Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information for: Exploration of Biochemical Reactivity with a QM/MM Growing String Method

Neil R. McFarlane[a], Jeremy N. Harvey[a]

[a] Department of Chemistry
KU Leuven
B-3001 Leuven Celestijnenlaan 200f- box 2404, Belgium
E-Mail: neilrory.mcfarlane@kuleuven.be
E-Mail: jeremy.harvey@kuleuven.be

Additional Figures and Tables

Table S1: *Chronological list of QM/MM studies of reactivity in BsCM. For each study, the starting PDB structure(s), the QM method and basis set, the MM forcefield, the barrier height (in kcal/mol) with standard deviation or range (where available) are provided.*

PDB ID	QM method/Basis set	MM forcefield	Barrier height (kcal/mol)	Reference
2CHT	AM1	CHARMM	17.8	[1]
2CHT, 1COM	HF/6-31G(d)	Amber	24.8	[2]
2CHT, 1COM	MP2/6-31G(d)	Amber	6.4	[2]
N/R	B3LYP/6-31G(d)	Amber	1.4	[3]
1COM	AM1	CHARMM	35.7	[4]
2CHT	HF/4-31G//cc-pVDZ	CHARMM	19.7	[5]
2CHT	B3LYP/6-311+G(2d,p)//6-31G(d)	CHARMM	2.7 – 11.9	[6]
2CHT	AM1	CHARMM	24.5 – 31.6	[6]
2CHT	MP2/6-31+G(d)//6-31G(d)	CHARMM	7.4 – 11.0	[7]
2CHT	B3LYP/6-311+G(2d,p)//6-31G(d)	CHARMM	12.7 – 16.1	[7]
2CHT	B3LYP/6-31G(d)	CHARMM	12.0 ± 1.7	[8]
2CHT	B3LYP/aug-cc-pVTZ//cc-pVTZ	CHARMM	10.2 ± 1.8	[9]
2CHT	LCCSD(T0)/aug-cc-pVTZ//cc-pVTZ	CHARMM	13.1 ± 1.1	[9]
2CHT	B3LYP/6-31G(d)	CHARMM	11.3 ± 1.8	[10]
2CHT	B3LYP/6-31G(d)	Amber	12.1 ± 0.2	[11]
2CHT	B3LYP/def2-svp	Amber	15.1	[12]
N/A	Exp.		12.7 ± 0.4	[13]

We have searched the literature for the most well-known QM/MM studies into this system (collated chronologically in ***Table S1***). Overall, five QM methods have emerged, and some minor variations in MM forcefield choice. The barrier has been most accurately calculated using LCCSD(T0), where the energy barrier is predicted to be 13.1 ± 1.1 kcal/mol [9] – in excellent agreement with experiment.[13] However, the high computational expense means that this method has only been used for single-point calculations on a B3LYP optimised pathway. Taking the averages of the remaining four methods (using the upper end of the ranges from ref. [6,7]), we find barriers of 8.7 kcal/mol (MP2),[2,7] 28.4 kcal/mol (AM1),[1,4,6] 22.3 kcal/mol (HF),[2,5] and 11.3 kcal/mol (B3LYP).[3,6–12] When comparing to the known experimental value for this reaction,[13] the most common choice for QM/MM reaction pathway generation is B3LYP. Removing the outlier predicted in reference [3], who calculate the barrier to be unreasonably small likely due to insufficient system preparation and/or sampling, we find the average barrier to be 12.7 kcal/mol – very similar to the experimental value.[13] Across the studies reported herein, the size of the QM region, methods for system preparation, extent of sampling, among others, is not consistent, so these factors may also have some effect on the energy barriers.

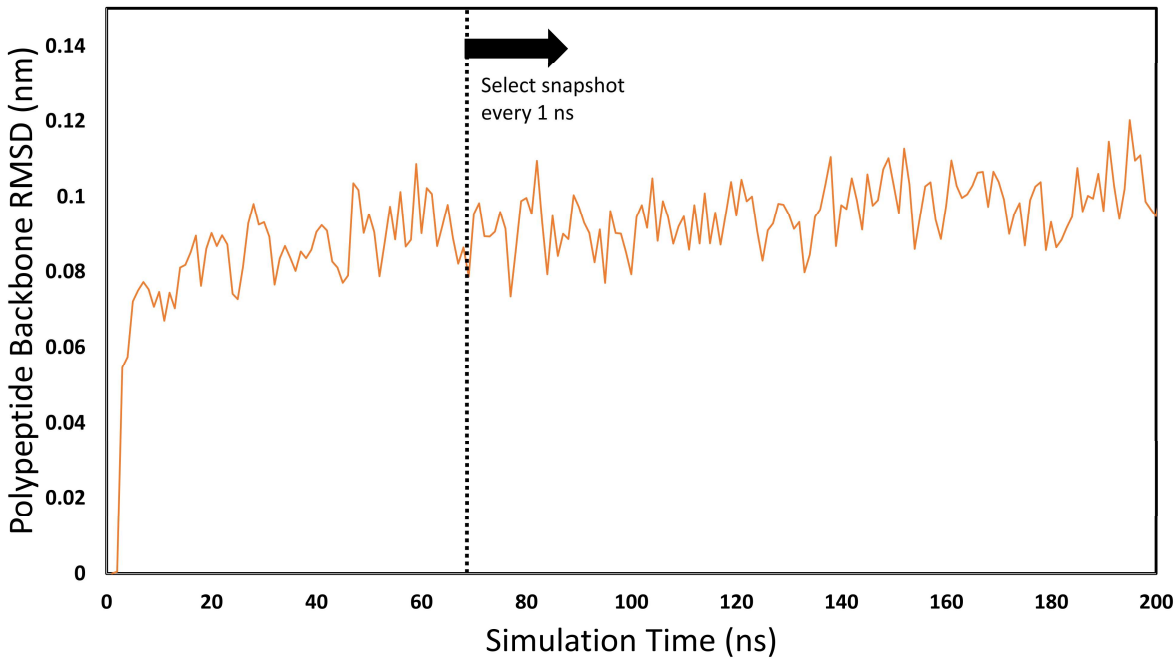

Figure S1: *RMSD of the polypeptide backbone α-carbons with respect to the starting configuration over 200 ns of simulation time, calculating the RMSD at every ns. From 68 ns, MD snapshots were taken systematically every 1 ns, as illustrated.*

