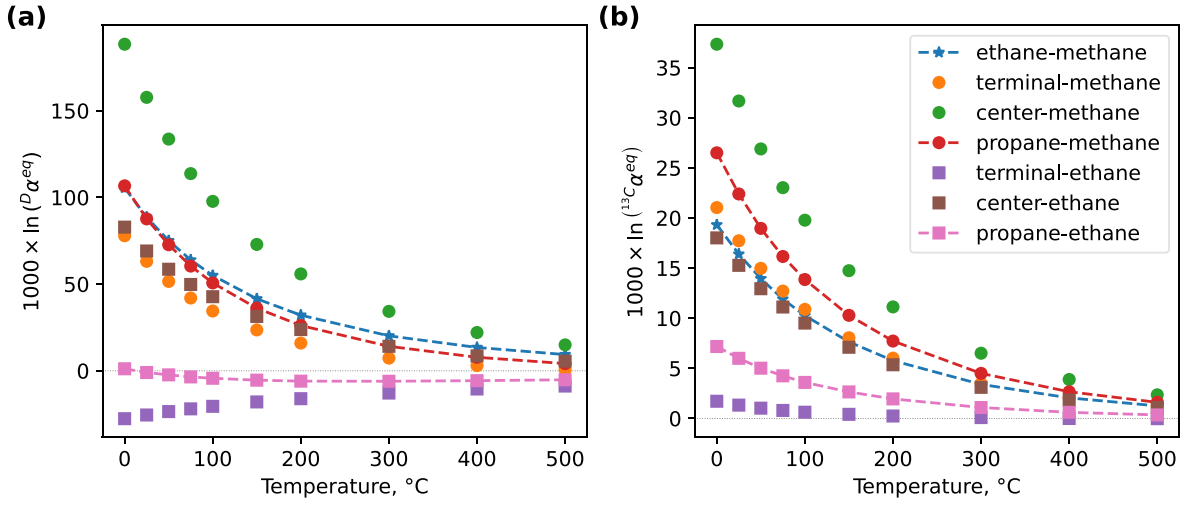

Fig. 14. (a) Deuterium and (b) carbon-13 fractionation between alkanes calculated with PIMC and after applying the DBO corrections. Fractionations of ethane with methane are stars, propane with methane are circles, and propane with ethane are squares. Dashed lines signify total fractionations.

by Xie et al., (2018) is likely definitional. Specifically, the study here, Xie et al. (2018) and Piasecki et al. (2016b) all provide determinations of site-specific effects that quantify difference in isotopic composition between the sites. In contrast, as we indicated in section 2.1, the site-specific values reported by Webb and Miller (2014) are given as enrichments of the center position relative to random distribution of isotopes – the same nomenclature as is typically used for the clumped isotope effect. Once Δ_i values given in Webb and Miller (2014) are converted to site-specific effects, the disagreements largely disappear. This finding is in agreement with a recent independent examination of this issue by Yin et al. (2024), whose calculations are in excellent agreement with ours. For example, for deuterium our result differs by ~2.5‰ relative to theirs (see Fig. 15 and table S26) and for carbon the difference is on the order of 0.5‰. We also agree on the direction of the

