

# **Mechanistic Applications of Volume Profiles for Chemical and Biochemical Processes**

**by  
Heather Wiebe**

B.Sc., University of the Fraser Valley, 2011

Thesis Submitted in Partial Fulfillment of the  
Requirements for the Degree of  
Doctor of Philosophy

in the  
Department of Chemistry  
Faculty of Science

© Heather Wiebe 2018  
SIMON FRASER UNIVERSITY  
Spring 2018

Copyright in this work rests with the author. Please ensure that any reproduction or re-use is done in accordance with the relevant national copyright legislation.

# Approval

**Name:** Heather Wiebe

**Degree:** Doctor of Philosophy (Chemistry)

**Title:** *Mechanistic Applications of Volume Profiles for Chemical and Biochemical Processes*

**Examining Committee:** **Chair: Dr. Charles J. Walsby**  
Associate Professor

**Dr. Michael H. Eikerling**  
Senior Supervisor  
Professor

**Dr. Noham Weinberg**  
Co-Supervisor  
Adjunct Professor

**Dr. Tim Storr**  
Supervisor  
Associate Professor

**Dr. Nancy Forde**  
Supervisor  
Professor  
Department of Physics

**Dr. Andrew J. Bennet**  
Internal Examiner  
Professor

**Dr. Catherine A. Royer**  
External Examiner  
Professor  
Department of Biological Sciences  
Rensselaer Polytechnic Institute

**Date Defended/Approved:** April 16, 2018

## Abstract

Pressure has a profound effect on both the speed and direction of chemical reactions; an effect which is inextricably linked to the change in volume of the system. Experimentally, this effect is described by the activation and reaction volumes, defined as the pressure derivatives of rate constants and equilibrium constants respectively. These quantities provide information about the relationships between the partial molar volumes of the reactant, product and transition state (TS). However, mechanistic interpretation of these volumes has posed a challenge due to the lack of an accurate computational technique for relating the geometrical structure of a reaction system to its volume.

We have formulated a theoretical methodology that can be used to describe and predict the effects of pressure on reaction systems using the concept of a volume profile. These profiles, which detail how the partial molar volume of a system changes over the course of a reaction, can be calculated using the recently developed Archimedean displacement model of molecular volume. We explore two mechanistic applications of these profiles. The first is the calculation of Gibbs energy profiles at elevated pressures, allowing for the prediction of interesting and potentially useful chemical transformations that can occur with pressure. This technique was used to investigate high pressure structural transformations for a radical hydrogen transfer reaction, and also to examine the feasibility of pressure-driven molecular machines. The second application is the elucidation of TS structures from a comparison of the theoretical volume profile with the experimentally-determined activation volume. This technique is especially useful for systems with a high degree of conformational flexibility whose TSs are not readily identified using standard computational methods. We used this method to identify the TS ensemble for a flexible model chain, and for conformational changes in a cyclophane system. To apply this method to larger, more complex systems such as the unfolding of proteins, a reaction coordinate for the process is required. A proper definition for such a coordinate was investigated and some preliminary results are presented for biological systems.

**Keywords:** high pressure; volume profile; activation volume; Gibbs energy profile; transition state ensemble; reaction coordinate

*To my friends and family for their unending support, patience, love and understanding. Thank you for believing in me.*

## Acknowledgements

First, I would like to thank to my supervisor, Dr. Noham Weinberg, for his many years of encouragement, support and leadership. Thank you for inspiring me and igniting my passion for research. I would also like to extend my gratitude to my committee members, Dr. Michael Eikerling, Dr. Nancy Force and Dr. Tim Storr, for their advice and guidance throughout my time at SFU, and to my internal and external examiners, Dr. Andrew Bennet and Dr. Catherine Royer, for taking the time to review this work.

I would also like to thank the past and current members of the Weinberg group for their friendship and encouragement. In particular, I would like to acknowledge the contribution of three undergraduate students to this work: Abigail Sherwood for the elevated pressure enthalpy profiles in Chapter 2, Melissa Prachnau for the nonequilibrium solvation surfaces in Chapter 2, and Miranda Louwerse for the volume profile of the rotor of Chapter 3 and the volume and energy profiles of the paracyclophane system of Chapter 4.

I gratefully acknowledge financial support from the Natural Sciences and Engineering Council of Canada, the SFU Chemistry Department and the University of the Fraser Valley as well as computing resources from Compute Canada.

# Table of Contents

Approval.....	ii
Abstract.....	iii
Dedication.....	iv
Acknowledgements.....	v
Table of Contents.....	vi
List of Tables.....	ix
List of Figures.....	x
List of Abbreviations.....	xxi
List of Symbols.....	xxii
<b>Chapter 1. Introduction.....</b>	<b>1</b>
1.1. Effects of high pressure on reaction kinetics and equilibria: theoretical principles and experiments.....	1
1.1.1 General principles.....	1
1.1.2 Mechanistic applications for chemical systems.....	3
1.1.3 Application of pressure to biological systems.....	5
1.1.4 Pressure-induced viscosity effects.....	6
1.2. Theoretical calculations of molecular volumes, volumes of activation and volume profiles.....	7
1.2.1 Review of existing volume models.....	7
1.2.1.1 Stearn-Eyring model.....	7
1.2.1.2 Van der Waals volume.....	8
1.2.1.3 Solvent accessible and solvent excluded volumes.....	9
1.2.1.4 Quantum mechanical isodensity model.....	9
1.2.2 Archimedean displacement model.....	10
1.2.3 Volume profiles.....	11
1.2.4 High pressure reaction profiles.....	13
1.3. Computational methodologies.....	15
1.3.1 Molecular dynamics simulations.....	15
1.3.1.1. The molecular dynamics algorithm.....	15
1.3.1.2. Interaction potentials.....	16
1.3.1.3. Periodic boundary conditions.....	21
1.3.1.4. Temperature and pressure control.....	21
1.3.1.5. Thermodynamic integration.....	22
1.3.2. Quantum mechanical calculations.....	23
1.4. Motivation and objectives.....	23
1.5. Brief outline of this work.....	24
<b>Chapter 2. High pressure effect on radical hydrogen abstraction.....</b>	<b>26</b>
2.1. Radical hydrogen abstraction.....	26
2.2. Description of the system.....	28
2.2.1. Reaction coordinate.....	28

2.2.2.	Force field parameterization .....	29
2.3.	High pressure reaction profiles .....	31
2.3.1.	Activation volume and volume profile.....	31
2.3.2.	High pressure Gibbs energy profile.....	32
2.3.3.	High pressure effect on activation enthalpy .....	35
2.4.	Pressure-induced viscosity effect .....	38
2.4.1.	Solute and solvent coordinates.....	38
2.4.2.	Gibbs energy surfaces.....	40
2.5.	Conclusions.....	43
<b>Chapter 3. Pressure-induced conformational changes: Application to molecular machines.....</b>		<b>45</b>
3.1.	Molecular conformations and molecular machines .....	45
3.2.	Bithiophene as a prototype for molecular wires .....	46
3.2.1.	Reference pressure torsion profiles .....	46
3.2.2.	High pressure profiles and pressure controlled conformational changes.....	49
3.3.	Molecular rotors.....	54
3.3.1.	Mathematical model of pressure-controlled rotor .....	57
3.3.2.	Energy and volume profiles of possible molecular candidates .....	62
3.4.	Conclusions.....	66
<b>Chapter 4. Theoretical volume profiles as tools for probing transition states and transition state ensembles.....</b>		<b>68</b>
4.1.	The concept.....	68
4.2.	Model flexible chain as a prototype for protein folding .....	70
4.2.1.	System .....	70
4.2.2.	Gibbs energy profile and expected TSE.....	71
4.2.3.	MD high pressure kinetics .....	72
4.2.4.	Volume profile and predicted TSE .....	74
4.3.	Conformational changes in paracyclophane .....	76
4.3.1.	System .....	76
4.3.2.	Experimental high pressure kinetic data .....	82
4.3.3.	MD volume profile and predicted TSE .....	83
4.3.4.	Validation from Gibbs energy profile .....	86
4.4.	Conclusions.....	87
<b>Chapter 5. Reaction coordinates for systems with multiple reaction pathways..</b>		<b>88</b>
5.1.	Introduction.....	88
5.2.	Collective variables.....	89
5.2.1.	Brief overview .....	89
5.2.2.	Best-Hummer optimization.....	92
5.3.	Comparative analysis of reaction coordinates for transformations in a four-state system.....	95
5.3.1.	Chemical prototypes .....	95
5.3.2.	Model system .....	97

5.3.3.	2D Kinetics .....	98
5.3.4.	Linear combination coordinate .....	100
5.3.5.	RMSD coordinate .....	103
5.3.6.	Piecewise defined coordinate .....	104
5.4.	Reaction coordinates for multi-minima domains.....	109
5.4.1.	Model surface .....	109
5.4.2.	A single-minimum reactant domain.....	110
5.4.3.	A two-minima reactant domain .....	112
5.4.4.	A three-minima reactant domain.....	115
5.4.5.	A four-minima reactant domain.....	116
5.4.6.	The piecewise coordinate .....	117
5.5.	Conclusions.....	121
<b>Chapter 6. Towards applications to biological systems.....</b>		<b>122</b>
6.1.	Reaction coordinate for the unfolding of a model helix.....	122
6.1.1.	System .....	122
6.1.2.	Definition of reactant and product states.....	124
6.1.3.	Trial reaction coordinate: Weighted average angle .....	125
6.2.	Kinetic network analysis for an $\alpha$ -helical peptide .....	128
6.2.1.	The Kieffer peptide .....	128
6.2.2.	Model system: Polyalanine .....	129
6.2.3.	Coarse-grained kinetic network for polyalanine unfolding .....	132
6.3.	Staphylococcal nuclease .....	137
6.4.	Conclusions.....	141
<b>Chapter 7. Concluding remarks and outlook .....</b>		<b>142</b>
<b>References.....</b>		<b>145</b>

## List of Tables

Table 1.1	Typical ranges of $\Delta V^\ddagger$ for different types of reactions. Data from ref. 14.....	3
Table 2.1	Experimentally measured pressure-induced changes in activation Gibbs energy ( $\Delta\Delta G^\ddagger$ ) and activation enthalpy ( $\Delta\Delta H^\ddagger$ ) relative to ambient pressure for the Diels-Alder dimerization of cyclopentadiene. Data from ref. 139. ....	35
Table 2.2	Barrier heights $\Delta H^\ddagger$ as a function of pressure. ....	37
Table 3.1	The change in the electronic properties of the models assessed by change in HOMO-LUMO energy gap $\Delta\Delta E_{\text{HOMO/LUMO}}$ between the conformation at ambient pressure ( $\phi_0$ ) and the high pressure structure ( $\phi = 180^\circ$ ). For reference, the $\Delta\Delta E_{\text{HOMO/LUMO}}$ for the same structural change in butadiene is provided. Energies were calculated at the B3LYP/6-31++G(d,p) level. ....	54
Table 4.1	TSE positions $r^\ddagger$ (in nm) for the unfolding of the model chain with different fitting methods. Deviations from the TSE predicted by the Gibbs energy profile are listed in parentheses.....	76
Table 4.2	Gibbs energies for four TSE candidates calculated using eq. 4.9. ....	86
Table 6.1	Experimental $\Delta V^\ddagger$ and $\Delta V$ data for snase.....	138

## List of Figures

Figure 1.1	The A1 (top) and A2 (bottom) mechanisms of acetal hydrolysis. The A1 mechanism is stepwise and occurs via a carbocation intermediate, while the A2 mechanism is concerted. ....	4
Figure 1.2	The stepwise di-radical (top) and concerted mechanisms (bottom) of the Diels-Alder reaction between butadiene and ethene.....	5
Figure 1.3	The Stearn-Eyring cylindrical model of molecular volume, shown here for the example of an $S_N2$ reaction. ....	8
Figure 1.4	The van der Waals molecular volume model, shown for the example of <i>n</i> -hexane. The volume is calculated from the set of overlapping van der Waal spheres centered on each atom. ....	8
Figure 1.5	Solvent accessible (left) and solvent excluded (right) models of molecular volume. The difference between these two models lies in whether the molecular surface is defined by the center of the probe sphere, or the edge. ....	9
Figure 1.6	Isodensity surface for <i>n</i> -hexane, with isodensity level = 0.001 a.u.....	10
Figure 1.7	Four hydrocarbons (clockwise from top left corner: <i>n</i> -hexane, cyclopentadiene, benzene and toluene) immersed in a model solvent represented here by an overlay of solvent configurations at different instants of time. The solvent particles (white) avoid the solute, forming a cavity which matches the size and shape of the solute molecule.....	11
Figure 1.8	Experimentally-determined 3-point volume profile for the Diels-Alder reaction between isoprene and methyl acrylate. Figure from ref. 117, Copyright © 2002 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc. ....	12
Figure 1.9	Schematic continuous volume profile (red line) connecting the experimentally-determined $\Delta V^\ddagger$ and $\Delta V$ along the reaction coordinate. Experimental data is the same as in Figure 1.8. ....	13
Figure 1.10	An illustration of the use of eq. 1.10 for a hypothetical two-state system. Left: the reference-pressure $\Delta G_o(x)$ (blue) and $\Delta V_o(x)$ (red dashed) profiles for the system. Right: a series of high pressure $\Delta G(x;P)$ profiles showing a change in the relative stabilities of the reactant and product state, as well as the pressure dependence of the reaction barriers. At the highest pressure, the less compact state ceases to exist. ....	14
Figure 1.11	The workflow for an MD simulation in GROMACS.....	16
Figure 1.12	(a) The Lennard-Jones potential with $\sigma$ , $\epsilon$ and $r_e$ labelled. (b) The Coulomb potential, shown for the repulsive interaction between two positively charged particles. ....	18
Figure 1.13	The harmonic potential, which is used to describe both bond stretching and angle bending interactions. Note that the x-axis can be either $r$ or $\theta$ , depending on the interaction being modeled. The equilibrium bond or angle value is located at the bottom of the parabola.....	19