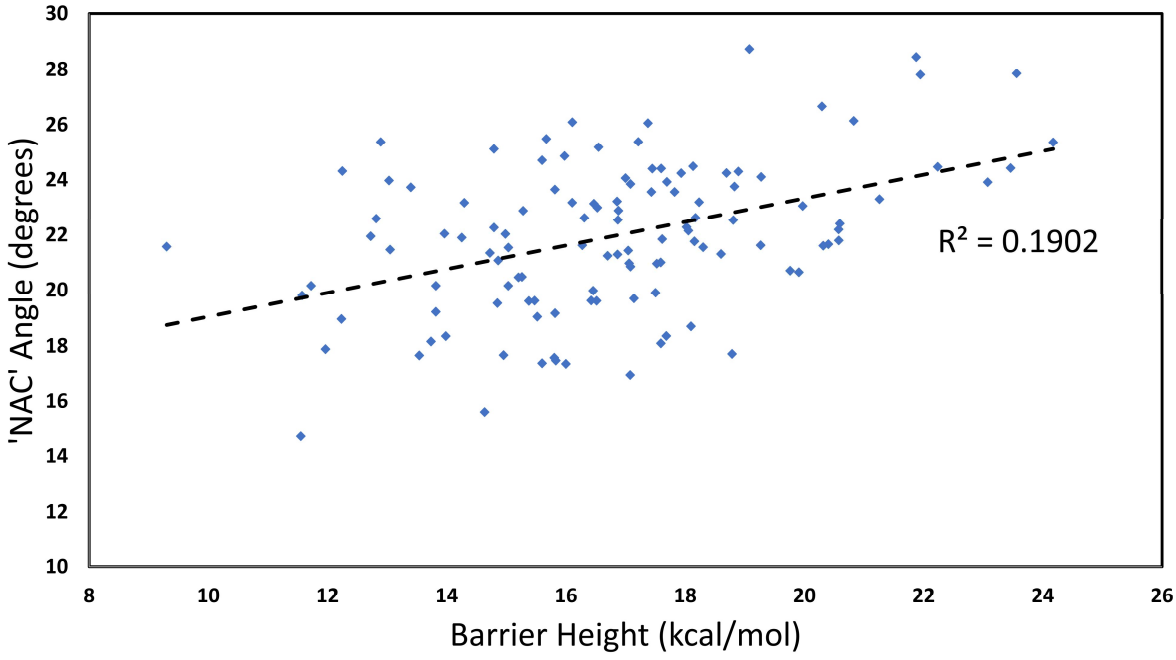

Figure S2: *The dependency of the 'NAC' angle on the barrier height of the 117 pathways calculated herein. For a definition of the 'NAC' angle, please refer to ref.* [14].

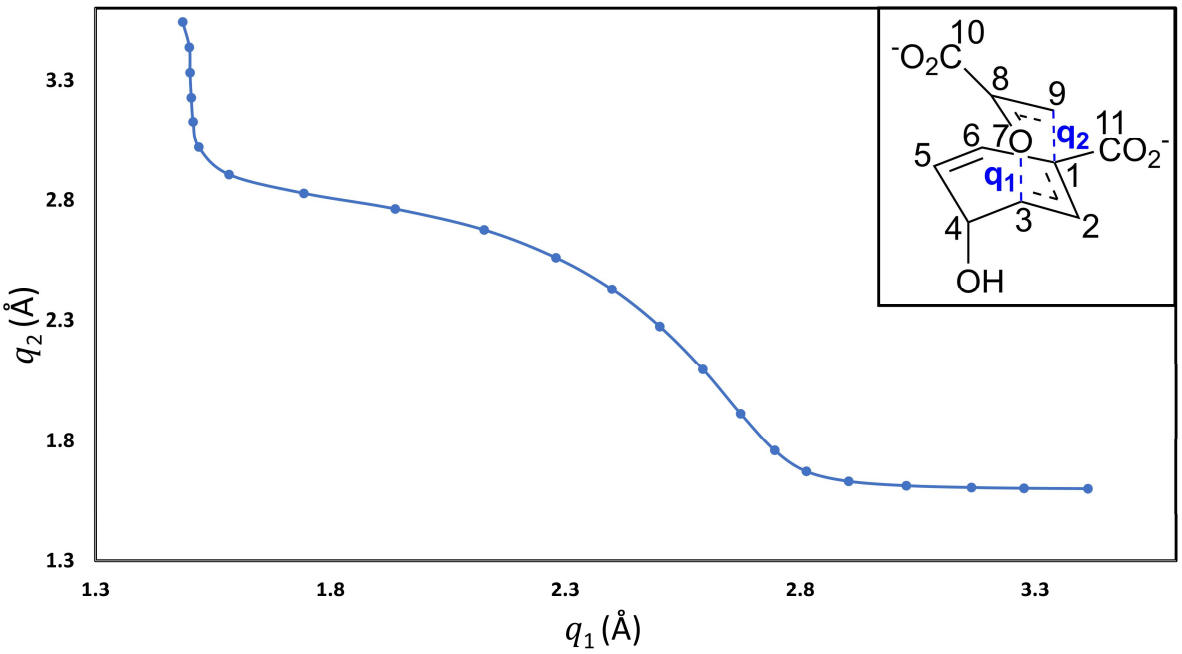

Figure S3: *Variation of internal coordinates q_1 and q_2 averaged along 117 pathways generated herein. q_1 is defined as $d(C_3 - O_7)$, and q_2 is defined as $d(C_1 - C_9)$. An illustration of the relevant internal coordinates is given in the figure inset.*