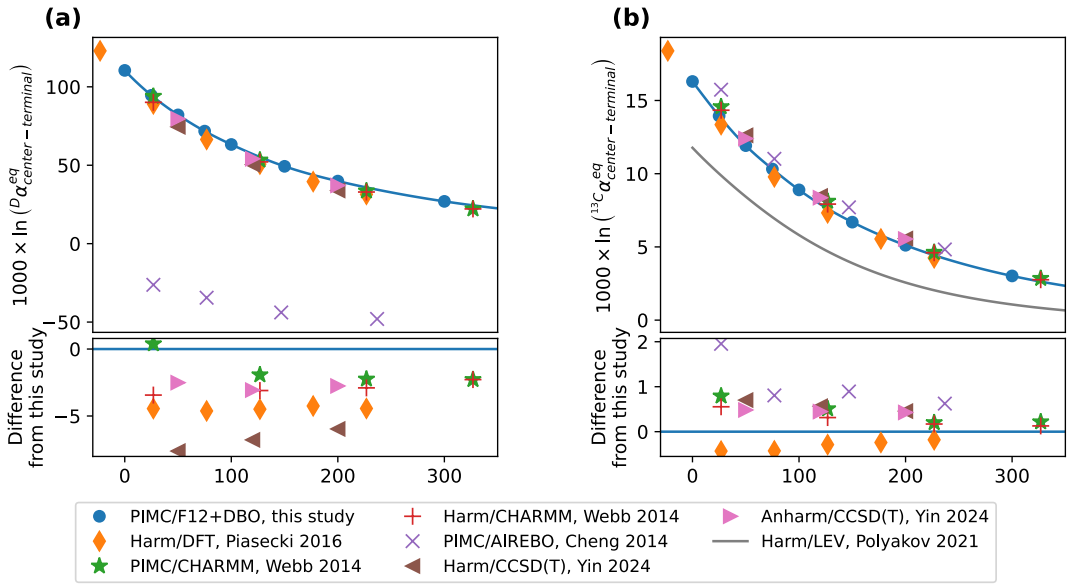

Fig. 15. Calculations of the site-specific isotope effect in propane from different studies for (a) hydrogen and (b) carbon. All results are presented as $1000 \times ln\,\alpha^{eq}_{center-terminal}$ to keep consistent with the rest of this study. In (a), we do not show the differences for the AIREBO calculations as they are off-scale.

shift due to anharmonicity and the DBOC. The remaining small differences are likely due to different basis set choice (we employed the Dunning-style as opposed to Popple-style basis set) and the use of the F12A coupled cluster variant.

We now turn our attention to the calculations performed by Cheng and Ceriotti (2014) with PIMC based on the AIREBO force field. Their results are generally in agreement with ours and others for ^{13}C (Fig. 15b), but they predict the opposite site preference for D (Fig. 15a). We interpret the disagreement with both our result based on coupled-cluster quality potential and other calculations as potentially arising from the use of AIREBO molecular potential. We did not repeat the PIMC calculations with AIREBO, but our harmonic result (Fig. S7) predicts that center site of propane is enriched with deuterium to a much higher extent than the reference harmonic method does (i.e., over +80% relative difference) over the temperature range studied. This is opposite the sign given in Cheng and Ceriotti (2014) and our recalculation is consistent with all other calculations of which we are aware. Based on results above, we consider it unlikely that anharmonic effects (i.e., in going from the harmonic result to the PIMC result) can shift the fractionation of deuterium by so much (e.g., at 0°C the shift would have to be over −200‰). We are unsure of the origin of this issue. But we note that during transfer from the Fortran-based implementation of AIREBO to that used here, various typographical errors were found and corrected for by Höhnerbach and Bientinesi (2018) relative to the code that was used in Cheng and Ceriotti (2014). It may be such changes resulted in the observed differences. Full resolution of this would require comparison of the PIMC-based calculations using the original Fortran AIREBO code to generate the molecular potentials, which was beyond the scope of this study.

7. Conclusions

We now summarize our key observations for the calculation of equilibrium isotopic fractionation including between different molecules, clumping, and site-specific effects in gaseous dihydrogen, water, and small alkanes from 0 to 500°C. We expect that many of these conclusions are applicable to other molecules as well and can therefore guide future studies.