Figure 1.14	The Morse potential, used to describe bond stretching for systems that cannot be accurately represented by the harmonic potential of Figure 1.13.....	19
Figure 1.15	Sample Ryckaert-Belleman potentials for rotation around a single bond (three-fold potential on the left) and a double bond (two-fold potential on the right). .....	20
Figure 1.16	The use of periodic boundary conditions allows particle <i>i</i> to interact with a translated image of particle <i>j</i> . In order to eliminate self-interaction, a cut-off radius is imposed for all through-space interactions (dashed circle). .....	21
Figure 2.1	The hydrogen abstraction reaction between methane and a methyl radical. ....	27
Figure 2.2	The effect of compression on the PES of the methyl-methane hydrogen abstraction system for (a) ambient conditions, (b) low, (c) medium, and (d) high compression. As compression increases, the reaction barrier decreases and the positions of the reactant and product begin to shift towards the TS. The blue arrows trace the position of the TS. The contour spacing is 10 kJ/mol. Figure reprinted from ref. 132, © 2016 Canadian Science Publishing or its licensors. Reproduced with permission. ....	27
Figure 2.3	(a) The two C-H bond lengths labelled $r_1$ and $r_2$ , which are used in the definition of the reaction coordinate for this system. (b) Gas phase B3LYP/6-31++G(d,p) potential energy surface in $r_1$ and $r_2$ coordinates. The white dashed line is the reaction path. ....	28
Figure 2.4	Gas phase B3LYP/6-31++G(d,p) energy profile $\Delta U_{\text{gas}}(x)$ along reaction coordinate $x$ . ....	29
Figure 2.5	The pyramidity parameter $\gamma$ of each carbon atom was assessed using the sum of the three H-C-H bond angles labelled in yellow or green.....	30
Figure 2.6	The ambient pressure volume profile $\Delta V_0(x)$ for the methyl-methane system (left). The dotted line is a fit to the profile obtained from a linear plot of volume against the distance $L$ between carbon atoms, $L = r_1 + r_2$ (right). ....	32
Figure 2.7	The ambient pressure Gibbs energy profile $\Delta G_0(x)$ (solid purple line) obtained as a sum of the gas phase potential energy $\Delta U_0(x)$ (blue data points) and the solvation Gibbs energy $\Delta G_{\text{solv}}(x)$ (red data points). ....	33
Figure 2.8	The calculated $\Delta G(x; P)$ profiles show that as pressure increases, the reaction barrier decreases. A pressure of 125 kbar was required to eliminate the reaction barrier. Above that pressure, the Gibbs energy profile possesses a single minimum at $x = 0$ representing a symmetric structure $\text{CH}_3\text{HCH}_3$ as a single stable state of the system. ....	34
Figure 2.9	The effect of pressure on the calculated $\Delta G^\ddagger$ values. The reaction barrier becomes practically negligible above 100 kbar. ....	35
Figure 2.10	The thermal expansion coefficient $\alpha = 1/V(\partial V/\partial T)_P$ for <i>n</i> -hexane decreases with increasing pressure, due to the tighter packing of the	

	molecules at higher pressures. The $\alpha$ values were obtained by numerical differentiation of the equation of state from data in ref. 140. ....	36
Figure 2.11	Enthalpy profiles for the methyl-methane system at four different pressures. The MD-calculated data points were fit with a Lorentzian function as a visualization aid. ....	37
Figure 2.12	Two types of Gibbs energy surfaces in solvent-solute coordinates: (a) strong solute-solvent coupling; (b) weak solute-solvent coupling. ....	39
Figure 2.13	Illustration of the solvent coordinate construct. The solvent-solute system is first equilibrated to the pilot solute in some configuration $y$ (left). The pilot solute is removed, but the solvent trajectory remains equilibrated to it (middle). The pilot solute is then replaced with the solute in a different configuration $x$ (right). ....	40
Figure 2.14	Nonequilibrium solvation energy for the methyl-methane system in solvent-solute coordinates. The properly solvated reactant is in the bottom left corner, and the properly solvated product in the top right. ....	41
Figure 2.15	Gibbs energy surface for the methyl-methane system in solvent-solute coordinates. The properly solvated reactant is in the bottom left and the properly solvated product is in the top right. This surface has two parallel valleys separated by the nearly uniform activation barrier, which implies that the kinetics of this reaction is nearly independent of solvent viscosity. ....	42
Figure 3.1	The three bithiophene systems used as models for polythiophene molecular wires. ....	46
Figure 3.2	Gas phase QM potential energy profiles for the three bithiophene models, 1 (black circles), 2 (red squares) and 3 (blue triangles). Lines are fit to smooth computational errors. ....	47
Figure 3.3	Gas phase potential energy (light data points, dotted lines) and solvation Gibbs energy (dark data points, dashed lines) summed to yield the reference pressure Gibbs energy profiles (solid lines) for the three bithiophene models. Dashed and dotted lines are fit to smooth computational errors. ....	48
Figure 3.4	Volume profiles $\Delta V_o(\phi)$ for rotation around $\phi$ in bithiophene models 1, 2, and 3. Error bars were calculated using block averaging. The portion of the volume profile around $180^\circ$ was fitted with a quadratic function to smooth out computational errors (dotted line). ....	50
Figure 3.5	High pressure Gibbs energy profiles $\Delta G(\phi; P)$ for bithiophene 1. Increasing pressure to 80 kbar causes the minimum of $\Delta G(\phi; P)$ to shift from $151^\circ$ to $180^\circ$ . ....	51
Figure 3.6	High pressure Gibbs energy profiles $\Delta G(\phi; P)$ for bithiophene 2. A pressure of 160 kbar is required to shift the minimum of $\Delta G(\phi; P)$ from $128^\circ$ to $180^\circ$ . ....	52
Figure 3.7	High pressure Gibbs energy profiles $\Delta G(\phi; P)$ for bithiophene 3. The increase in the Gibbs energy barrier at $180^\circ$ in $\Delta G_o(\phi)$ coupled with the relatively small volume change from model 2 means that model 3 requires a high pressure of 500 kbar to drive the system into its planar conformation. ....	53

Figure 3.8	The first synthetic chemically-driven rotor. <sup>168</sup> This system relies on a series of chemical modifications to the amine group to promote rotation of the three bladed propellor 120° in the clockwise direction. Reprinted by permission from Springer Nature: Nature, Unidirectional rotary motion in a molecular system, T. Ross Kelly, Harshani De Silva and Richard A. Silva. © 1999.....	55
Figure 3.9	The first photo-driven rotor. <sup>169</sup> Irradiation results in isomerization around the central double bond, and steric interactions promote rotation in a single direction. Reprinted by permission from Springer Nature: Nature, Light-driven monodirectional molecular rotor, Nagatoshi Koumura, Robert W. J. Zijlstra, Richard A. van Delden, Nobuyuki Harada & Ben L. Feringa. © 1999.....	56
Figure 3.10	A schematic of a hypothetical pressure-driven molecular rotor with energy and volume profiles shown in Figure 3.11. The mobile part (red) rotates within the stationary strap (blue) by a properly designed pressure pulse (see text).....	57
Figure 3.11	A hypothetical molecular rotor system described by a four-fold ambient Gibbs energy profile $\Delta G_o(\phi)$ (blue) and a two-fold volume profile $\Delta V_o(\phi)$ (red). For convenience, both $\Delta G_o(\phi)$ and $\Delta V_o(\phi)$ are dimensionless. Labels A to I serve to relate conformations of the rotor to the respective maxima and minima of $\Delta G_o(\phi)$ .....	58
Figure 3.12	The periodic two stage pressure pulse sequence. The first part of the pulse (dashed line) is in the negative direction, followed by a positive pressure pulse (dotted line).....	59
Figure 3.13	Progression of the pressure-dependent Gibbs energy profiles $\Delta G(\phi; P)$ for the hypothetical molecular rotor utilising the two stage pressure pulse of Figure 3.12. The red circles trace the shift in the position of the system with pressure. The conformational changes of the rotor for each pressure pulse are illustrated near their respective minima. ....	60
Figure 3.14	An exact alignment of extrema in $\Delta V_o(\phi)$ (red) and $\Delta G_o(\phi)$ (blue solid) produces unwanted behaviour in the high pressure $\Delta G(\phi; P)$ (blue dashed). (a) The maximum of $\Delta G_o(\phi)$ aligned with the minimum of $\Delta V_o(\phi)$ results in a 50% chance of reverse rotation. (b) The alignment of the minima of both profiles results in no change in the conformation of the system at high pressure. ....	61
Figure 3.15	A 1:1 ratio between $\Delta G_o(\phi)$ (blue solid) and $\Delta V_o(\phi)$ (red) results in insufficient movement of the high pressure $\Delta G(\phi; P)$ minimum (blue dashed). Therefore, upon release of pressure, the system returns to the previous minimum of $\Delta G_o(\phi)$ .....	61
Figure 3.16	A 3:1 ratio between $\Delta G_o(\phi)$ (blue solid) and $\Delta V_o(\phi)$ (red) results in the system becoming stuck between 120° and 180°. (a) A negative pressure pulse (dark blue dashed) eliminates the barrier at 150° and causes the system to fall into the 180° minimum, where it remains. (b) The next positive stage of the pressure pulse (light blue dashed) results in the 180° minimum becoming destabilized, and the system falls back into the 120° minimum. ....	62

Figure 3.17	A real molecular rotor <b>4</b> designed based on the hypothetical system of section 3.3.1. Torsion angles $\phi$ and $\tau$ are shown in blue and green respectively. Both angles use the two yellow carbons as reference.....	63
Figure 3.18	(a) The gas phase potential energy profile $\Delta U_{\text{gas}}(\phi)$ of <b>4</b> , showing 2-fold periodicity as with the hypothetical rotor. (b) Volume profile $\Delta V_o(\phi)$ for <b>4</b> obtained using 1000 Lennard-Jones particles as solvent with $\sigma = 0.35$ nm. Error bars are too small to be seen on this scale.....	63
Figure 3.19	Cavity images for rotor <b>4</b> with a $\sigma = 0.35$ nm solvent: Top – flat conformation ( $\phi = 0^\circ$ ); bottom – orthogonal conformation ( $\phi = 90^\circ$ ); left – front view; right – side view.....	64
Figure 3.20	Volume profiles $\Delta V_o(\phi)$ for <b>4</b> obtained with (a) $\sigma = 0.45$ nm solvent and (b) with $\sigma = 0.25$ nm solvent. The larger solvent exacerbates the unwanted volume dip near $\phi = 90^\circ$ whereas the smaller solvent smoothes $\Delta V_o(\phi)$ towards a single maximum at $\phi = 90^\circ$ (the dashed line, showing the general shape of the “ideal” profile of Figure 3.11).....	65
Figure 3.21	Cavity images for rotor <b>4</b> with a $\sigma = 0.25$ nm solvent: Top – flat conformation ( $\phi = 0^\circ$ ); bottom – orthogonal conformation ( $\phi = 90^\circ$ ); left – front view; right – side view.....	65
Figure 4.1	An illustration of the use of eq. 4.3. Combination of the experimentally determined $\Delta V^\ddagger$ and the MD-calculated $\Delta V(x)$ yields the location of the unknown TS. ....	69
Figure 4.2.	30 unit model chain. The distance, $r$ , between terminal atoms serves as the reaction coordinate for the unfolding process. ....	70
Figure 4.3	(a) The averaged Gibbs energy profile for unfolding of the model chain (black), fitted with a cubic parabola (grey) to obtain the maximum. (b) The quality of the profile is assessed by the level of convergence between data from three MD trajectories with the initial $r$ values of 0.44 nm (blue), 0.9 nm (red) and 1.5 nm (green). ....	71
Figure 4.4	A sample survival probability plot, $\ln N$ , for $r^\ddagger = 0.638$ nm and $P = 1$ kbar.....	73
Figure 4.5	Fit to MD simulated high pressure kinetic data with linear (red) and quadratic (green) regression functions. ....	73
Figure 4.6	Volume profile for the unfolding of the model chain system. The region of interest (between 0.44 nm and 1.20 nm) was fitted with both a linear (red) and quadratic (green) regression function. ....	74
Figure 4.7	Correlation between volume profile $\Delta V(r)$ (blue) and the entropy profile $T\Delta S(r)$ (green) for the unfolding of the model chain. The inset shows a direct linear correlation between the two. ....	75
Figure 4.8	The structure of 1,12-dimethoxy-[12]-paracyclophane. ....	77
Figure 4.9	The torsion angles that control the rotation of the benzene ring: (a) $\phi$ , in blue, and (b) $\tau$ , in green. Both angles are defined relative to one terminal segment of the tether, shown in yellow. ....	77
Figure 4.10	A Newman projection showing the relationship between $\phi$ and $\tau$ . Benzene ring is represented as the bold line.....	78