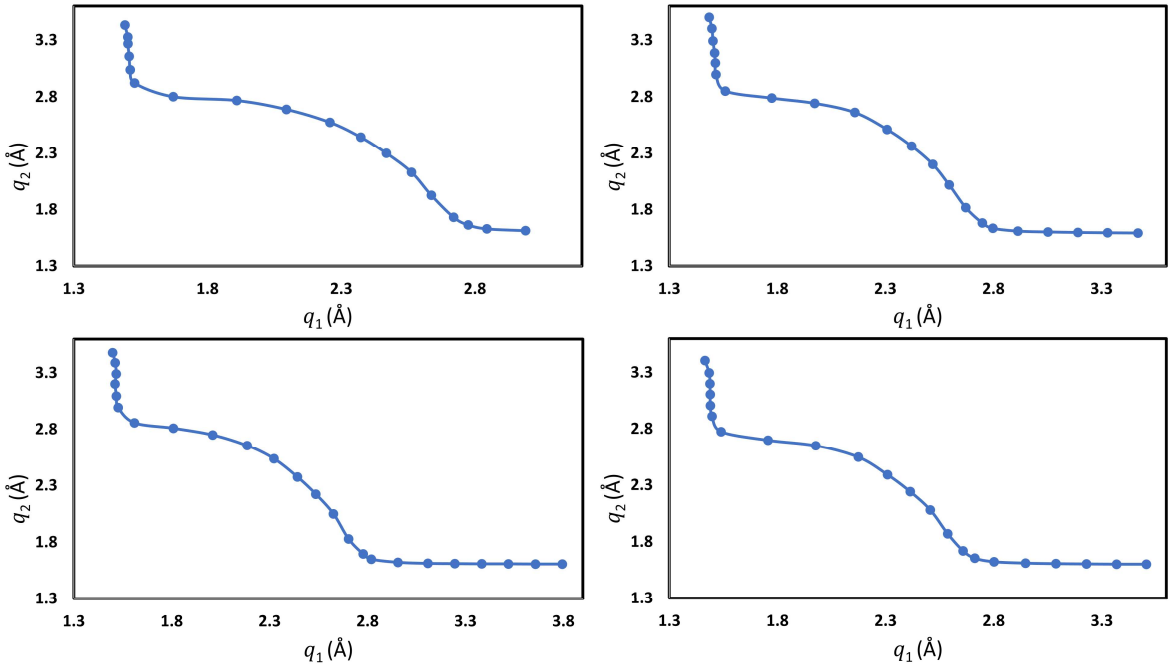

Figure S4: *Variation of internal coordinates q_1 and q_2 for four of the 117 pathways generated herein. q_1 is defined as $d(C_3 - O_7)$, and q_2 is defined as $d(C_1 - C_9)$. An illustration of the relevant internal coordinates is given in the figure inset of* ***Figure S3****.*

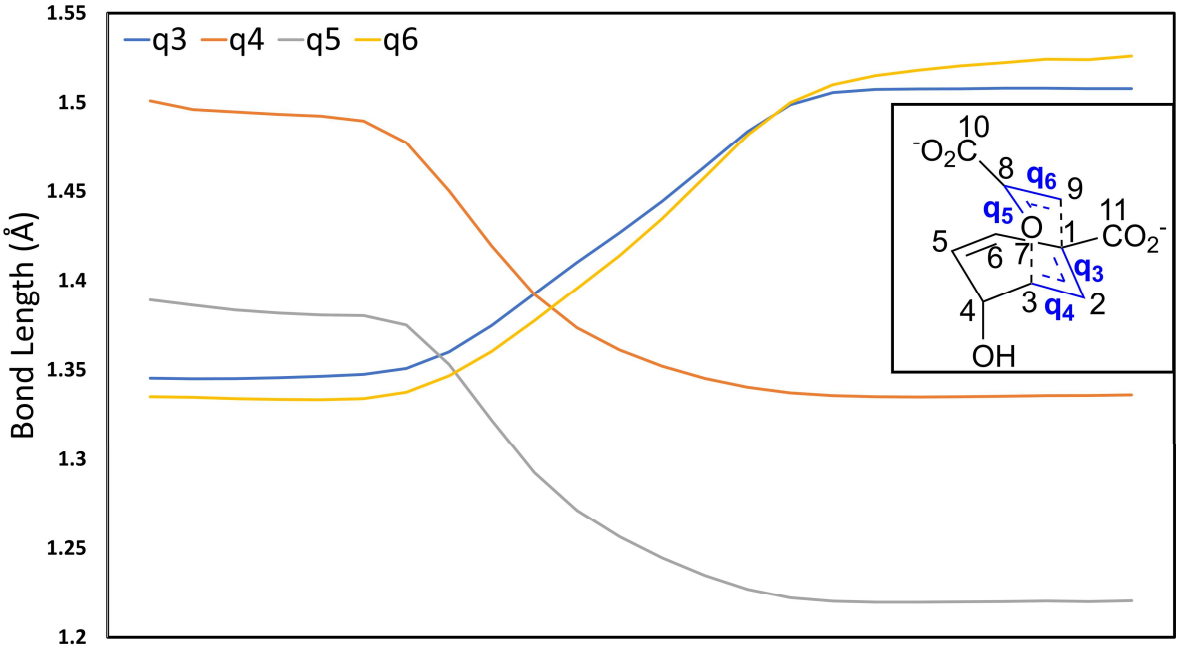

Figure S5: *Variation of internal coordinates q_3, q_4, q_5, and q_6 averaged along 117 pathways generated herein. q_3 is defined as $d(C_1 - C_{11})$, q_4 is defined as $d(C_2 - C_3)$, q_5 is defined as $d(O_7 - C_8)$, and q_6 is defined as $d(C_8 - C_9)$. An illustration of the relevant internal coordinates is given in the figure inset.*