1. Generally, isotopic fractionations between different molecules and site-specific isotope effects computed with DFT (B3LYP functional) or the post-Hartree-Fock methods (CCSD(T), CCSD and MP2) are within ±10% relative difference, while clumped isotope effects are typically within ±5% from the reference method (F12/ATZ).
2. We do not observe a uniform trend in terms of convergence of values with lower quality methods differing more vs. higher quality methods. In particular, the (lower quality) MP2 results are typically closer to the reference F12/ATZ values than the (higher quality) CCSD results. This highlights the unpredictable nature of error cancellations when RPFRs are compared to find a fractionation factor or clumped isotope effect.
3. We take 1. and 2. above to indicate that at the levels of theory commonly used in isotope geochemistry, one should anticipate about ±5–10% relative deviation if a different molecular potential is used. An additionally important point we wish to make here is that it is the relative differences as opposed to absolute differences that appear consistent. Thus, the 5% relative difference for a large fractionation factor (e.g., 1000‰) would result in meaningful differences in terms of measurements, but perhaps not for a smaller fractionation factor (e.g., 0.5‰) unless exceptional accuracy (sub-0.01‰) is needed.
4. The DFT/B3LYP is the method with the smallest computational cost out of those listed in 1. and its performance, for intercomparison of harmonically derived fractionation factors, is generally adequate (except for fractionations involving water, which also require larger basis set size). Based on this, DFT/B3LYP calculations with 6–311++G** basis set is likely sufficient if relative uncertainties of ±5 to ±10% are acceptable. We note that DFT/B3LYP calculations of equilibria involving carbon-13 (bulk fractionation, site-specific and ^{13}C+D as well as ^{13}C+^{13}C clumping) tend to show the largest relative deviations (i.e., about ±10%) compared to the reference result.
5. We do not recommend decreasing the quality of the molecular potential to empirical force-field (AIREBO, CHARMM) or RHF methods. These can yield differences in isotopic equilibria from those predicted by the higher quality methods by more than 75%, overwhelming all other approximations considered in this study.
6. Clumped heavy isotope effects enjoy significant error cancellation and can be computed to accuracies similar to or less than typical measurement precisions using a simple Bigeleisen-Mayer-Urey harmonic approach and a computationally inexpensive DFT potentials like B3LYP/6–311G**.
7. On the contrary, fractionation of heavy isotopes between alkanes, water, and dihydrogen as well as site-specific isotope effect in

propane do benefit from the use of high-quality potentials and the accounting of anharmonic effects.
8. DBO corrections are negligible for equilibria involving carbon-13. However, they can be important for the fractionations and site-specific effects involving deuterium.

CRediT authorship contribution statement

Roman Korol: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Andrew C. Turner:** Writing – review & editing, Methodology, Investigation. **Apurba Nandi:** Writing – review & editing, Software, Methodology. **Joel M. Bowman:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition. **William A. Goddard:** Writing – review & editing, Supervision, Resources, Funding acquisition. **Daniel A. Stolper:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization.

Data availability

Data are available through Mendeley Data at https://doi.org/10.17632/2kn87g7d46.1.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

DAS acknowledges support from the US Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, under Award Numbers DE-AC02-05CH11231 and DE-SC0022949. WGIII acknowledges support from the National Science Foundation under award number CBET-2311117 The computations presented here were conducted in the Resnick High Performance Computing Center, a facility supported by Resnick Sustainability Institute at the California Institute of Technology. RK thanks Dr. Tomislav Begušić for helpful discussions.

Appendix A. Supplementary material

We attach two supplementary files. The first file provides details on the electronic structure, path integral and the equilibrium isotope calculations as well as our analysis of error cancellation in harmonic calculations based on the zero-point (vibrational) energies. Supplementary Figure S1 shows rotamers of ethane and propane, Figures S2-S12 present comparison of performance of the less computationally expensive molecular potentials (based on CHARMM and AIREBO force fields, restricted Hartree-Fock) as well as the post-Hartree-Fock and DFT/B3LYP potentials with smaller basis sets. Fig. S13 illustrates the effects of anharmonicity and DBOC on fractionation with dihydrogen. Figs. S14-S17 present polynomial fits to PIMC calculations of clumped effects and Fig. S18 illustrates numerical difficulties of calculating small clumped isotope effects with PIMC. This file also includes supplementary tables S1-S6 which provide details and validation of the computational methodology. The second file is an Excel worksheet with various supplementary tables. These supplementary tables provide precise numerical values for all computations (tables S7-S11) and comparisons (tables S12-S20) performed in this study. We also give our recommended values for fractionations (table S21), site preferences (table S22) and clumped effects (table S24). Tables S23 and S25 contain interpolative polynomial fit coefficients as well as ready-to-use calculators for any temperature in the interpolative range. Finally, we provide values for site preference in propane compiled from the literature in table S26. Supplementary material to this article can be found online at https://doi.org/10.1016/j.gca.2025.02.028.