Figure 4.11	Gas phase AM1 potential energy surface of <i>1</i> in $\phi$ and $\tau$ coordinates. The chosen constant $\tau$ reaction path (dotted line) has its highest point (white diamond) at $\phi = 34.5^\circ$ and $\tau = 6.5^\circ$ , which is a reasonable approximation of the true TS (white star) at $\phi = 30^\circ$ and $\tau = 3^\circ$ . . . . .	78
Figure 4.12	The gas phase B3LYP/6-31++G(d,p) potential energy profile for ring rotation along $\phi$ . The minima are at $-70^\circ$ and $110^\circ$ , and maxima at $-150^\circ$ and $30^\circ$ . . . . .	79
Figure 4.13	The structure of <i>para</i> -di(1'-methoxyhexyl)benzene. . . . .	79
Figure 4.14	The gas phase B3LYP/6-31++G(d,p) potential energy profile for rotation in <i>2</i> . The severing of the tether into two independent chains has restored the perfect periodicity of the profile. . . . .	80
Figure 4.15	The MD-calculated potential energy profile (red) agrees well with the QM-calculated profile of Figure 4.12 (black). . . . .	80
Figure 4.16	The effect of the position of the ring on the conformational ensemble of the tether illustrated here for $\phi = -150^\circ$ (red), $-70^\circ$ (green) and $30^\circ$ (red). The front (left) and side (right) views were generated as overlays of instantaneous conformations of <i>1</i> sampled along MD trajectories generated for constant $\tau = 6.5^\circ$ . . . . .	81
Figure 4.17	Experimental high pressure kinetic data (black data points) from ref. 217, fitted with eq. 4.7. . . . .	82
Figure 4.18	The pressure dependence of $\Delta V^\ddagger$ as described by eq. 4.8 (red line). Our calculated zero pressure value $\Delta V^\ddagger = -12.8 \text{ cm}^3/\text{mol}$ (blue diamond) constitutes a starting point. The experimentally-determined $\Delta V_{\text{ave}}^\ddagger$ values (black squares) are consistent with the graph. . . . .	83
Figure 4.19	Volume profile $\Delta V(\phi)$ (black data points) for one $180^\circ$ rotation of the ring along the $\phi$ reaction coordinate. The experimental $\Delta V^\ddagger$ value is shown in red. Top: solvent cavities corresponding to selected values of $\phi$ . . . . .	84
Figure 4.20	Further detail in the TSE region of the volume profile. There are four points, indicated by arrows, at which $\Delta V(\phi)$ (black data points) appears to cross the experimental $\Delta V^\ddagger$ . The value of $\phi$ at the intersection points was determined using linear interpolation (blue dashed lines). . . . .	85
Figure 4.21	Gibbs energy profile for a single $180^\circ$ rotation of the benzene ring (left) and its expansion (red rectangle) in the vicinity of the maximum (right). The TSE (the maximum of the profile) is located at $\phi^\ddagger = 10^\circ$ . . . . .	86
Figure 5.1	(a) Schematic free energy surface for the folding of a protein, showing an extremely rugged landscape. (b) Transformations between the folded (F) and unfolded (U) domains can occur through a number of different reaction pathways. . . . .	88
Figure 5.2	Gibbs energy profiles for the model chain system of Chapter 4 as a function of (a) $R_g$ and (b) $r$ . A change of coordinate results in dramatic changes in both the barrier height and the relative stability of the reactant and product states. . . . .	91
Figure 5.3	A 2D Gibbs energy surface for the model chain system in $r$ and $R_g$ coordinates. The position of the TSE is denoted by the black circle	

	and hypersurfaces orthogonal to $r$ and $R_g$ are denoted by the dashed lines. ....	92
Figure 5.4	Schematic of the configurational space of a flexible system with two choices of collective variables: (a) $x = q_1$ and (b) $x = 1/2(q_1 + q_2)$ . See text for details. ....	93
Figure 5.5	Probability that a point on a hypersurface of constant $x$ lies in the transition path region for (a) $x = q_1$ and (b) $x = 1/2(q_1 + q_2)$ as defined in Figure 5.4. Red circles label positions of the individual TSEs of the TSE. ....	93
Figure 5.6	Long equilibrium trajectory that visits reactant ( $R$ ) and product ( $P$ ) domains multiple times. The transition path segments of this trajectory are shown in green, and non-reactive segments are shown in red. The $P(\text{TP} x)$ distribution (bottom) reaches its highest value in the middle of the TP region. ....	94
Figure 5.7	Two possible reaction pathways for the rotation of the carboxylic acid groups in 2,5-dihydroxyterephthalic acid resulting in restructuring of the two intramolecular hydrogen bonds (red dashed lines). ....	95
Figure 5.8	Four-state B3LYP/6-31++G(d,p) PESs for terephthalic acid (top) and 2,5-dihydroxyterephthalic acid (bottom) showing two possible reaction pathways. Torsion angles $\phi_1$ and $\phi_2$ are as defined on the left. ....	96
Figure 5.9	Top: Model linear triatomic system with two bonds labelled $r_1$ and $r_2$ . Bottom left: The double well 1D potential used to describe interactions between neighboring ("bonded") atoms in this system. Bottom right: the four-state 2D PES for the triatomic obtained as a combination of the 1D potentials. ....	97
Figure 5.10	MD-generated Gibbs energy surface $\Delta G(r_1, r_2)$ for the triatomic system, labelled with the location of states A, B, C and D. ....	98
Figure 5.11	Kinetic schemes for the depletion of the A state (left) and for the accumulation of the D state (right). ....	99
Figure 5.12	Kinetic curves for (a) exit from the A state and (b) entry into the D state, determined using MD simulations (grey circles) and the kinetic schemes of Figure 5.11 (black squares). ....	99
Figure 5.13	(a) The optimized coordinate (white dashed line) shown on the GES for the triatomic system. (b) The one-dimensional $\Delta G(x)$ profile along this coordinate. ....	100
Figure 5.14	1D kinetic scheme representing projections of 2D reaction networks of Figure 5.11 onto the optimized reaction coordinate for depletion of the reactant state (left) and accumulation of the product state (right). ..	101
Figure 5.15	Kinetic curves for (a) depletion of the reactant state and for (b) accumulation of the product state for the kinetic network from the 2D surface (black) and for the 1D profile for the optimized linear coordinate (red). ....	101
Figure 5.16	(a) Hypersurfaces of constant $x$ at the reactant $x = 0.35$ (black), and the TS $x = 0.375$ (grey). (b) The Gibbs energy profiles along these	

	hypersurfaces. The maxima and minima of the $x = 0.375$ profile are labeled by red triangles and orange circles respectively.....	102
Figure 5.17	(a) The optimized circular coordinate (white concentric circles) shown on the GES for the triatomic system. (b) The one-dimensional $\Delta G(x)$ profile along this coordinate. ....	103
Figure 5.18	Kinetic curves for the depletion of the reactant state for the kinetic network from the 2D surface (black) and for the 1D profile for the optimized RMSD coordinate (red). ....	104
Figure 5.19	The hypersurfaces of constant $x$ (dashed lines) for (a) $x$ defined by eq. 5.9 (the A domain), and (b) $x$ defined by eq. 5.10 (the D domain)...	105
Figure 5.20	Constant $x$ hypersurfaces in the intermediate zone (i.e. the B and C domains), providing an interpolation between the border of the A domain (black) and the border of the D domain (white). ....	106
Figure 5.21	Hypersurfaces of constant $x$ for the piecewise reaction coordinate defined by eq. 5.12.....	106
Figure 5.22	Sample $P(TP x)$ distributions for the piecewise coordinate of eq. 5.12. Only $a = b = 0.4$ nm produces a distribution contained entirely within the intermediate region of $1 < x < 2$ . ....	107
Figure 5.23	One dimensional profile $\Delta G(x)$ along the optimal two dimensional coordinate. ....	108
Figure 5.24	Comparison of the kinetic curves derived from the 2D kinetic networks of Figure 5.11 (black) and from the 1D profile for the optimized piecewise coordinate of eq 5.12 (red) for (a) the decomposition of the reactant and (b) the accumulation of the product. ....	108
Figure 5.25	The model 16-state Gibbs energy surface of eq. 5.13. Blue configurations are minima, and red configurations are maxima. ....	109
Figure 5.26	The boundary of the $R_1$ domain (white solid line) with hypersurfaces of constant $x_{R_1}$ (white dashed lines). Note that for clarity the surface has been zoomed in. ....	110
Figure 5.27	Gibbs energy profile $\Delta G(x_{R_1})$ along the $x_{R_1}$ coordinate for the $R_1$ domain. ....	111
Figure 5.28	(a) Kinetic scheme for exit from the $R_1$ domain (represented by the grey square) on the 2D GES. (b) Comparison of the kinetic curves derived from the 2D kinetic network (black) and from the 1D $\Delta G(x_{R_1})$ (red). ....	112
Figure 5.29	(a) The boundary of the $R_2$ domain (solid white line) and hypersurfaces of constant $x_{R_2}$ (dashed white lines). (b) The 2D reaction network for exit from the $R_2$ domain, represented by the grey rectangle. ....	113
Figure 5.30	(a) Gibbs energy profile $\Delta G(x_{R_2})$ for the $R_2$ domain along the $x_{R_2}$ coordinate defined by eq. 5.17. (b) A comparison of the kinetic curve derived from the full reaction network of Figure 5.29(b) (black) with the two-state kinetics from $\Delta G(x_{R_2})$ (red) shows a poorer quality match. ....	113

Figure 5.31	(a) The hypersurfaces of constant $x_{R_2}$ (white dashed lines) for the redefined coordinate of eq. 5.18. The boundary of the $R_2$ domain (solid line) has not changed. (b) Gibbs energy profile $\Delta G(x_{R_2})$ for the $R_2$ domain along redefined $x_{R_2}$ coordinate. (c) A comparison of the kinetic curves derived from the 2D reaction network of Figure 5.29(b) (black) and from the new 1D profile $\Delta G(x_{R_2})$ (red). .....	114
Figure 5.32	(a) The boundary of the $R_3$ domain (white solid line), encapsulating the $S_{11}$ , $S_{21}$ and $S_{12}$ minima. Hypersurfaces of constant $x_{R_3}$ are shown as white dashed lines. (b) The 2D reaction network for exit from the $R_3$ domain (grey). .....	115
Figure 5.33	(a) The Gibbs energy profile $\Delta G(x_{R_3})$ for the $R_3$ domain along the $x_{R_3}$ coordinate defined by eq. 5.19. (b) A comparison of the kinetic curves derived from the 2D reaction network of Figure 5.29(b) (black) and from the 1D profile $\Delta G(x_{R_3})$ (red). .....	116
Figure 5.34	(a) The boundary of the $R_4$ domain (white solid line), encapsulating the $S_{11}$ , $S_{21}$ , $S_{12}$ and $S_{22}$ minima. Hypersurfaces of constant $x$ are shown as white dashed lines. (b) The 2D reaction network for exit from the $R_4$ domain (grey square). .....	116
Figure 5.35	(a) The Gibbs energy profile for the $R_4$ domain. (b) A comparison of the kinetic curves derived from the 2D reaction network (black) and from the 1D profile (red). .....	117
Figure 5.36	(a) The reactant A, intermediate B,C and product D domains for the 16-state surface, each domain consisting of 4 minima. (b) The hypersurfaces of constant $x$ for the piecewise coordinate.....	118
Figure 5.37	The Gibbs energy profile $\Delta G(x)$ along the piecewise coordinate $x$ , with minima labelled R-1 to P-2.....	118
Figure 5.38	Kinetic schemes for accumulation of product for (a) the Gibbs energy profile $\Delta G(x)$ and (b) the 2D surface. The states that comprise the D domain are highlighted with the grey square. (c) The resulting kinetic curves obtained from the 1D profile (red) and the 2D surface (black) agree well with each other.....	119
Figure 5.39	The modified 16-well surface with increased activation barriers between the domains.....	120
Figure 5.40	(a) The Gibbs energy profile $\Delta G(x)$ along the piecewise coordinate $x$ , with minima labelled R-1 to P-2. These minima correspond to the labels used in Figure 5.37(b). The kinetic curves obtained using 1D (red) and 2D (black) reaction networks show good agreement for both (b) depletion of reactant and (c) accumulation of product. ....	121
Figure 6.1	The model helix system showing the five dihedral angles $\phi_1$ - $\phi_5$ (left) and the two hydrogen bonds $r_1$ and $r_2$ (right). .....	122
Figure 6.2	(a) Double well potential for the dihedral angles in the model helix. The unfolded state at $180^\circ$ is slightly lower in energy than the folded state at $60^\circ$ . (b) The Morse bond potential for the two hydrogen bonds in this system. ....	123