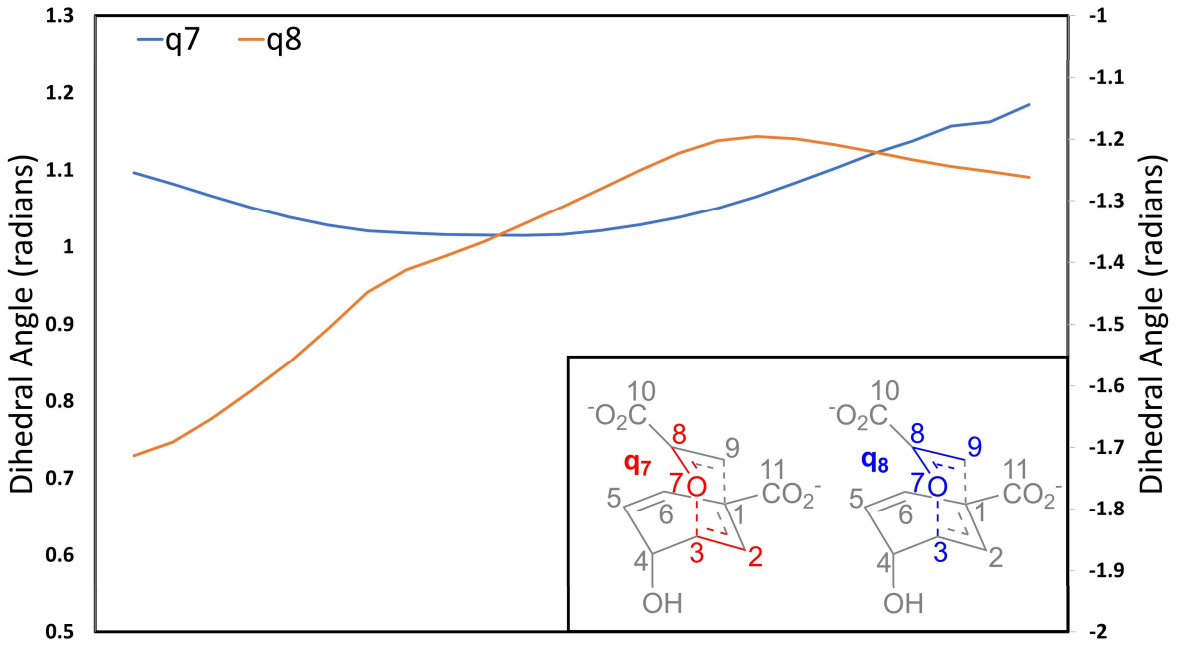

Figure S6: *Variation of internal coordinates q_7 and q_8 averaged along 117 pathways generated herein. q_7 is defined as $\varphi(C_2 - C_3 - O_7 - C_8)$, and q_8 is defined as $\varphi(C_3 - O_7 - C_8 - C_9)$. An illustration of the relevant internal coordinates is given in the figure inset.*