References

Adler, T.B., Knizia, G., Werner, H.J., 2007. A simple and efficient CCSD(T)-F12 approximation. J. Chem. Phys. 127, 221106.
Anadu J. S., Thiagarajan N., Kitchen N., Shuai Y., Warr O., Lollar B. S. and Eiler J. (2024) Insights into ethane creation, evolution, and destruction from clumped isotopologues. In 2024 Goldschmidt Conference. GOLDSCHMIDT.
Bardo, R.D., Wolfsberg, M., 1976. A Theoretical Calculation of the Equilibrium Constant for the Isotopic Exchange Reaction between H_2O and HD. J. Phys. Chem. 80, 1068.
Bardo, R.D., Wolfsberg, M., 1978. The adiabatic correction for nonlinear triatomic molecules: Techniques and calculations. J. Chem. Phys. 68, 2686–2695.
Bardo, R.D., Wolfsberg, M., 1975. The nuclear mass dependence of the adiabatic correction to the Born–Oppenheimer approximation. J. Chem. Phys. 62, 4555–4558.
Becke, A.D., 1993. Density-functional thermochemistry. III. The role of exact exchange. J. Chem. Phys. 98, 5648.
Bigeleisen, J., Mayer, M.G., 1947. Calculation of Equilibrium Constants for Isotopic Exchange Reactions. J. Chem. Phys. 15, 261–267.
Blanchard, M., Balan, E., Schauble, E.A., 2017. Equilibrium fractionation of non-traditional isotopes: A molecular modeling perspective. In Non-Traditional Stable Isotopes Walter De Gruyter GmbH. 27–64.
Born, M., Oppenheimer, R., 1927. Zur Quantentheorie Der Molekeln. Ann. Phys. 389, 457–484.
Cheng, B., Ceriotti, M., 2014. Direct path integral estimators for isotope fractionation ratios. J. Chem. Phys. 141, 244112.
Clog, M., Lawson, M., Peterson, B., Ferreira, A.A., Santos Neto, E.V., Eiler, J.M., 2018. A reconnaissance study of 13C–13C clumping in ethane from natural gas. Geochim. Cosmochim. Acta 223, 229–244.
Csernica, T., Sessions, A.L., Eiler, J.M., 2023. High-dimensional isotomics, part 2: Observations of over 100 constraints on methionine's isotome. Chem. Geol. 642, 121771.
Dai, J., Xia, X., Li, Z., Coleman, D.D., Dias, R.F., Gao, L., Jian, L.i., Deev, A., Jin, L.i., Dessort, D., Duclerc, D., Li, L., Liu, J., Schloemer, S., Zhang, W., Ni, Y., Hu, G., Wang, X., Tang, Y., 2012. Inter-laboratory calibration of natural gas round robins for δ^2H and $\delta^{13}C$ using off-line and on-line techniques. Chem. Geol. 310–311, 49–55.
Ditchfield, R., Hehre, W.J., Pople, J.A., 1971. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. J. Chem. Phys. 54, 724.
Dong, G., Xie, H., Formolo, M., Lawson, M., Sessions, A., Eiler, J., 2021. Clumped isotope effects of thermogenic methane formation: insights from pyrolysis of hydrocarbons. Geochim. Cosmochim. Acta 303, 159–183.
Dunning, T.H., 1989. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J. Chem. Phys. 90, 1007–1023.
Eldridge, D.L., Korol, R., Lloyd, M.K., Turner, A.C., Webb, M.A., Miller, T.F., Stolper, D. A., 2019. Comparison of Experimental vs Theoretical Abundances of $^{13}CH_3D$ and $^{12}CH_2D_2$ for Isotopically Equilibrated Systems from 1 to 500 °C. ACS Earth Space Chem. 3, 2747–2764.
Eldridge, D.L., Turner, A.C., Bill, M., Conrad, M.E., Stolper, D.A., 2023. Experimental determinations of carbon and hydrogen isotope fractionations and methane clumped isotope compositions associated with ethane pyrolysis from 550 to 600 °C. Geochim. Cosmochim. Acta 355, 235–265.
Feynman, R.P., Hibbs, A.R., 1965. Quantum mechanics and path integrals. McGraw-Hill.
Frenkel, D., Smit, B., 2002. Understanding Molecular Simulation: From Algorithms to Applications. Academic press.
Gao, C., Zhang, Y., Liu, Q., Yang, Y., Liu, Y., 2023. Path-integral molecular dynamics predictions of equilibrium H and O isotope fractionations between brucite and water. Geochim. Cosmochim. Acta 346, 207–222.
Gilbert, A., 2021. The Organic Isotopologue Frontier. Annu. Rev. Earth Planet. Sci. 49, 435–464.
Gilbert, A., Yamada, K., Suda, K., Ueno, Y., Yoshida, N., 2016. Measurement of position-specific ^{13}C isotopic composition of propane at the nanomole level. Geochim. Cosmochim. Acta 177, 205–216.
Gonzalez, Y., Nelson, D.D., Shorter, J.H., McManus, J.B., Dyroff, C., Formolo, M., Wang, D.T., Western, C.M., Ono, S., 2019. Precise measurements of $^{12}CH_2D_2$ by tunable infrared laser direct absorption spectroscopy. Anal. Chem. 91, 14967–14974.
Gropp, J., Iron, M.A., Halevy, I., 2021. Theoretical estimates of equilibrium carbon and hydrogen isotope effects in microbial methane production and anaerobic oxidation of methane. Geochim. Cosmochim. Acta 295, 237–264.
Handy, N.C., Yamaguchi, Y., Schaefer, H.F., 1986. The diagonal correction to the Born–Oppenheimer approximation: Its effect on the singlet–triplet splitting of CH_2 and other molecular effects. J. Chem. Phys. 84, 4481–4484.
Höhnerbach, M., Bientinesi, P., 2019. Accelerating AIREBO: Navigating the Journey from Legacy to High-Performance Code. J. Comput. Chem. 40, 1471–1482.
Horita, J., 2001. Carbon isotope exchange in the system CO_2-CH_4 at elevated temperatures. Geochim. Cosmochim. Acta 65, 1907–1919.
Hunt, J.M., 1996. Petroleum geochemistry and geology, 2nd ed. Freeman, New York, W. H.
Iron, M.A., Gropp, J., 2019. Cost-effective density functional theory (DFT) calculations of equilibrium isotopic fractionation in large organic molecules. Phys. Chem. Chem. Phys. 21, 17555.