Figure 6.3	Probability distribution functions for the five dihedral angles in the model helix grouped by symmetry: $\phi_1$ and $\phi_5$ (solid line), $\phi_2$ and $\phi_4$ (dashed line) and $\phi_3$ (dotted line).....	124
Figure 6.4	Population density for (a) $\phi_5$ and (b) $\phi_3$ as a function of time. For $\phi_3$ , there is a clear region of low population density between $70^\circ < \phi_3 < 90^\circ$ . .....	125
Figure 6.5	$P(\text{TP} \text{x})$ distributions for a selection of different weights: equal weighting (green), only $\phi_1$ and $\phi_5$ (dark blue), only $\phi_2$ and $\phi_4$ (yellow) only $\phi_3$ (light blue) and the optimized weighting (red). .....	126
Figure 6.6	The $P(\text{TP} \text{x})$ distribution (black) and the component distributions $P(\text{x})$ (light grey dashed) and $P(\text{x} \text{TP})$ (dark grey dashed) for (a) sample weights $w_2 = w_4 = 0.5$ , $w_3 = 0$ and (b) the optimized weights. ....	127
Figure 6.7	(a) The three-state model used to describe the unfolding of the Kieflaber peptide. The helix must unfold in the region between the xanthone (blue) and naphthalene (red) groups in order for triplet-triplet energy transfer (TTET) to occur. (b) Bi-exponential experimental kinetic data for TTET at different pressures. Figure reprinted from ref. 271 with permission. ....	128
Figure 6.8	The structure of the model polyaniline peptide showing the orientation of the phenylalanine R groups in the folded state: (a) a side view and (b) a view from above. ....	130
Figure 6.9	Polyexponential kinetic curve for the quenching of the polyaniline peptide obtained via first exit MD simulations. ....	131
Figure 6.10	(a) The $\psi$ and $\phi$ backbone torsion angles. (b) Ramachandran plot for the polyaniline helix. The folded and unfolded domains are cleanly separated from each other along the $\psi$ axis. ....	131
Figure 6.11	Coarse-grained kinetic network for the unfolding and quenching of polyaniline. ....	132
Figure 6.12	(a) Survival probability curve for the $S_0 \rightarrow S_1$ transition. The darker portion of the curve was used to determine the rate constant for the transition. (b) The survival probability curve for the $S_4 \rightarrow S_3$ , showing highly nonlinear behaviour. (c) Comparison of the kinetic curves for the coarse-grained kinetic network (red) and MD simulations (black). ....	133
Figure 6.13	Survival probability curves for residence in (a) the $S_0$ state and (b) the $S_4$ state. The darker portion of each curve was used to determine the rate constant. (c) Comparison of the kinetic curves for the new partitioning method (green) and MD simulations (black). ....	134
Figure 6.14	Markov chain model where quenching transitions are not considered. ....	135
Figure 6.15	Population vs. time curves for (a) $S_0$ , (b) $S_1$ , (c) $S_2$ , (d) $S_3$ , (e) $S_4$ and (f) $S_5$ based on the kinetic network of Figure 6.14 where quenching is no longer considered. For all plots, the black curve is from MD simulations and the green curve is from the kinetic network. ....	135
Figure 6.16	Probabilities of exit from state $S_2$ forward into state $S_3$ (black) and back into state $S_1$ (red). ....	136

Figure 6.17	Population vs. time curves for (a) $S_0$ , (b) $S_1$ , (c) $S_2$ , (d) $S_3$ , (e) $S_4$ and (f) $S_5$ . The black curves are obtained from MD simulations and the green curves are from the kinetic network of Figure 6.14 applied to the model helix of section 6.1. ....	136
Figure 6.18	(a) The staphylococcal nuclease protein (PDB: 1SNO) shown in red at the N-terminus and blue at the C-terminus. (b) The tryptophan residue in snase is shown in the coloured space-filling representation. ....	137
Figure 6.19	The value of the TRP-water RDF taken at $r = 0.4$ nm clearly shows an increasing access of water to the tryptophan as the trajectory progresses. The dashed and dotted lines show best fit lines to the data before and after 600 ps. ....	139
Figure 6.20	Snapshots of the unfolding trajectory of snase at 0 ps (left), 600 ps (middle) and 1 ns (right). The amino acid side chains which are close to the tryptophan in the folded state are shown in the grey space-filling representation. At 600 ps, the surrounding residues have moved away from the TRP, allowing the access of water. ....	139
Figure 6.21	(a) The amino acid side chains that are in close proximity to tryptophan in the folded structure of snase. (b) Correlation of the average distance between the centers of mass of these amino acids and tryptophan with the value of the TRP-water RDF $g$ at 0.4 nm. ....	140

## List of Abbreviations

---

Abbreviation	Definition
1D	One-dimensional
2D	Two-dimensional
B3LYP	Becke, 3-Parameter, Lee-Yang-Parr functional
BEBO	Bond energy bond order
GES	Gibbs energy surface
MD	Molecular dynamics
OPLS	Optimized potential for liquid simulations
PDF	Probability density function
PES	Potential energy surface
QM	Quantum mechanics
RDF	Radial distribution function
RMSD	Root mean square deviation
SAXS	Small angle X-ray scattering
Snase	Staphylococcal nuclease
TRP	Tryptophan
TS	Transition state
TSE	Transition state ensemble
TTET	Triplet-triplet energy transfer

---

## List of Symbols

Symbol	Definition
$A$	BEBO parameter
$C_n$	Parameter for Ryckaert-Bellemans potential
$D$	Depth parameter for Morse potential
$E$	Energy
$E_{kin}$	Kinetic energy
$\Delta\Delta E_{HOMO/LUMO}$	HOMO-LUMO energy gap
$\mathbf{F}_i$	Atomic force vector
$G$	Gibbs energy
$\Delta G$	Gibbs energy of reaction
$\Delta G^\ddagger$	Activation Gibbs energy
$\Delta G_o(x)$	Gibbs energy profile at ambient pressure
$\Delta G(x; P)$	Gibbs energy profile at elevated pressure
$\Delta G_{solv}(x)$	Solvation Gibbs energy profile
$\Delta\Delta G^\ddagger$	Change in activation Gibbs energy
$\mathcal{H}$	Hamiltonian
$\Delta H$	Reaction enthalpy
$\Delta H^\ddagger$	Activation enthalpy
$\Delta\Delta H^\ddagger$	Change in activation enthalpy
$K$	Equilibrium constant
$L$	Length
$M$	Total mass
$N$	Number of particles
$N_{df}$	Number of degrees of freedom
$P$	Pressure
$P_o$	Reference pressure
$P(x)$	Probability
$Q$	Fraction of native contacts
$Q_i$	Atomic charge
$R$	Gas constant
$R_g$	Radius of gyration
$S$	Solvent particle

$\Delta S(x)$	Entropy profile
$T$	Temperature
$U$	Potential energy
$U_{angle}$	Angle bending potential
$U_{bond}$	Bond stretching potential
$U_{es}$	Electrostatic potential
$U_{morse}$	Morse potential
$U_{torsion}$	Torsion potential
$U_{vdW}$	van der Waals potential
$U_{MD}(x)$	Average MD potential energy profile
$\Delta U_{gas}(x)$	Gas phase potential energy profile
$\Delta U_{nb}(x)$	Non-bonded potential energy profile
$\Delta U_{solv}(x, y)$	Nonequilibrium solvation energy
$V$	Volume
$V^\ddagger$	Volume of transition state
$V_P$	Volume of product species
$V_R$	Volume of reactant species
$\Delta V$	Reaction volume
$\Delta V^\ddagger$	Activation volume
$\Delta V_{ave}^\ddagger$	Average activation volume
$\Delta V_o(x)$	Volume profile at ambient pressure
$\Delta V(x; P)$	Volume profile at elevated pressure
$X$	Solute particle
$Z(t)$	Solvent trajectory
$a$	Optimizable parameter
$b$	Optimizable parameter
$g(r)$	Radial distribution function
$h$	Planck constant
$k$	Rate constant
$k_B$	Boltzmann constant
$k_a$	Angle bending force constant
$k_b$	Bond stretching force constant
$k_i$	Partial rate constants

$m_i$	Atomic mass
$n$	Bond order
$n_i$	Number of transitions into state $i$
$q_i$	Structural parameter
$r$	Distance
$\mathbf{r}_{COM}$	Position vector of center of mass
$r_e$	Equilibrium distance
$\mathbf{r}_i$	Atomic position vector
$\mathbf{r}_i^o$	Atomic position vector in reference state
$r_{ij}$	Inter-atomic distance
$r_o$	Equilibrium bond length
$t$	Time
$v_i$	Atomic velocity
$w_i$	Weighting factors
$x$	Reaction coordinate
$z(t)$	Solute trajectory
$\alpha$	Thermal expansion coefficient
$\beta$	Width parameter for Morse potential
$\gamma$	Pyramidal parameter
$\varepsilon_{ij}$	Lennard-Jones potential parameter
$\theta$	Angle
$\theta_{sp^2}$	Trigonal planar angle (120°)
$\theta_{sp^3}$	Tetrahedral angle (109.47°)
$\theta_o$	Equilibrium angle
$\kappa$	Thermal expansion factor
$\lambda$	Coupling parameter for thermodynamic integration
$\lambda_{max}$	Fluorescence emission maximum wavelength
$\sigma$	Cross-sectional area
$\sigma_{ij}$	Lennard-Jones potential parameter
$\tau$	Coupling constant for barostat/thermostat
$\phi$	Dihedral angle & protein backbone dihedral angle
$\psi$	Protein backbone dihedral angle

---

# Chapter 1.

## Introduction

### 1.1. Effects of high pressure on reaction kinetics and equilibria: theoretical principles and experiments

#### 1.1.1 General principles

Temperature and pressure are the two most important physical parameters that determine the selectivity, direction, and rates of chemical reactions. As such, manipulation of these physical conditions is a useful experimental technique for investigation of the thermodynamics and kinetics of a process, and the results can be used to elucidate a mechanism for the reaction. Temperature has received considerable attention in the literature and its effects are well understood. Pressure, which can have effects of comparable scale, has received relatively little attention due to the increased complexity of both experimentation and theory. In the latter case, the major source of this complexity lies in the crucial importance of the solvent medium. Whereas the effects of elevated temperature, which simply increases the internal energy of the reaction system, can be adequately described using gas phase models, the effects of pressure are manifested through the solvent-solute interactions and so an explicit consideration of the solvent medium is required. This is a computationally demanding task that was, until recently, not feasible. The latest advancement in computer technology has considerably changed the balance and we are now in a position to formulate a theoretical method that can be used to both describe and predict the effects of pressure on chemical reactions which, in combination with the appropriate experimental measurements, can turn pressure into a powerful mechanistic tool.

Unlike temperature, which typically has an accelerating effect on reaction rates regardless of their mechanisms, pressure is more sensitive to the reaction mechanism and can either accelerate or decelerate a reaction depending on its type.<sup>1-4</sup> If a reaction results in a decrease in volume of the system, such as in bond formation, an increase in pressure will favour this reaction. This can result in the kinetic effect of an increase in reaction rate, or the thermodynamic effect of a shift of the equilibrium towards the

products. Alternatively, reactions which result in an increased volume, such as bond cleavage, will be disfavoured by pressure and will experience a decrease in reaction rate, or a shift in equilibrium towards the reactants.