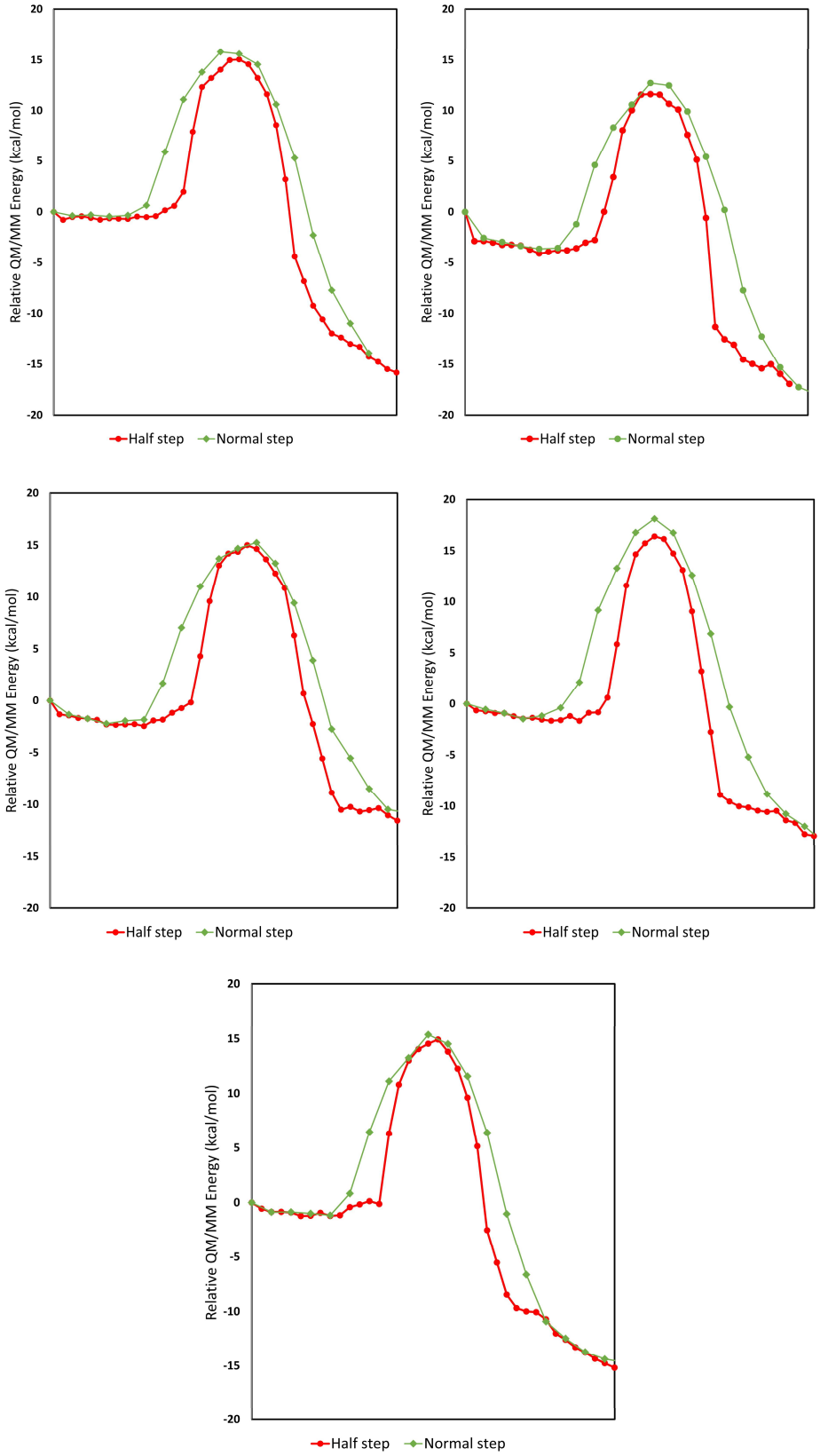

Figure S7: *Comparison of using half-step or normal-step in the growth-phase of the QM/MM SE-GSM algorithm for five of the 117 pathways generated herein. A normal step is defined as steps of 0.1 Å and 5° for bonds and valence and/or dihedral angles, respectively. The half-step is half of these values.*

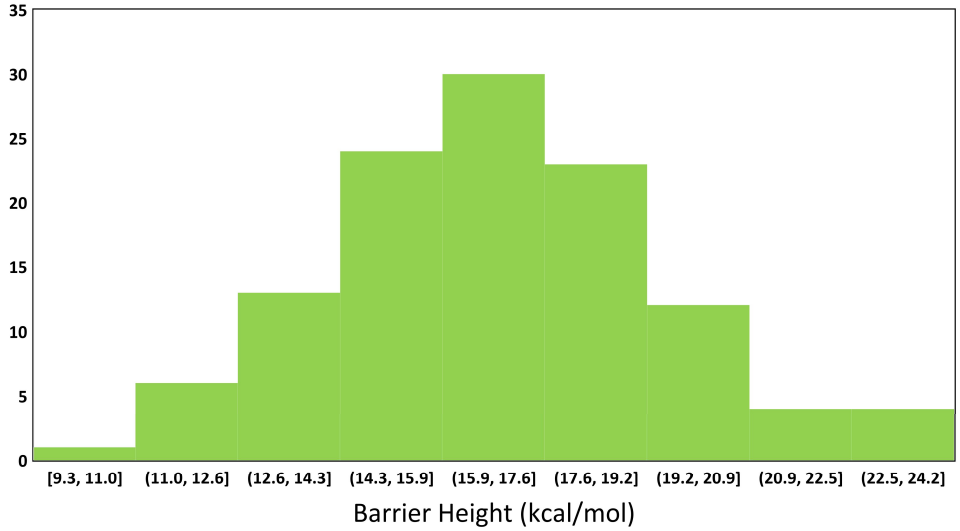

Figure S8: *Histogram of the distribution of barrier heights for the 117 reaction pathways obtained herein.*

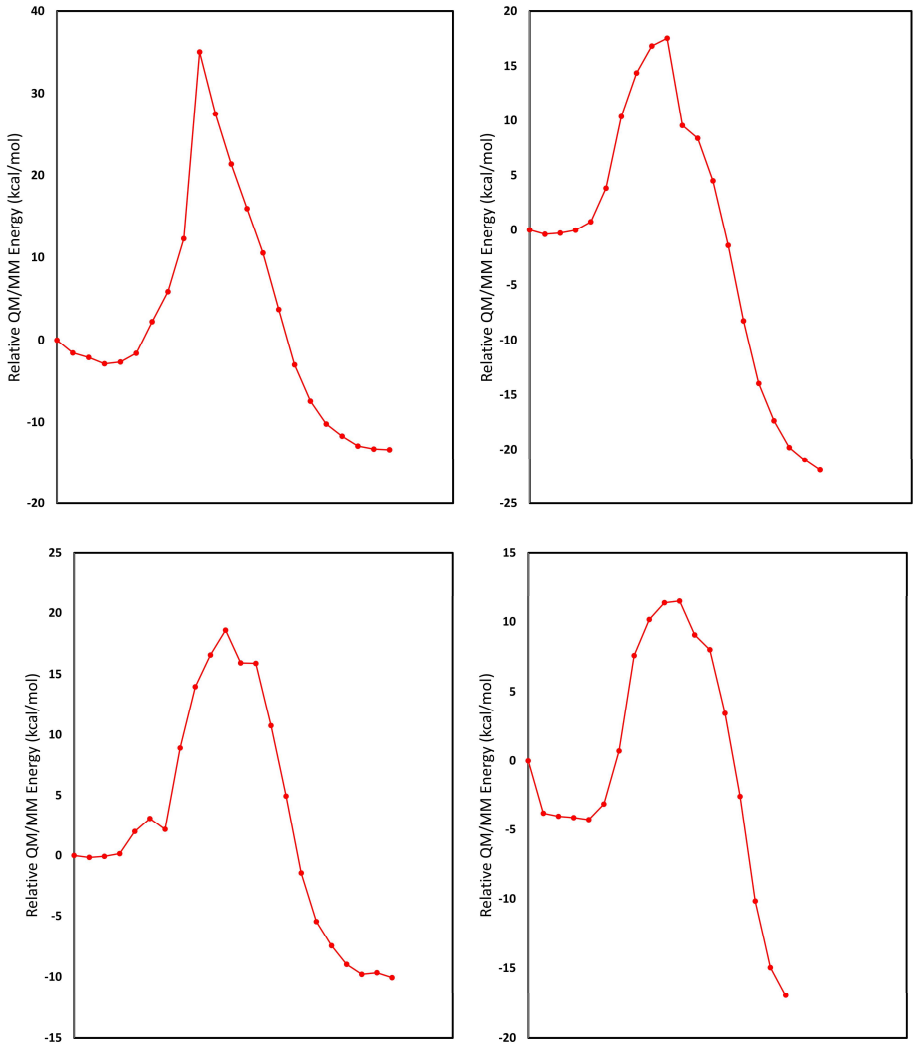

Figure S9: *Four examples of QM/MM energy profiles which were deemed as discontinuous and therefore discarded from further analysis.*