Julien, M., Goldman, M.J., Liu, C., Horita, J., Boreham, C.J., Yamada, K., Green, W.H., Yoshida, N., Gilbert, A., 2020. Intramolecular 13C isotope distributions of butane from natural gases. Chem. Geol. 541, 119571.
Kueter, N., Schmidt, M.W., Lilley, M.D., Bernasconi, S.M., 2019. Experimental determination of equilibrium CH_4–CO_2-CO carbon isotope fractionation factors (300–1200 °C). Earth Planet. Sci. Lett. 506, 64–75.
Lee, C., Yang, W., Parr, R.G., 1988. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B 37, 785.
Liu, C., McGovern, G.P., Liu, P., Zhao, H., Horita, J., 2018. Position-specific carbon and hydrogen isotopic compositions of propane from natural gases with quantitative NMR. Chem. Geol. 491, 14–26.
Liu, Q., Liu, Y., 2016. Clumped-isotope signatures at equilibrium of CH_4, NH_3, H_2O, H_2S and SO_2. Geochim. Cosmochim. Acta 175, 252–270.
Liu, Q., Tossell, J.A., Liu, Y., 2010. On the proper use of the Bigeleisen-Mayer equation and corrections to it in the calculation of isotopic fractionation equilibrium constants. Geochim. Cosmochim. Acta 74, 6965–6983.
Liu, Q., Yin, X., Zhang, Y., Julien, M., Zhang, N., Gilbert, A., Yoshida, N., Liu, Y., 2021. Theoretical calculation of position-specific carbon and hydrogen isotope equilibriums in butane isomers. Chem. Geol. 561, 120031.
Lloyd, M.K., Eldridge, D.L., Stolper, D.A., 2021. Clumped $^{13}CH_2D$ and $^{12}CHD_2$ compositions of methyl groups from wood and synthetic monomers: Methods, experimental and theoretical calibrations, and initial results. Geochim. Cosmochim. Acta 297, 233–275.
Lloyd, M.K., Stein, R.A., Ibarra, D.E., Barclay, R.S., Wing, S.L., Stahle, D.W., Dawson, T. E., Stolper, D.A., 2023. Isotopic clumping in wood as a proxy for photorespiration in trees. Proc. Natl. Acad. Sci. 120, e2306736120.
Mangenot, X., Xie, H., Crémière, A., Giunta, T., Lilley, M., Sissmann, O., Orphan, V., Schimmelmann, A., Gaucher, E.C., Girard, J.-P., Eiler, J., 2023. ^2H-^2H clumping in molecular hydrogen method and preliminary results. Chem. Geol. 621, 121278.
Nandi, A., Qu, C., Houston, P.L., Conte, R., Bowman, J.M., 2021. Δ-machine learning for potential energy surfaces: A PIP approach to bring a DFT-based PES to CCSD(T) level of theory. J. Chem. Phys. 154, 051102.
Neese, F., 2018. Software update: the ORCA program system, version 4.0. Wiley Interdiscip. Rev. Comput. Mol. Sci. 8, e1327.
Neese, F., Wennmohs, F., Becker, U., Riplinger, C., 2020. The ORCA quantum chemistry program package. J. Chem. Phys. 152, 224108.
Neese, F., Wiley, J., 2012. The ORCA program system. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2, 73–78.
Neubauer, C., Kantnerová, K., Lamothe, A., Savarino, J., Hilkert, A., Juchelka, D., Hinrichs, K.-U., Elvert, M., Heuer, V., Elsner, M., Bakkour, R., Julien, M., Öztoprak, M., Schouten, S., Hattori, S., Dittmar, T., 2023. Discovering Nature's Fingerprints: Isotope Ratio Analysis on Bioanalytical Mass Spectrometers. J. Am. Soc. Mass Spectrom. 34, 525–537.