In an ideal system, the equilibrium constant  $K$  is related to the standard Gibbs energy of reaction  $\Delta G$  by eq. 1.1.

$$\Delta G = -RT \ln K \quad (1.1)$$

Since volume is the isothermal pressure derivative of Gibbs energy (eq. 1.2), the effect of pressure on chemical equilibria can be quantified using eq. 1.3, as the logarithmic pressure derivative of the equilibrium constant,<sup>5</sup>

$$\left(\frac{\partial \Delta G}{\partial P}\right)_T = \Delta V \quad (1.2)$$

$$\Delta V = -RT \left(\frac{\partial \ln K}{\partial P}\right)_T \quad (1.3)$$

where  $\Delta V$  is the reaction volume, defined as the difference in partial molar volume between the reactant and product species (eq. 1.4).

$$\Delta V = \sum V_P - \sum V_R \quad (1.4)$$

Similarly, the kinetic effect of pressure is quantified through the activation volume  $\Delta V^\ddagger$ , defined as the logarithmic pressure derivative of the rate constant  $k^6$  (eq. 1.5).

$$\Delta V^\ddagger = -RT \left(\frac{\partial \ln k}{\partial P}\right)_T \quad (1.5)$$

While the relationship in eq. 1.5 was known as early as the late 19<sup>th</sup> century,<sup>5,7</sup> at that time there did not exist a theoretical framework within which to provide a physical interpretation of  $\Delta V^\ddagger$ . It wasn't until 1935, with the advent of transition state (TS) theory by Eyring,<sup>8,9</sup> Evans and Polanyi,<sup>10</sup> that the mechanistic utility of  $\Delta V^\ddagger$  became truly apparent. Within TS theory,  $\Delta V^\ddagger$  is interpreted as the difference in partial molar volume between the TS and the reactant state (eq. 1.6).

$$\Delta V^\ddagger = V^\ddagger - \sum V_R \quad (1.6)$$

Since the partial molar volume of reactant states are easily obtainable experimentally,  $\Delta V^\ddagger$  can be used as a direct measurement of the volume of the TS. Therefore, volume is one of the few structural properties of TSs that can be assessed from direct experimental measurements.

### 1.1.2. Mechanistic applications for chemical systems

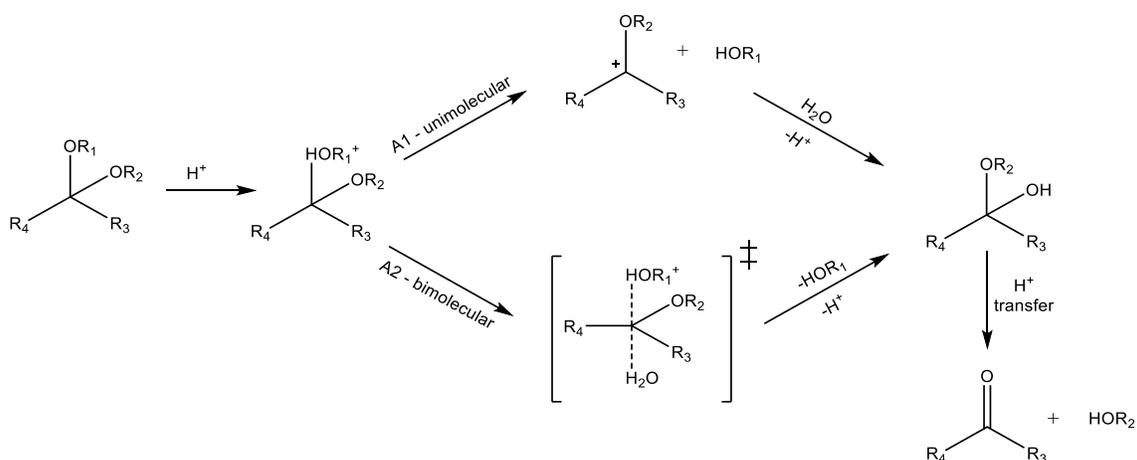
Activation and reaction volumes have been tabulated for a wide variety of chemical processes for the past seventy years,<sup>3,4,11-14</sup> and it appears that different types of reactions fall into typical ranges of  $\Delta V^\ddagger$  values (Table 1.1). While the large degree of overlap in the  $\Delta V^\ddagger$  ranges limits their usefulness in the context of mechanistic discrimination, there have been instances in the literature where the sign and relative magnitude of the  $\Delta V$  and  $\Delta V^\ddagger$  values have provided useful mechanistic insight.<sup>2</sup> Two examples of this kind are discussed below.

**Table 1.1** Typical ranges of  $\Delta V^\ddagger$  for different types of reactions. Data from ref. 14.

Type of Reaction	$\Delta V^\ddagger$ (cm <sup>3</sup> /mol)
Homolysis	5 to 20
Radical polymerization	≈ - 20
Diels-Alder cyclization	-25 to -40
Intramolecular cyclizations	-25 to -30
Dipolar cycloadditions	-40 to -50
Ester hydrolysis (basic)	-10 to - 15
Ester hydrolysis (acidic)	> - 10
Epoxide ring-opening	-15 to - 20
Wittig reactions	-20 to -30

In 1892, Röntgen performed the first high pressure kinetic study by investigating the hydrochloric acid-catalyzed hydrolysis of sucrose.<sup>3,15,16</sup> His objective was to determine if the previously observed increase in the conductivity of HCl solutions with pressure was due to an increase in dissociation, which would cause an increase in the rate of hydrolysis, or an increase in ion mobility in solution, which would have no effect on the rate. Surprisingly, he found that pressure slowed the hydrolysis reaction. This result was reproduced by other researchers.<sup>3,17</sup> Even more puzzlingly, a subsequent study of the effects of pressure on the HCl-catalyzed hydrolysis of methyl and ethyl acetate showed

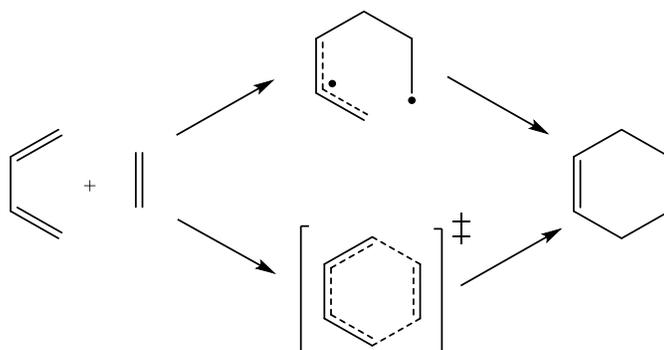
that these reactions were accelerated by pressure.<sup>3,16,18</sup> These observations were left unexplained for nearly sixty years, until Hamann<sup>1</sup> interpreted them as evidence that the hydrolysis of the ketal linkage in sucrose occurs via an A1 mechanism, while the ester hydrolysis occurs via an A2 mechanism (Figure 1.1). In both mechanisms the rate limiting step is the departure of the leaving group following protonation, but in the A1 mechanism this occurs in a stepwise fashion with a carbocation intermediate while in the A2 mechanism it is concerted. The different pressure effect on the rate of hydrolysis is simply a reflection of a difference in reaction mechanism and the sign of  $\Delta V^\ddagger$  has been held up as a classic indicator of an A1 versus an A2 mechanism<sup>2,11,19,20</sup>



**Figure 1.1** The A1 (top) and A2 (bottom) mechanisms of acetal hydrolysis. The A1 mechanism is stepwise and occurs via a carbocation intermediate, while the A2 mechanism is concerted.

Perhaps the most well-known example of the mechanistic application of high pressure is the discrimination between the concerted and stepwise pathways for Diels-Alder reactions.<sup>21,22</sup> This reaction, discovered in the late 1920s,<sup>23</sup> is a [4+2] cycloaddition reaction between a diene and an alkene, commonly referred to as a dienophile. The Diels-Alder reaction has become a very useful tool in synthetic organic chemistry since it forms 6-membered rings with predictable regio- and stereochemistry.<sup>20</sup> Two competing mechanisms for the Diels-Alder reaction have been proposed, shown in Figure 1.2. The first is known as the concerted mechanism,<sup>24</sup> where the two new bonds are partially formed in the transition state. There is no intermediate structure in this mechanism as both bonds are formed concurrently. The second is known as the stepwise or diradical mechanism,<sup>25</sup> in which the bonds form one after the other and an intermediate structure exists with one bond between the diene and dienophile fully formed. These two

mechanisms were hotly debated for almost 60 years,<sup>22</sup> with evidence piling up on both sides. High pressure studies of Stewart,<sup>26</sup> and Klärner<sup>27</sup> showed that the reaction volume  $\Delta V$  and activation volume  $\Delta V^\ddagger$  were similar in both sign and magnitude for a wide variety of Diels-Alder reactions. For this to occur, the transition state must have a similar volume as the cycloadduct product. This clearly favours the concerted mechanism over the stepwise mechanism, as the transition state is more product-like with partial formation of both bonds.



**Figure 1.2** The stepwise di-radical (top) and concerted mechanisms (bottom) of the Diels-Alder reaction between butadiene and ethene.

### 1.1.3. Application of pressure to biological systems

Proteins are naturally-occurring polymers that perform a wide variety of functions essential for life. The building blocks of proteins are amino acids, which differ by the identity of their R group. The R groups of amino acids vary in shape, size, polarity and charge, and as such the properties of a protein will depend on its amino acid composition. One of the most important properties of a protein is its three-dimensional structure. When a protein is synthesized, it will fold into a structure that is uniquely determined by its amino acid sequence<sup>28</sup> This structure is intimately linked with the protein's function, and a change in structure is generally accompanied by a loss of function or gain of a new function; both of which can lead to disease.<sup>28</sup> As such, it is extremely important that the mechanism by which this folding process occurs is well understood.

The earliest use of elevated pressure on a biological system was by Bridgman in 1914.<sup>29</sup> He applied hydrostatic pressure to egg white, and noted that at a pressure of 7 kbar it coagulated as if it had been cooked. In the years that followed, studies were carried out on other proteins, such as trypsin, chymotrypsin, pepsin and chymosin.<sup>30–32</sup> In all

cases, the application of pressure resulted in either inactivation of enzymatic activity or aggregation of the protein, indicating that the proteins were unfolding and that the volume change for unfolding was negative. In the hundred years since those early studies, reaction volumes have been tabulated for the unfolding of a wide variety of proteins.<sup>32</sup>

The negative reaction volume for protein unfolding tells us that the unfolded state has a smaller partial molar volume than the folded state. This is due to two distinct contributions: the effect of internal cavities and the effect of solvation.<sup>33–37</sup> The intricate three-dimensional structures of proteins allows for void volumes, also known as cavities, to exist in the core of the protein due to empty space between the amino acid side chains. Upon unfolding, these cavities cease to exist, decreasing the volume of the protein. The second effect is a result of water-amino acid interaction. As the protein unfolds, the hydrophobic amino acids at its core become exposed to water. The water molecules must solvate the newly exposed amino acids, resulting in a net increase in water density near the protein. This would result in a shrinkage in volume of the protein-water system. Recently, it has been suggested that the cavity effect is dominant over the solvation effect.<sup>34,35,37</sup>

The development of the fast pressure jump and high pressure fluorescence technique in the 1970s<sup>38,39</sup> allowed for measurement of high pressure rate constants for protein unfolding, and consequently activation volumes. These values are typically positive, but vary widely between different proteins.<sup>40–52</sup> Conclusions have been drawn about the relative size and hydration of the transition state based on the sign and magnitude of the activation volume,<sup>40</sup> but further analysis is needed to fully understand these results.

#### **1.1.4. Pressure-induced viscosity effects**

In 1940, the then-newly developed TS theory<sup>8–10</sup> was challenged by Kramers,<sup>53</sup> who showed that an increase in solvent viscosity could significantly affect the reaction dynamics and that the resultant reaction rates would be lower than those predicted by TS theory. If the solvent motion is slower than the chemical reaction being studied, the thermodynamic equilibrium between the reactant and transition state is disturbed and the reaction rate becomes limited by the solvent motion. In the years since, Kramers' original work has been improved upon by many authors,<sup>54–61</sup> leading to a greater theoretical

understanding of the effect of solvent dynamics on chemical reactions. However, it wasn't until nearly 50 years after Kramers that experimental confirmation of these solvent effects was obtained,<sup>62-71</sup> through the use of high pressure as a way to increase the solvent viscosity.

Some of the most extensive work on solvent viscosity in high pressure reactions was performed by Asano and coworkers,<sup>65-69</sup> who studied high pressure E/Z isomerizations<sup>66</sup> and cyclizations<sup>68</sup> in viscous solvents and their non-viscous analogues. They showed solvent viscosity to be an important factor influencing reaction kinetics at elevated pressures, thus providing an unequivocal experimental confirmation for Kramers' ideas. Their results also indicate that at high pressures the reaction kinetics may no longer be accurately described by transition state theory. This is an important consideration for all high pressure reactions and will be expanded upon in Chapter 2 of this work.