Additional Parameters

Chorismate Forcefield Parameters

cho.itp

```
; Created by cgenff_charmm2gmx.py

[ moleculetype ]
; Name                                 nrexcl
CHO                                     3

[ atoms ]
;             nr                       type       resnr residue  atom   cgnr             charge                 mass  typeB    chargeB
            massB
; residue    1 CHO rtp CHO q       qsum
    1    CG2DC1     1    CHO    C1     1    -0.200    12.011  ;
    2    CG2DC2     1    CHO    C2     2    -0.067    12.011  ;
    3    CG2DC2     1    CHO    C3     3    -0.386    12.011  ;
    4     CG311     1    CHO    C4     4     0.295    12.011  ;
    5     OG311     1    CHO    O5     5    -0.638    15.999  ;
    6     CG311     1    CHO    C6     6     0.407    12.011  ;
    7    CG2DC1     1    CHO    C7     7    -0.319    12.011  ;
    8     OG301     1    CHO    O8     8    -0.348    15.999  ;
    9    CG2D1O     1    CHO    C9     9    -0.205    12.011  ;
   10    CG2DC3     1    CHO   C10    10    -0.611    12.011  ;
   11     CG2O3     1    CHO   C11    11     0.666    12.011  ;
   12     OG2D2     1    CHO   O12    12    -0.740    15.999  ;
   13     OG2D2     1    CHO   O13    13    -0.740    15.999  ;
   14     CG2O3     1    CHO   C14    14     0.897    12.011  ;
   15     OG2D2     1    CHO   O15    15    -0.742    15.999  ;
   16     OG2D2     1    CHO   O16    16    -0.742    15.999  ;
   17      HGP1     1    CHO   H17    17     0.414     1.008  ;
   18      HGA5     1    CHO   H18    18     0.230     1.008  ;
   19      HGA5     1    CHO   H19    19     0.230     1.008  ;
   20      HGA4     1    CHO   H20    20     0.145     1.008  ;
   21      HGA4     1    CHO   H21    21     0.124     1.008  ;
   22      HGA4     1    CHO   H22    22     0.150     1.008  ;
   23      HGA1     1    CHO   H23    23     0.090     1.008  ;
   24      HGA1     1    CHO   H24    24     0.090     1.008  ;

[ bonds ]
;             ai             aj funct                          c0                          c1                          c2
                            c3
   1    2     1
   1   11     1
   1    7     1
   2    3     1
   2   20     1
   3    4     1
   3   22     1
   4    5     1
   4    6     1
   4   23     1
   5   17     1
   6    8     1
   6    7     1
   6   24     1
   7   21     1
   8    9     1
   9   10     1
   9   14     1
  10   18     1
  10   19     1
  11   12     1
  11   13     1
  14   15     1
  14   16     1

[ pairs ]
```

```
;          ai        aj funct          c0          c1          c2          c3
   1   24    1
   1    4    1
   1   22    1
   1    8    1
   2    5    1
   2    6    1
   2   12    1
   2   13    1
   2   21    1
   2   23    1
   3   17    1
   3   11    1
   3   24    1
   3    7    1
   3    8    1
   4    9    1
   4   20    1
   4   21    1
   5   24    1
   5   22    1
   5    7    1
   5    8    1
   6   10    1
   6   11    1
   6   14    1
   6   17    1
   6   22    1
   7    9    1
   7   12    1
   7   13    1
   7   20    1
   7   23    1
   8   15    1
   8   16    1
   8   18    1
   8   19    1
   8   21    1
   8   23    1
   9   24    1
  10   15    1
  10   16    1
  11   20    1
  11   21    1
  14   18    1
  14   19    1
  17   23    1
  20   22    1
  21   24    1
  22   23    1
  23   24    1

[ angles ]
;          ai        aj        ak funct          c0          c1          c2          c3
   2    1   11        5
   2    1    7        5
  11    1    7        5
   1    2    3        5
   1    2   20        5
   3    2   20        5
   2    3    4        5
   2    3   22        5
   4    3   22        5
   3    4    5        5
   3    4    6        5
   3    4   23        5
   5    4    6        5
   5    4   23        5
   6    4   23        5
```

```
 17    5    4           5
  8    6    4           5
  8    6    7           5
  8    6   24           5
  4    6    7           5
  4    6   24           5
  7    6   24           5
  1    7   21           5
  1    7    6           5
 21    7    6           5
  9    8    6           5
 10    9   14           5
 10    9    8           5
 14    9    8           5
  9   10   18           5
  9   10   19           5
 18   10   19           5
  1   11   12           5
  1   11   13           5
 12   11   13           5
  9   14   15           5
  9   14   16           5
 15   14   16           5

[ dihedrals ]
;          ai          aj          ak          al funct                    c0                    c1
                 c2                    c3                    c4                    c5
 11    1    2    3        9
 11    1    2   20        9
  7    1    2    3        9
  7    1    2   20        9
  2    1   11   12        9
  2    1   11   13        9
  7    1   11   12        9
  7    1   11   13        9
  2    1    7   21        9
  2    1    7    6        9
 11    1    7   21        9
 11    1    7    6        9
  1    2    3    4        9
  1    2    3   22        9
 20    2    3    4        9
 20    2    3   22        9
  2    3    4    5        9
  2    3    4    6        9
  2    3    4   23        9
 22    3    4    5        9
 22    3    4    6        9
 22    3    4   23        9
  3    4    5   17        9
  6    4    5   17        9
 23    4    5   17        9
  3    4    6    8        9
  3    4    6    7        9
  3    4    6   24        9
  5    4    6    8        9
  5    4    6    7        9
  5    4    6   24        9
 23    4    6    8        9
 23    4    6    7        9
 23    4    6   24        9
  4    6    8    9        9
  7    6    8    9        9
 24    6    8    9        9
  8    6    7    1        9
  8    6    7   21        9
  4    6    7    1        9
  4    6    7   21        9
 24    6    7    1        9
 24    6    7   21        9
  6    8    9   10        9
```

```
  6    8    9   14          9
 14    9   10   18          9
 14    9   10   19          9
  8    9   10   18          9
  8    9   10   19          9
 10    9   14   15          9
 10    9   14   16          9
  8    9   14   15          9
  8    9   14   16          9

[ dihedrals ]
;           ai            aj          ak          al funct                    c0                    c1
            c2                        c3
  9   10   14    8          2
 11   13   12    1          2
 14   16   15    9          2
```