O'Neil J. R. (1986) Chapter 1. Theoretical and Experimental Aspects of Isotopic Fractionation eds. J. W. Valley, H. P. Taylor, and J. R. O'Neil. **16**.
Ono, S., Wang, D.T., Gruen, D.S., Sherwood, L.B., Zahniser, M.S., McManus, B.J., Nelson, D.D., 2014. Measurement of a Doubly Substituted Methane Isotopologue, $^{13}CH_3D$, by Tunable Infrared Laser Direct Absorption Spectroscopy. Anal. Chem. 86, 6487–6494.
Piasecki, A., Sessions, A., Lawson, M., Ferreira, A.A., Neto, E.V.S., Eiler, J.M., 2016a. Analysis of the site-specific carbon isotope composition of propane by gas source isotope ratio mass spectrometer. Geochim. Cosmochim. Acta 188, 58–72.
Piasecki, A., Sessions, A., Peterson, B., Eiler, J., 2016b. Prediction of equilibrium distributions of isotopologues for methane, ethane and propane using density functional theory. Geochim. Cosmochim. Acta 190, 1–12.
Pinilla, C., Blanchard, M., Balan, E., Ferlat, G., Vuilleumier, R., Mauri, F., 2014. Equilibrium fractionation of H and O isotopes in water from path integral molecular dynamics. Geochim. Cosmochim. Acta 135, 203–216.
Polyakov, V.B., Horita, J., 2021. Equilibrium carbon isotope fractionation factors of hydrocarbons: Semi-empirical force-field method. Chem. Geol. 559, 119948.
Popa, M.E., Paul, D., Janssen, C., Röckmann, T., 2019. H_2 clumped isotope measurements at natural isotopic abundances. Rapid Commun. Mass Spectrom. 33, 239–251.
Richet, P., Bottinga, Y., Javoy, M., 1977. A Review of Hydrogen, Carbon, Nitrogen, Oxygen, Sulphur, and Chlorine Stable Isotope Fractionation Among Gaseous Molecules. Annu. Rev. Earth Planet. Sci. 5, 65–110.
Rustad, J.R., 2009. Ab initio calculation of the carbon isotope signatures of amino acids. Org. Geochem. 40, 720–723.
Schweizer, K.S., Stratt, R.M., Chandler, D., Wolynes, P.G., 1981. Convenient and accurate discretized path integral methods for equilibrium quantum mechanical calculations. J. Chem. Phys. 75, 1347–1364.
Seewald, J.S., 2003. Organic–inorganic interactions in petroleum-producing sedimentary basins. Nature 426, 327–333.
Stern, M.J., Spindel, W., Monse, E.U., 1968. Temperature dependences of isotope effects. J. Chem, Phys, p. 48.
Stolper D. A., Lawson M., Formolo M. J., Davis C. L., Douglas P. M. J. and Eiler J. M. (2018) The utility of methane clumped isotopes to constrain the origins of methane in natural gas accumulations. In From Source to Seep: Geochemical Applications in Hydrocarbon Systems (eds. M. Lawson, M.J. Formolo, and J.M. Eiler). Geological Society of London. p. 0.
Stolper, D.A., Sessions, A.L., Ferreira, A.A., Santos, N.E., V., Schimmelmann A., Shusta S. S., Valentine D. L. and Eiler J. M., 2014. Combined ^{13}C-D and D-D clumping in methane: Methods and preliminary results. Geochim. Cosmochim. Acta 126, 169–191.
Taguchi, K., Yamamoto, T., Nakagawa, M., Gilbert, A., Ueno, Y., 2020. A fluorination method for measuring the ^{13}C-^{13}C isotopologue of C_2 molecules. Rapid Commun. Mass Spectrom. 34, e8761.