## **1.2. Theoretical calculations of molecular volumes, volumes of activation and volume profiles**

### **1.2.1. Review of existing volume models**

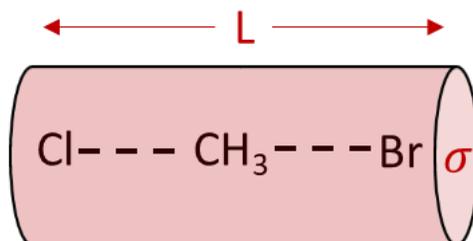
As detailed in section 1.1.2, the theoretical interpretation of  $\Delta V^\ddagger$  has thus far been limited to a qualitative discussion based on the sign and magnitude of their values. However, extraction of more detailed structural information about the TS from the experimentally-determined  $\Delta V^\ddagger$  is not possible without an accurate theoretical method for calculating volumes of molecular species. In solution, contributions to the molecular volume come not only from the volume occupied by the constituent atoms, but also from temperature effects (expansion volume), and the effects of packing (void volume) and solvent-solute interactions (solvation volume). For the past seventy years, various models of molecular volume have been proposed in order to address this complex problem. In this section, the existing theoretical models of molecular volume will be reviewed.

#### **1.2.1.1 Stearn-Eyring model**

In 1941, Stearn and Eyring proposed the first model for molecular volume.<sup>72</sup> In this model, the reaction system is approximated with a cylinder of constant cross-sectional area  $\sigma$  and variable length  $L$ , as illustrated in Figure 1.3. The volume of the molecular system can therefore be calculated using eq. 1.7.

$$V = \sigma L \quad (1.7)$$

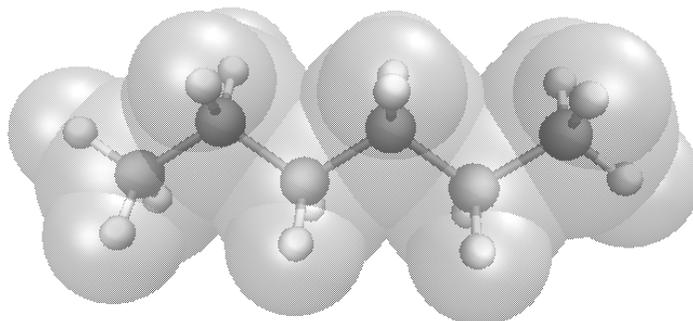
This model, while useful for qualitative discussions of simple bond breaking and forming reactions, is not broadly useful since it does not take into account the molecular shape or any change in structure other than elongation along  $L$ .



**Figure 1.3** The Stearn-Eyring cylindrical model of molecular volume, shown here for the example of an  $S_N2$  reaction.

#### 1.2.1.2 Van der Waals volume

The hard sphere, or van der Waals, volume model offers an alternative to the structurally insensitive Stearn-Eyring model.<sup>73-76</sup> In this model, the volume of a molecule is calculated to be the total volume encompassed by overlapping van der Waals spheres centered on each atom in the molecule, as illustrated in Figure 1.4. While this model was an improvement due to the stronger link between molecular structure and molecular volume, the choice of van der Waals radii used was quite arbitrary. In addition, the model completely neglects the effect of solvent-solute interaction as well as contributions to the volume from both thermal expansion and packing, which is particularly bad for macromolecules.



**Figure 1.4** The van der Waals molecular volume model, shown for the example of  $n$ -hexane. The volume is calculated from the set of overlapping van der Waal spheres centered on each atom.

### 1.2.1.3 Solvent accessible and solvent excluded volumes

The van der Waals model does not consider the effect of the size of the solvent particles on the volume of the solute, which is an important contribution to molecular volume. In the 1970s, Lee and Richards proposed two new models for molecular volume that include a consideration of the solvent: the solvent accessible<sup>77,78</sup> and solvent excluded volumes.<sup>79–82</sup> In both models, a probe sphere with a radius chosen to represent the solvent size is rolled over the van der Waals surface of the molecule in question. The solvent accessible volume is defined as the volume encapsulated by the surface traced by the center of the probe sphere. The solvent excluded volume, on the other hand, is defined as the volume encapsulated by the surface traced by the edge of the probe sphere. From the schematic in Figure 1.5, one can see that, for a given probe radius, the solvent accessible volume is greater than the solvent excluded volume and likely provides an overestimation of the molecular volume.



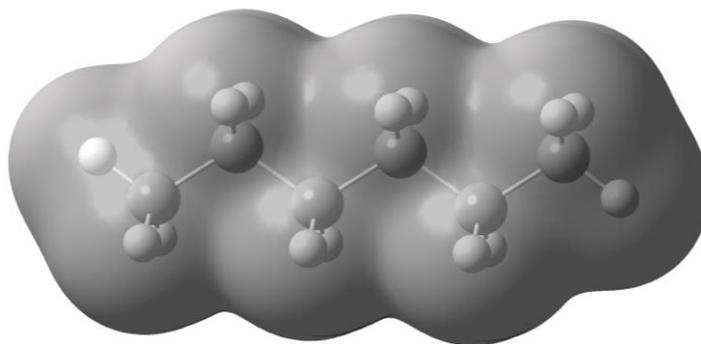
**Figure 1.5** Solvent accessible (left) and solvent excluded (right) models of molecular volume. The difference between these two models lies in whether the molecular surface is defined by the center of the probe sphere, or the edge.

Both of these volumes critically depend on the choice of the probe sphere and van der Waals radii, a choice which remains rather arbitrary. These models also do not take into account the effects of temperature. A temperature-dependent probe sphere radius may be used to account for this,<sup>83</sup> but with no *a priori* method for determining the correct radii this technique has limited applicability for calculation of molecular volumes.

### 1.2.1.4 Quantum mechanical isodensity model

A more sophisticated alternative to the geometrical hard sphere models, proposed by Bader in 1967,<sup>84,85</sup> uses the electron density to define the boundaries of the molecule as the quantum mechanics-based electronic isodensity surface. The volume of the

molecule is then calculated as the interior of the surface as shown in Figure 1.6. The isodensity level is typically chosen to be 0.001 a.u., but this parameter is not universal and can vary between different molecule types.<sup>86–90</sup> A comparison of molecular volumes calculated using the isodensity model with their experimental values showed that the model underestimates the volume by up to 25%.<sup>89</sup> While the isodensity model is certainly more advanced than the hard sphere models, it suffers from the same drawback: a lack of solvent and temperature dependency. A comparison of Figure 1.4 and Figure 1.6 clearly shows the visual similarity between the isodensity and the hard sphere volume models.



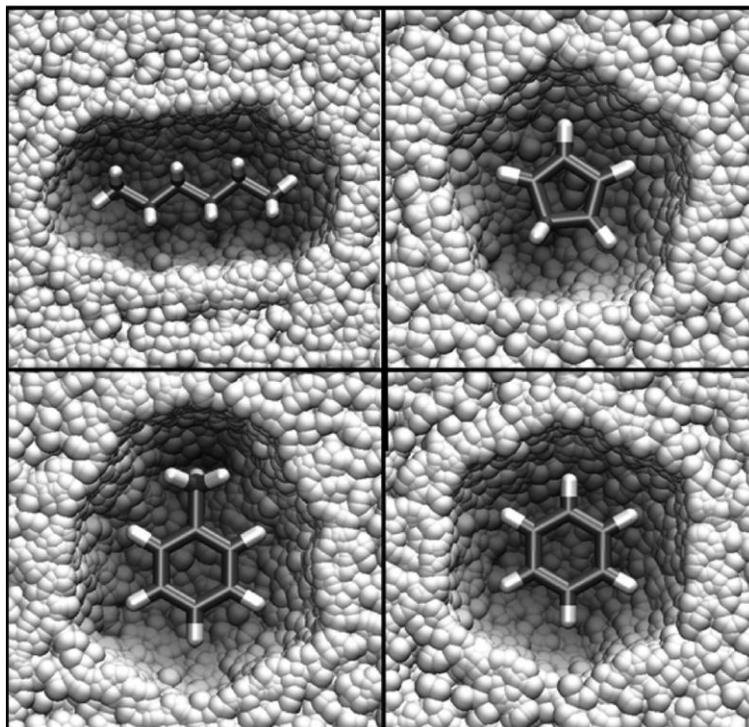
**Figure 1.6** Isodensity surface for *n*-hexane, with isodensity level = 0.001 a.u.

### 1.2.2. Archimedean displacement model

All of the aforementioned models suffer from the same fundamental problem: molecules are treated as geometric objects with distinct boundaries defined by model-specific parameters, and thus the effects of temperature and solvent-solute interactions are ignored. Interaction-based molecular dynamics (MD) and Monte Carlo simulations offer a physically more realistic alternative for calculating molecular volumes.<sup>27,34,91–109</sup>

Recently, a new model has been developed in our lab that utilizes MD simulations to accurately calculate molecular volumes.<sup>83,110,111</sup> When a solute molecule *X* is inserted into *N* molecules of solvent *S*, it causes displacement of the latter as illustrated by Figure 1.7. This displacement, combined with the solute-induced restructuring of the solvent, leads to the change in the overall volume of the solvent-solute system relative to the volume previously occupied by the pure solvent. The resulting incremental volume change represents the partial molar volume of the solute and can be used as a measure of the solute volume in a given solvent. Practically, this volume can be calculated by subtracting the volume of pure solvent from the volume of the solvent plus the target solute (eq. 1.8).

$$V(X) = V(N \cdot S + X) - V(N \cdot S) \quad (1.8)$$



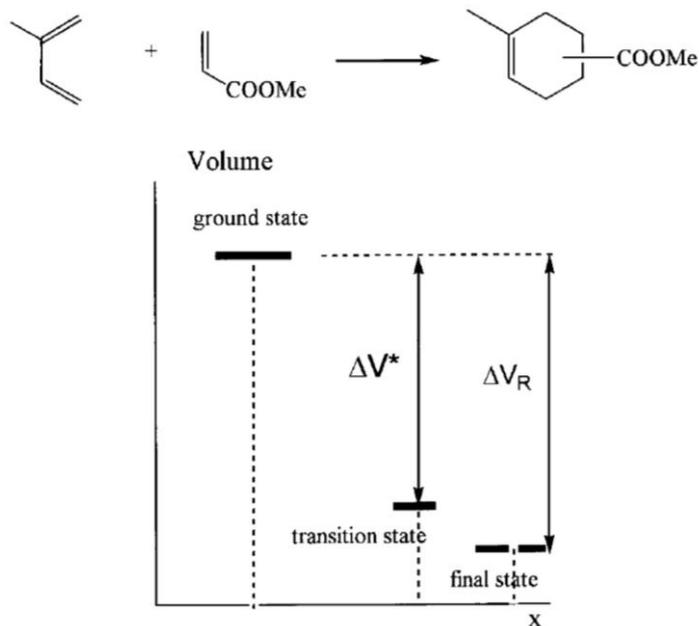
**Figure 1.7** Four hydrocarbons (clockwise from top left corner: *n*-hexane, cyclopentadiene, benzene and toluene) immersed in a model solvent represented here by an overlay of solvent configurations at different instants of time. The solvent particles (white) avoid the solute, forming a cavity which matches the size and shape of the solute molecule.

Molecular volume, so defined, is sensitive to temperature, pressure, the nature and strength of solvent-solute interactions, and most importantly, the geometrical configuration of the solute. This method, referred to as the Archimedean displacement model, has been shown to produce accurate results for molecular, reaction and activation volumes for a large variety of different molecules and reactions in solvents of different polarity and at a variety of temperatures.<sup>83,110,111</sup>

### 1.2.3. Volume profiles

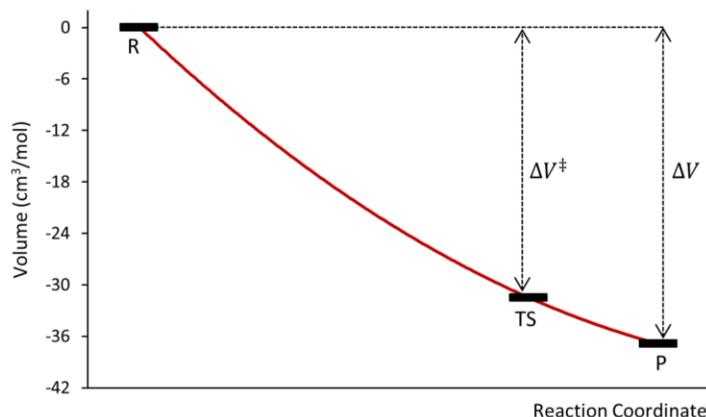
When activation and reaction volumes are used to discuss reaction mechanisms, the concept of a volume profile is often invoked.<sup>112–116</sup> A volume profile describes how the partial molar volume of a reaction system varies along its reaction coordinate, and serves

as a visualization of the volume changes that occur during a reaction. Since the only transient state that has an experimentally measurable volume is the TS, experimentally-determined volume profiles are limited to only 3 points as can be seen in Figure 1.8, and thus provide only limited information about the volume changes along the reaction coordinate.