cho.prm

```
[ bondtypes ]
;    i      j  func       b0        kb
 CG2D1O   CG2DC3    1  0.13400000   368192.00
 CG2D1O    CG2O3    1  0.14925000   239341.54
 CG2DC1    CG311    1  0.15020000   305432.00
 CG2DC2    CG311    1  0.15020000   305432.00


[ angletypes ]
;    i      j      k  func      theta0      ktheta        ub0        kub
 CG2DC3   CG2D1O    CG2O3    5   108.060000   502.080000   0.00000000       0.00
 CG2DC3   CG2D1O    OG301    5   123.500000   543.920000   0.00000000       0.00
  CG2O3   CG2D1O    OG301    5   111.880000   574.128480   0.00000000       0.00
 CG2DC1   CG2DC1    CG311    5   123.500000   401.664000   0.00000000       0.00
 CG2DC2   CG2DC1    CG2O3    5   121.570000   403.002880   0.00000000       0.00
  CG311   CG2DC1    HGA4    5   116.000000   334.720000   0.00000000       0.00
 CG2DC2   CG2DC2    CG311    5   123.500000   401.664000   0.00000000       0.00
  CG311   CG2DC2    HGA4    5   116.000000   334.720000   0.00000000       0.00
 CG2D1O   CG2DC3    HGA5    5   120.500000   376.560000   0.00000000       0.00
 CG2D1O    CG2O3    OG2D2    5   116.000000   334.720000   0.23530000    41840.00
 CG2DC1    CG311    CG311    5   112.200000   267.776000   0.00000000       0.00
 CG2DC1    CG311    OG301    5    99.000000   167.360000   0.00000000       0.00
 CG2DC1    CG311    HGA1    5   111.500000   376.560000   0.00000000       0.00
 CG2DC2    CG311    CG311    5   112.200000   267.776000   0.00000000       0.00
 CG2DC2    CG311    OG311    5   110.100000   633.457600   0.00000000       0.00
 CG2DC2    CG311    HGA1    5   111.500000   376.560000   0.00000000       0.00
 CG2D1O    OG301    CG311    5   109.000000   443.504000   0.00000000       0.00


[ dihedraltypes ]
;    i      j      k      l  func       phi0       kphi  mult
  CG2O3   CG2D1O   CG2DC3     HGA5     9   180.000000    13.439845    2
  OG301   CG2D1O   CG2DC3     HGA5     9   180.000000    12.552000    2
 CG2DC3   CG2D1O    CG2O3    OG2D2     9     0.000000     3.884426    1
 CG2DC3   CG2D1O    CG2O3    OG2D2     9   180.000000     9.211494    2
 CG2DC3   CG2D1O    CG2O3    OG2D2     9   180.000000     0.108784    3
 CG2DC3   CG2D1O    CG2O3    OG2D2     9     0.000000     1.703306    4
  OG301   CG2D1O    CG2O3    OG2D2     9     0.000000    14.343589    2
  OG301   CG2D1O    CG2O3    OG2D2     9     0.000000     6.443360    3
  OG301   CG2D1O    CG2O3    OG2D2     9   180.000000     0.196648    6
 CG2DC3   CG2D1O    OG301    CG311     9   180.000000     3.765600    1
 CG2DC3   CG2D1O    OG301    CG311     9   180.000000    12.970400    2
 CG2DC3   CG2D1O    OG301    CG311     9   180.000000     5.020800    3
  CG2O3   CG2D1O    OG301    CG311     9   180.000000     3.347200    1
  CG2O3   CG2D1O    OG301    CG311     9   180.000000    12.552000    2
  CG2O3   CG2D1O    OG301    CG311     9   180.000000     4.602400    3
 CG2DC2   CG2DC1   CG2DC1    CG311     9   180.000000     2.343040    1
 CG2DC2   CG2DC1   CG2DC1    CG311     9   180.000000    29.288000    2
  CG2O3   CG2DC1   CG2DC1    CG311     9   180.000000     2.343040    1
  CG2O3   CG2DC1   CG2DC1    CG311     9   180.000000    29.288000    2
  CG2O3   CG2DC1   CG2DC1     HGA4     9   180.000000     4.537548    2
```

```
 CG2O3  CG2DC1  CG2DC2  CG2DC2    9  180.000000    3.204944    1
 CG2O3  CG2DC1  CG2DC2  CG2DC2    9    0.000000   14.230202    2
 CG2O3  CG2DC1  CG2DC2  CG2DC2    9    0.000000    4.152620    3
 CG2O3  CG2DC1  CG2DC2   HGA4    9  180.000000    4.248434    2
CG2DC2  CG2DC1   CG2O3   OG2D2    9  180.000000    5.439200    2
CG2DC1  CG2DC1   CG311   CG311    9    0.000000    2.092000    2
CG2DC1  CG2DC1   CG311   CG311    9    0.000000    1.255200    3
CG2DC1  CG2DC1   CG311   OG301    9    0.000000    2.928800    3
CG2DC1  CG2DC1   CG311    HGA1    9    0.000000    0.125520    3
 HGA4   CG2DC1   CG311   CG311    9    0.000000    0.836800    3
 HGA4   CG2DC1   CG311   OG301    9    0.000000    0.836800    3
 HGA4   CG2DC1   CG311    HGA1    9    0.000000    0.836800    3
CG2DC1  CG2DC2  CG2DC2   CG311    9  180.000000    2.343040    1
CG2DC1  CG2DC2  CG2DC2   CG311    9  180.000000   29.288000    2
 CG311  CG2DC2  CG2DC2   HGA4    9  180.000000   21.756800    2
CG2DC2  CG2DC2   CG311   CG311    9    0.000000    2.092000    2
CG2DC2  CG2DC2   CG311   CG311    9    0.000000    1.255200    3
CG2DC2  CG2DC2   CG311   OG311    9  180.000000    7.949600    1
CG2DC2  CG2DC2   CG311   OG311    9  180.000000    1.673600    2
CG2DC2  CG2DC2   CG311   OG311    9  180.000000    2.510400    3
CG2DC2  CG2DC2   CG311    HGA1    9    0.000000    0.125520    3
 HGA4   CG2DC2   CG311   CG311    9    0.000000    0.836800    3
 HGA4   CG2DC2   CG311   OG311    9    0.000000    0.836800    3
 HGA4   CG2DC2   CG311    HGA1    9    0.000000    0.836800    3
CG2DC1   CG311   CG311  CG2DC2    9  180.000000    3.765600    1
CG2DC1   CG311   CG311  CG2DC2    9  180.000000    1.255200    3
CG2DC1   CG311   CG311   OG311    9    0.000000    0.836800    3
CG2DC1   CG311   CG311    HGA1    9    0.000000    0.815880    3
CG2DC2   CG311   CG311   OG301    9    0.000000    0.836800    3
CG2DC2   CG311   CG311    HGA1    9    0.000000    0.815880    3
 OG301   CG311   CG311   OG311    9    0.000000    0.836800    3
CG2DC1   CG311   OG301  CG2D1O    9  180.000000    3.807440    1
CG2DC1   CG311   OG301  CG2D1O    9    0.000000    2.092000    2
 CG311   CG311   OG301  CG2D1O    9    0.000000    1.004160    1
 CG311   CG311   OG301  CG2D1O    9    0.000000    1.213360    2
 CG311   CG311   OG301  CG2D1O    9    0.000000    0.083680    3
 HGA1    CG311   OG301  CG2D1O    9    0.000000    0.271960    3
CG2DC2   CG311   OG311   HGP1    9    0.000000    5.439200    1
CG2DC2   CG311   OG311   HGP1    9    0.000000    2.928800    2
CG2DC2   CG311   OG311   HGP1    9    0.000000    0.585760    3


[ dihedraltypes ]
; 'improper' dihedrals
;    i      j      k      l  func       phi0       kphi
 CG2D1O  CG2DC3   CG2O3   OG301    2    0.000000  251.040000
  CG2O3   OG2D2   OG2D2  CG2D1O    2    0.000000  803.328000
```