Ten-no, S., Noga, J., 2012. Explicitly correlated electronic structure theory from R12/F12 ansätze. Wires Comput. Mol. Sci. 2, 114–125.
Thiagarajan, N., Xie, H., Ponton, C., Kitchen, N., Peterson, B., Lawson, M., Formolo, M., Xiao, Y., Eiler, J., 2020. Isotopic evidence for quasi-equilibrium chemistry in thermally mature natural gases. Proc. Natl. Acad. Sci. 117, 3989–3995.
Tuckerman, M.E., Berne, B.J., Martyna, G.J., Klein, M.L., 1993. Efficient molecular dynamics and hybrid Monte Carlo algorithms for path integrals. J. Chem. Phys. 99, 2796–2808.
Turner, A.C., Korol, R., Eldridge, D.L., Bill, M., Conrad, M.E., Miller, T.F., Stolper, D.A., 2021. Experimental and theoretical determinations of hydrogen isotopic equilibrium in the system CH_4-H_2-H_2O from 3 to 200 °C. Geochim. Cosmochim. Acta 314, 223–269.
Turner, A.C., Korol, R., Bill, M., Stolper, D.A., 2025. Stable isotope equilibria in the dihydrogen-water-methane-ethane-propane system. Part 2: Experimental determination of hydrogen isotopic equilibrium for ethane-H_2 from 30–200 °C and propane-H_2 from 75–200 °C. Geochim. Cosmochim. Acta.
Urey, H.C., 1947. The Thermodynamic Properties of Isotopic Substances. J. Chem. Soc., 562–581.
Wang, Y., Sessions, A.L., Nielsen, R.J., Goddard, W.A., 2009a. Equilibrium $^2H/^1H$ fractionations in organic molecules: I. Experimental calibration of ab initio calculations. Geochim. Cosmochim. Acta 73, 7060–7075.
Wang, Z., Schauble, E.A., Eiler, J.M., 2004. Equilibrium thermodynamics of multiply substituted isotopologues of molecular gases. Geochim. Cosmochim. Acta 68, 4779–4797.
Wassenaar, L.I., Coplen, T.B., Aggarwal, P.K., 2014. Approaches for achieving long-term accuracy and precision of $\delta^{18}O$ and δ^2H for waters analyzed using laser absorption spectrometers. Environ. Sci. Technol. 48, 1123–1131.
Watts, H.D., Kubicki, J.D., Pedentchouk, N., Freeman, K.H., 2024. Position-Specific $^2H/H$ Equilibrium Isotopic Fractionation Factors in Alkane, Alkene, and Aromatic Molecules: A Density Functional Theory Approach. ACS Earth Space Chem. 8, 21–35.
Webb, M.A., Miller, T.F., 2014. Position-Specific and Clumped Stable Isotope Studies: Comparison of the Urey and Path-Integral Approaches for Carbon Dioxide, Nitrous Oxide, Methane, and Propane. J. Phys. Chem. A 118, 467–474.
Webb, M.A., Wang, Y., Braams, B.J., Bowman, J.M., Miller, T.F., 2017. Equilibrium clumped-isotope effects in doubly substituted isotopologues of ethane. Geochim. Cosmochim. Acta 197, 14–26.
Weiss, G.M., Sessions, A.L., Julien, M., Csernica, T., Yamada, K., Gilbert, A., Freeman, K. H., Eiler, J.M., 2023. Analysis of intramolecular carbon isotope distributions in alanine by electrospray ionization Orbitrap mass spectrometry. Int. J. Mass Spectrom. 493, 117128.
Werner, H.J., Knowles, P.J., Knizia, G., Manby, F.R., Schütz, M., 2012. Molpro: A general-purpose quantum chemistry program package. Wiley Interdiscip. Rev. Comput. Mol, Sci, p. 2.