**Figure 1.8** Experimentally-determined 3-point volume profile for the Diels-Alder reaction between isoprene and methyl acrylate. Figure from ref. 117, Copyright © 2002 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

Similarity in volume between the TS and either the reactant or product state has been used in the past to discriminate between proposed reaction mechanisms<sup>22,26,27</sup> as discussed in section 1.1.2, but detailed determination of the structural characteristics of the TS was not possible. With the development of the Archimedean displacement model, this is no longer the case. By calculating the volume of a set of molecular structures taken sequentially along the reaction coordinate, it is now possible to fill in the missing points on the volume profile to generate a smooth curve as shown in Figure 1.9. These volume profiles are valuable mechanistic tools not only because they allow for a more detailed picture of how the size of a system changes as the reaction progresses, but also because they offer a method for constructing high pressure reaction profiles, as discussed in the next section, and because they can be used to assess the structure of an unknown TS, detailed in Chapter 4.



**Figure 1.9** Schematic continuous volume profile (red line) connecting the experimentally-determined  $\Delta V^\ddagger$  and  $\Delta V$  along the reaction coordinate. Experimental data is the same as in Figure 1.8.

### 1.2.4. High pressure reaction profiles

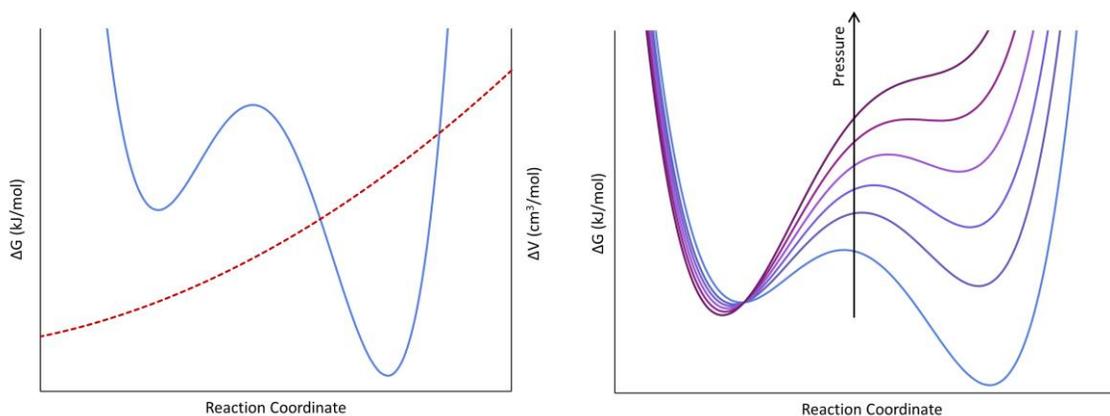
An important application of volume profiles is their use in the calculation of high pressure Gibbs energy profiles. When pressure is applied to a reaction system, it modifies the Gibbs energy landscape, resulting in an alteration of reactant, product and TS structures and a change in the reaction barriers. This effect is described by eq. 1.9, which is obtained by integration of eq. 1.2. Here  $\Delta G(x; P)$  is the elevated pressure Gibbs energy profile along reaction coordinate  $x$ ,  $\Delta G_o(x)$  is the Gibbs energy profile at a reference pressure, and  $\Delta V(x; P)$  is the pressure-dependent volume profile for the process that can be calculated using our Archimedean displacement model.

$$\Delta G(x; P) = \Delta G_o(x) + \int_{P_o}^P \Delta V(x; P) dP \quad (1.9)$$

Although the volume of a molecular system displays a significant dependence on pressure, this dependence is different for different degrees of freedom. The reaction coordinate is typically the softest degree of freedom, and if we assume that the other degrees of freedom do not contribute significantly to the volume change, *i.e.* that the volume profile is independent of pressure, eq. 1.9 can be modified to produce eq. 1.10,

$$\Delta G(x; P) \approx \Delta G_o(x) + (P - P_o)\Delta V_o(x) \quad (1.10)$$

where  $\Delta V_o(x)$  is the reference pressure volume profile. This assumption has been tested by its application to collinear exchange reactions in a number of triatomic systems,<sup>118</sup> and was found to work well.\*



**Figure 1.10** An illustration of the use of eq. 1.10 for a hypothetical two-state system. Left: the reference-pressure  $\Delta G_o(x)$  (blue) and  $\Delta V_o(x)$  (red dashed) profiles for the system. Right: a series of high pressure  $\Delta G(x; P)$  profiles showing a change in the relative stabilities of the reactant and product state, as well as the pressure dependence of the reaction barriers. At the highest pressures, the less compact state ceases to exist.

The use of eq. 1.10 is illustrated in Figure 1.10 for the case of a hypothetical two-state system where the state that is more stable at ambient conditions has a higher volume. As pressure increases, the relative Gibbs energy of the higher volume state begins to increase due to the larger contribution of the  $(P - P_o)\Delta V_o(x)$  term. At the highest pressures, the barrier separating the two states completely disappears and the system becomes trapped in the more compact state.

We used this approach to investigate pressure-dependent Gibbs energy landscapes in reactive and conformationally flexible systems exemplified by a study of hydrogen transfer at high pressures (Chapter 2) and pressure-induced conformational changes in molecular machines (Chapter 3).

---

\* Note that this may not be the case for more conformationally rich systems like polypeptides and proteins.

## 1.3. Computational methodologies

In this work, we utilized two of the most common methodologies in computational chemistry: classical molecular dynamics simulations and quantum mechanical calculations. A brief overview of these two methodologies is presented here.

### 1.3.1. Molecular dynamics simulations

Classical MD simulations are the technique of choice for systems where dynamic information is required, solvent-solute interactions are important or where finite temperature and pressure are needed. The bulk of the computation performed in this work was done using MD simulations, from which volumes, Gibbs energies, potential energies, enthalpies and other properties of chemical systems were obtained.

#### 1.3.1.1. *The molecular dynamics algorithm*

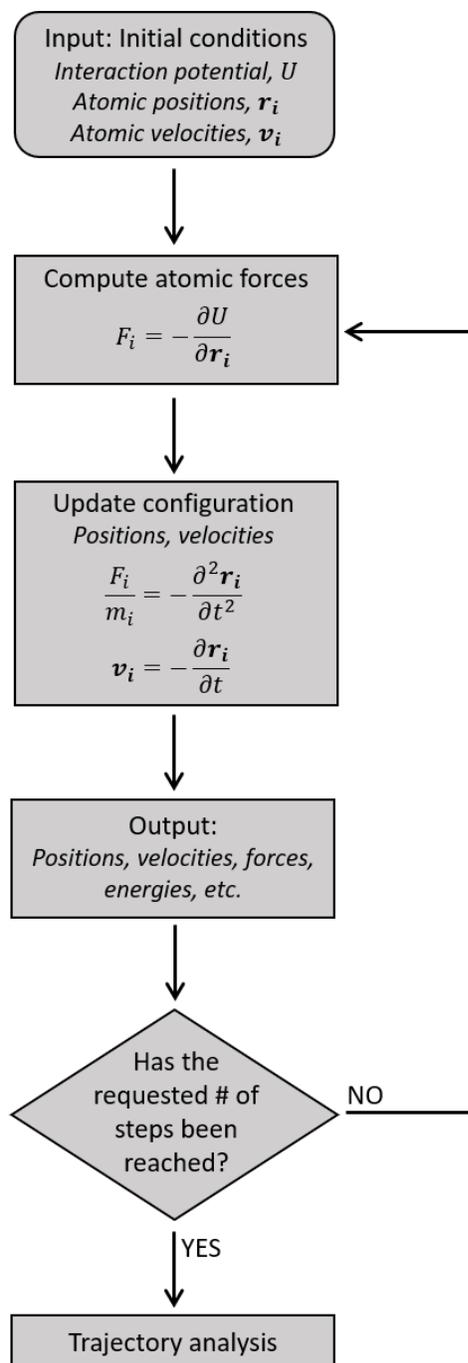
In MD simulations, the motion of each atom  $i$  in the system is described by Newton's equation (eq. 1.11), where  $t$  is time,  $m_i$  is the atomic mass, and  $\mathbf{r}_i$  and  $\mathbf{F}_i$  are the atomic position and force vectors, respectively.

$$m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = \mathbf{F}_i \quad (1.11)$$

The force on each atom  $\mathbf{F}_i$  is obtained as the negative gradient of potential energy  $U$  (eq. 1.12):

$$-\frac{\partial U}{\partial \mathbf{r}_i} = \mathbf{F}_i \quad (1.12)$$

The time dependence of the atomic positions  $\mathbf{r}_i$  is obtained through numerical integration of eq. 1.11 & 1.12 using small time steps, typically of the order of 1 fs. We used the leap-frog algorithm for time integration.<sup>119</sup> The general workflow of an MD simulation is outlined in Figure 1.11. In this work, the GROMACS<sup>120–122</sup> software package was used for all MD simulations.



**Figure 1.11** The workflow for an MD simulation in GROMACS.

### 1.3.1.2. Interaction potentials

The potential energy function  $U$  used in classical MD simulations includes contributions of two types: non-bonded interactions such as van der Waals and electrostatics, and bonded interactions such as bond stretching, angle bending and torsions (eq. 1.13). The set of parameters for  $U$  is known as a force field.

$$\begin{aligned}
U = & \sum_{i>j}^N U_{vdW}(r_{ij}) + \sum_{i>j}^N U_{es}(r_{ij}) + \sum_{bonds} U_{bond}(r) \\
& + \sum_{angles} U_{angle}(\theta) + \sum_{torsions} U_{torsion}(\phi)
\end{aligned}
\tag{1.13}$$

The van der Waals interaction  $U_{vdW}(r_{ij})$  between two atoms  $i$  and  $j$  is modeled with the Lennard Jones potential<sup>123</sup> (eq. 1.14).

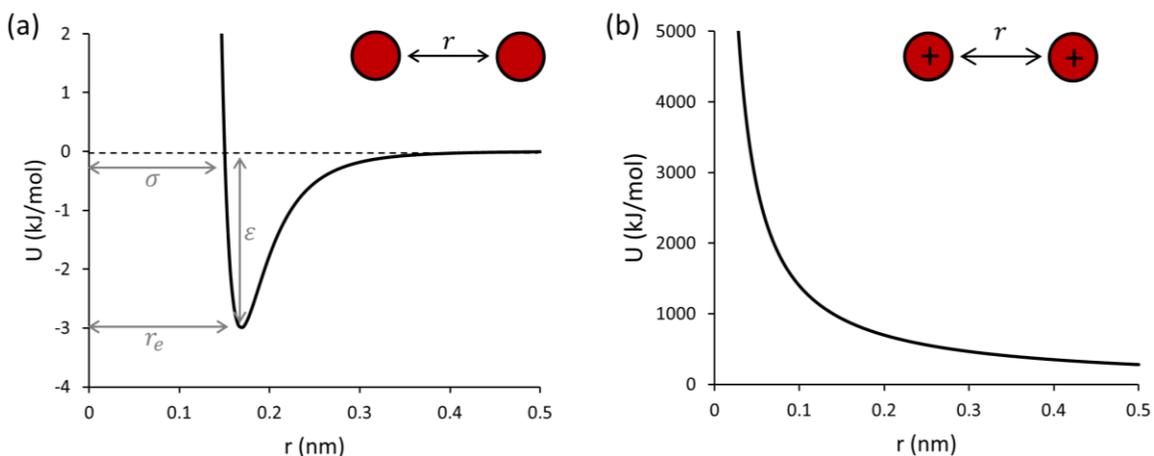
$$U_{vdW}(r_{ij}) = 4\varepsilon_{ij} \left( \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right)
\tag{1.14}$$

A sample Lennard Jones potential is shown in Figure 1.12(a). The twelfth power term represents repulsion between the two atoms at close distances, and the negative sixth power term represents the long-range attraction. The  $\sigma_{ij}$  parameter is the distance at which the potential is zero. It is related to the equilibrium distance  $r_e$  between the two atoms by eq. 1.15. The  $\varepsilon_{ij}$  parameter is the depth of the attractive well.

$$r_e = 2^{1/6} \sigma_{ij}
\tag{1.15}$$

Parameters  $\varepsilon_{ij}$  and  $\sigma_{ij}$  for atoms of different types are obtained as a geometric mean of the individual  $\varepsilon$  and  $\sigma$  values of the atoms involved (eq. 1.16).

$$\sigma_{ij} = \sqrt{\sigma_i \sigma_j}, \varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}
\tag{1.16}$$



**Figure 1.12 (a) The Lennard-Jones potential with  $\sigma$ ,  $\epsilon$  and  $r_e$  labelled. (b) The Coulomb potential, shown for the repulsive interaction between two positively charged particles.**

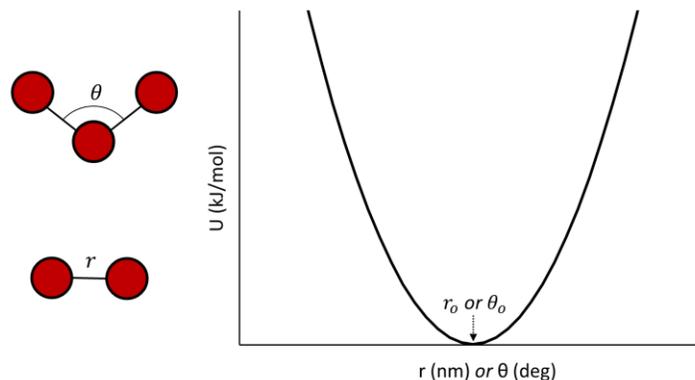
The electrostatic interaction  $U_{es}(r_{ij})$  between two charged atoms is modeled by the Coulomb potential (eq. 1.17)

$$U_{es}(r_{ij}) = f \frac{Q_i Q_j}{r_{ij}} \quad (1.17)$$

where  $Q_i$  and  $Q_j$  are the charges on each atom, and  $f = \frac{1}{4\pi\epsilon_0} = 138.94 \text{ kJ mol}^{-1} \text{ e}^{-2}$  is a conversion factor. A sample Coulomb potential is shown in Figure 1.12(b). Note that both van der Waals and electrostatic interactions are limited to interactions between atoms of different molecules, or between atoms of the same molecule that are at least 3 atoms away from each other.