Optimisation Convergence Criteria

Growth Phase

Change in energy:	0.000400 (Ha)
Maximum change of X:	0.014000 (bohr)
RMS change of X:	0.008000 (bohr)
Maximum gradient element:	0.010000 (Ha/bohr)
RMS gradient element:	0.004000 (Ha/bohr)

Optimisation Phase & Normal Criteria

Change in energy:	0.000100 (Ha)
Maximum change of X:	0.007000 (bohr)
RMS change of X:	0.004000 (bohr)
Maximum gradient element:	0.005000 (Ha/bohr)
RMS gradient element:	0.002000 (Ha/bohr)

References

1 P. D. Lyne, A. J. Mulholland and W. G. Richards, *J. Am. Chem. Soc.*, 1995, **117**, 11345–11350.
2 M. M. Davidson, I. R. Gould and I. H. Hillier, *J. Chem. Soc. Perkin Trans. 2*, 1996, 525–532.
3 R. J. Hall, S. A. Hindle, N. A. Burton and I. H. Hillier, *J. Comput. Chem.*, 2000, **21**, 1433–1441.
4 S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, J. Bertrán and M. J. Field, *J. Am. Chem. Soc.*, 2001, **123**, 1709–1712.
5 Y. S. Lee, S. E. Worthington, M. Krauss and B. R. Brooks, *J. Phys. Chem. B*, 2002, **106**, 12059–12065.
6 K. E. Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann and A. J. Mulholland, *Mol. Phys.*, 2003, **101**, 2695–2714.
7 K. E. Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann and A. J. Mulholland, *Org. Biomol. Chem.*, 2004, **2**, 968.
8 F. Claeyssens, K. E. Ranaghan, F. R. Manby, J. N. Harvey and A. J. Mulholland, *Chem. Commun.*, 2005, 5068.
9 F. Claeyssens, J. N. Harvey, F. R. Manby, R. A. Mata, A. J. Mulholland, K. E. Ranaghan, M. Schütz, S. Thiel, W. Thiel and H.-J. Werner, *Angew. Chem. Int. Ed.*, 2006, **45**, 6856–6859.
10 F. Claeyssens, K. E. Ranaghan, N. Lawan, S. J. Macrae, F. R. Manby, J. N. Harvey and A. J. Mulholland, *Org. Biomol. Chem.*, 2011, **9**, 1578.
11 X. Pan, R. Van, E. Epifanovsky, J. Liu, J. Pu, K. Nam and Y. Shao, *J. Phys. Chem. B*, 2022, **126**, 4226–4235.
12 S. A. Siddiqui and K. D. Dubey, *Phys. Chem. Chem. Phys.*, 2022, **24**, 1974–1981.
13 P. Kast, M. Asif-Ullah and D. Hilvert, *Tetrahedron Lett.*, 1996, **37**, 2691–2694.
14 S. Hur and T. C. Bruice, *J. Am. Chem. Soc.*, 2003, **125**, 10540–10542.