Wilkes, E.B., Sessions, A.L., Zeichner, S.S., Dallas, B., Schubert, B., Jahren, A.H., Eiler, J. M., 2022. Position-specific carbon isotope analysis of serine by gas chromatography/Orbitrap mass spectrometry, and an application to plant metabolism. Rapid Commun. Mass Spectrom. 36, e9347.
Xie, H., Dong, G., Formolo, M., Lawson, M., Liu, J., Cong, F., Mangenot, X., Shuai, Y., Ponton, C., Eiler, J., 2021. The evolution of intra- and inter-molecular isotope equilibria in natural gases with thermal maturation. Geochim. Cosmochim. Acta 307, 22–41.
Xie, H., Formolo, M.J., Sessions, A.L., Eiler, J.M., 2024. Theoretical and experimental constraints on hydrogen isotope equilibrium in C_1-C_5 alkanes. Geochim. Cosmochim. Acta 386, 63–73.
Xie, H., Ponton, C., Formolo, M.J., Lawson, M., Peterson, B.K., Lloyd, M.K., Sessions, A. L., Eiler, J.M., 2018. Position-specific hydrogen isotope equilibrium in propane. Geochim. Cosmochim. Acta 238, 193–207.
Yin, X., Zhang, Y., Liu, Q., Gilbert, A., Liu, F., Gao, C., Zhang, S., Ridley, M.K., Liu, Y., 2024. Position-specific and clumped isotope equilibria in propane: Ab initio calculations beyond the harmonic and Born-Oppenheimer approximations. Chem, Geol.
Young, E.D., Rumble, D., Freedman, P., Mills, M., 2016. A large-radius high-mass-resolution multiple-collector isotope ratio mass spectrometer for analysis of rare isotopologues of O_2, N_2, CH_4 and other gases. Int. J. Mass Spectrom. 401, 1–10.
Zeichner, S.S., Aponte, J.C., Bhattacharjee, S., Dong, G., Hofmann, A.E., Dworkin, J.P., Glavin, D.P., Elsila, J.E., Graham, H.V., Naraoka, H., Takano, Y., Tachibana, S., Karp, A.T., Grice, K., Holman, A.I., Freeman, K.H., Yurimoto, H., Nakamura, T., Noguchi, T., Okazaki, R., Yabuta, H., Sakamoto, K., Yada, T., Nishimura, M., Nakato, A., Miyazaki, A., Yogata, K., Abe, M., Okada, T., Usui, T., Yoshikawa, M., Saiki, T., Satoshi, T., Terui, F., Nakazawa, S., Watanabe, S., Tsuda, Y., Hamase, K., Fukushima, K., Aoki, D., Hashiguchi, M., Mita, H., Chikaraishi, Y., Ohkouchi, N., Ogawa, N.O., Sakai, S., Parker, E.T., McLain, H.L., Orthous-Daunay, F.-R., Vuitton, V., Wolters, C., Schmitt-Kopplin, P., Hertkorn, N., Thissen, R., Ruf, A., Isa, J., Oba, Y., Koga, T., Yoshimura, T., Araoka, D., Sugahara, H., Furusho, A., Furukawa, Y., Aoki, J., Kano, K., Nomura, S.M., Sasaki, K., Sato, H., Yoshikawa, T., Satoru, T., Morita, M., Onose, M., Kabashima, F., Fujishima, K., Yamazaki, T., Kimura, Y., Eiler, J.M., 2023. Polycyclic aromatic hydrocarbons in samples of ryugu formed in the interstellar medium. Science 382, 1411–1416.
Zhang, Y., Liu, Y., 2024. On accurate calculations of equilibrium H/D fractionations at super-cold conditions (≤200 K) I: Full Partition Function Ratio (FPFR) vs. Path Integral Monte Carlo (PIMC). Chem. Geol. 666, 122303.
Zhang, Y., Liu, Y., 2018. The theory of equilibrium isotope fractionations for gaseous molecules under super-cold conditions. Geochim. Cosmochim. Acta 238, 123–149.