The harmonic potential  $U_{bond}(r)$  (eq. 1.18) is most commonly used to describe bond stretching energy. It is illustrated in Figure 1.13.

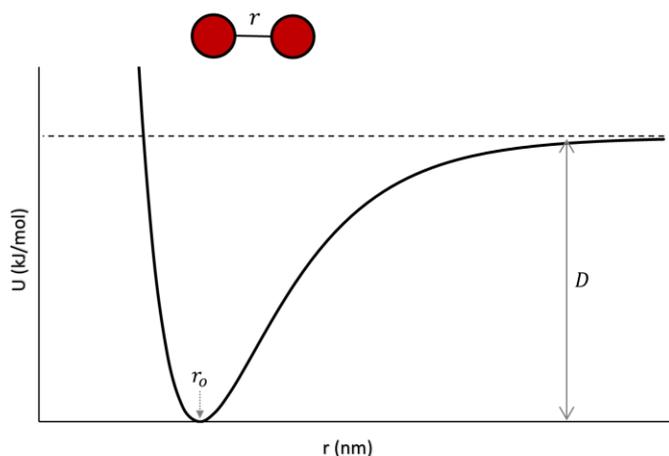
$$U_{bond}(r) = \frac{1}{2} k_b (r - r_o)^2 \quad (1.18)$$



**Figure 1.13** The harmonic potential, which is used to describe both bond stretching and angle bending interactions. Note that the x-axis can be either  $r$  or  $\theta$ , depending on the interaction being modeled. The equilibrium bond or angle value is located at the bottom of the parabola.

In this potential,  $k_b$  is the bond stretching force constant and  $r_0$  is the equilibrium bond length. This potential works well for small-amplitude vibrations but is a poor approximation to highly stretched or highly compressed bonds (*i.e.* the anharmonic region), in which case more sophisticated models, such as the Morse potential  $U_{morse}(r)$  (eq. 1.19, Figure 1.14), can be used. For this potential,  $D$  and  $\beta$  control the depth and width of the attractive well respectively, and  $r_0$  is the equilibrium bond length.

$$U_{morse}(r) = D(1 - \exp[-\beta(r - r_0)])^2 \quad (1.19)$$



**Figure 1.14** The Morse potential, used to describe bond stretching for systems that cannot be accurately represented by the harmonic potential of Figure 1.13.

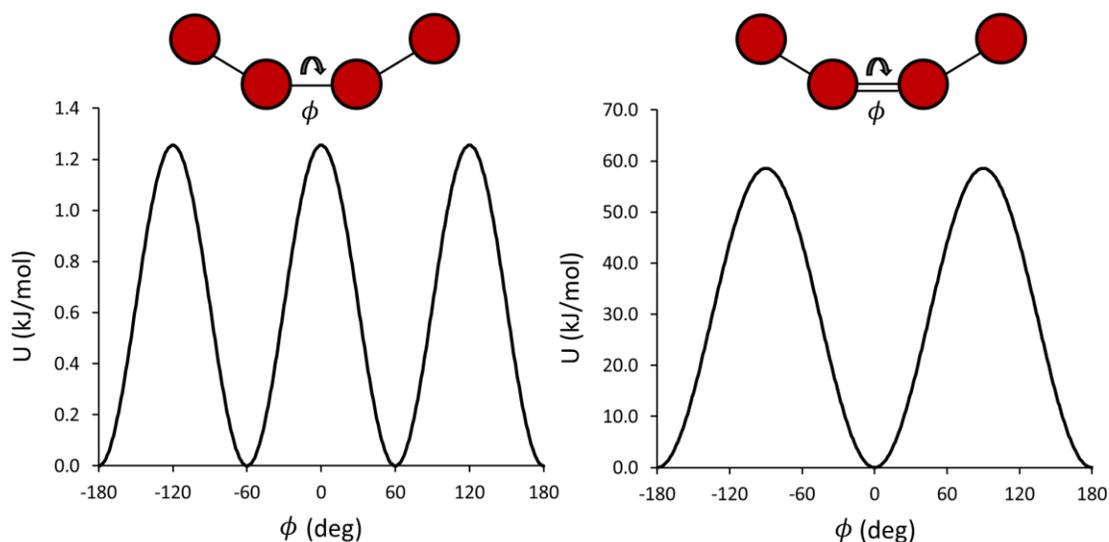
The angle bending interaction  $U_{angle}(\theta)$  is given a similar treatment to the bond stretching interaction and is represented by a harmonic potential (eq. 1.20, Figure 1.13) with force constant  $k_a$  and equilibrium angle  $\theta_o$ .

$$U_{angle}(\theta) = \frac{1}{2}k_a(\theta - \theta_o)^2 \quad (1.20)$$

Dihedral angle interactions are periodic in nature, and thus cannot be modeled with a harmonic potential. Instead, the Ryckaert-Bellemans potential<sup>124</sup> (eq. 1.21) is frequently used.

$$V_{torsion}(\phi) = \sum_{n=0}^5 C_n \cos(\phi)^n \quad (1.21)$$

By different choice of  $C_n$  the shape of the torsion profile can be adjusted to what is appropriate for the identity of the atoms involved in a specific dihedral (Figure 1.15).



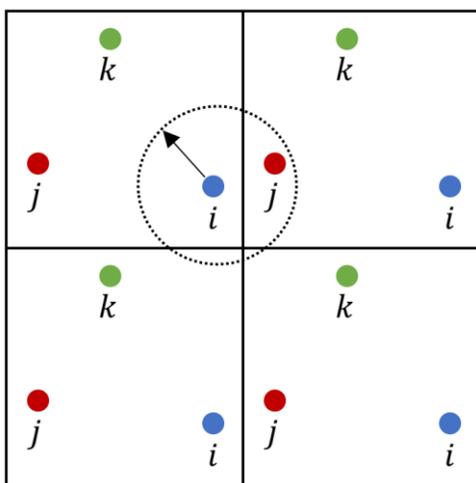
**Figure 1.15** Sample Ryckaert-Belleman potentials for rotation around a single bond (three-fold potential on the left) and a double bond (two-fold potential on the right).

The force field parameters for these different potentials are optimized to reproduce experimental data such as density and heat of vaporization,<sup>125,126</sup> or to fit to gas phase *ab*

*initio* potential energies.<sup>127</sup> For our simulations, we used the Optimized Potential for Liquid Simulation (OPLS) force field<sup>125,126</sup> unless otherwise stated.

### 1.3.1.3. Periodic boundary conditions

To improve the computational efficiency of MD simulations, periodic boundary conditions are typically employed.<sup>119</sup> This technique allows for a large system to be described by a relatively small number of particles. The unit cell, referred to as the simulation box, is surrounded by translated copies of itself to yield an essentially infinite system. Unfortunately, long-range interactions such as electrostatics pose a problem, as self-interaction through the periodic boundary would result in serious artifacts. Therefore, all long-range interactions are terminated beyond a user-defined cut-off radius, as illustrated in Figure 1.16. The chosen cut-off length must be long enough to encapsulate all interactions of interest, but short enough that a particle cannot interact with its own periodic image. Here, a cut-off length of 0.9 nm was used for all simulations.



**Figure 1.16** The use of periodic boundary conditions allows particle  $i$  to interact with a translated image of particle  $j$ . In order to eliminate self-interaction, a cut-off radius is imposed for all through-space interactions (dashed circle).

### 1.3.1.4. Temperature and pressure control

In this work, simulations were performed in either the canonical  $NVT$  ensemble (constant number of particles, volume, and temperature) or the isobaric-isothermal  $NPT$  ensemble (constant number of particles, pressure, and temperature). In MD simulations, the temperature  $T$  of an  $N$  particle system is calculated from the average kinetic energy

using eq. 1.22, where  $k_B$  is the Boltzmann constant and  $N_{df}$  is the number of degrees of freedom in the system.

$$T = \frac{\sum_{i=1}^N m_i v_i^2}{N_{df} k_B} \quad (1.22)$$

The hydrostatic (isotropic) pressure  $P$  is calculated according to the virial theorem,<sup>119</sup> and consists of a contribution from the kinetic energy and from the forces between atoms in the system (eq. 1.23).

$$P = \frac{2}{3V} \left[ E_{kin} + \frac{1}{2} \sum_{i<j}^N \mathbf{r}_{ij} \mathbf{F}_{ij} \right] \quad (1.23)$$

To maintain a constant temperature or pressure, the system is coupled to an external bath referred to as a thermostat or a barostat. We chose to use the Berendsen algorithm<sup>128</sup> for both temperature and pressure coupling. The deviation of pressure or temperature from the desired values is corrected by either rescaling the velocities of all atoms, or by rescaling the volume of the simulation box, respectively. This is controlled according to eq. 1.24, where  $X$  is either temperature or pressure,  $X_0$  is the desired value and  $\tau$  is the coupling constant, which determines how often the quantity is updated.

$$\frac{\partial X}{dt} = \frac{X_0 - X}{\tau} \quad (1.24)$$

A smaller  $\tau$  value indicates a stronger coupling, as the pressure or temperature is updated more frequently. Unless otherwise stated, values of  $\tau_T = 0.1$  ps and  $\tau_P = 1.0$  ps were used for temperature and pressure control, respectively.

### 1.3.1.5. Thermodynamic integration

Calculation of Gibbs energy  $\Delta G$  was accomplished using thermodynamic integration.<sup>119</sup> For this, the Gibbs energy difference  $\Delta G_{A \rightarrow B}$  between two states  $A$  and  $B$  is assessed by slowly tuning the Hamiltonian  $\mathcal{H}$  of the system from that describing state  $A$ , to that describing state  $B$ . Practically, this involves the introduction of a coupling parameter

$\lambda$  such that  $\mathcal{H}(\lambda = 0) = \mathcal{H}_A$  and  $\mathcal{H}(\lambda = 1) = \mathcal{H}_B$ . A series of simulations are run for systems with  $\lambda$  values ranging from 0 to 1, and the average derivative with respect to  $\lambda$  is obtained for each simulation. The Gibbs energy difference  $\Delta G_{A \rightarrow B}$  can then be assessed from these partial derivatives via numerical integration, according to eq. 1.25.

$$\Delta G_{A \rightarrow B} = \int_0^1 \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (1.25)$$

### 1.3.2. Quantum mechanical calculations

In this work, QM calculations were used to obtain gas-phase potential energy profiles and surfaces, to optimize TS structures and also to generate force field parameters for chemical species that are not parameterized in the standard OPLS parameter set. All calculations were performed using Gaussian 09<sup>129</sup> at the B3LYP/6-31++G(d,p) level, unless otherwise specified. All TS structures obtained using QM calculations were verified using frequency calculations.

## 1.4. Motivation and objectives

The concept of a reaction profile, describing the change of potential energy, free energy, or Gibbs energy along the reaction coordinate, has been widely used in chemical kinetics for the discussion of chemical and biochemical reaction mechanisms. However, the closely related concept of a volume profile has not received the same level of recognition, mostly due to the lack of a theoretical methodology that could effectively utilize this concept. The goal of this work is to fill this void by developing the necessary computational protocols and thus to pave the way towards interpreting the large amount of experimental high pressure kinetic data in the literature<sup>11–13,40–52</sup> which has, so far, remained largely unutilized at a mechanistic level.

Two novel applications of volume profiles are explored in this work. The first is their use for the calculation of elevated pressure Gibbs energy profiles, as introduced in section 1.2.4. This enables one to predict the structural changes that occur in molecular systems with an increase in pressure. Such ability to predict these changes enables one to design systems with pressure-sensitive properties and to assess the range of pressures

necessary to achieve the desired effect. The second application of volume profiles formulated in this work is their use, in combination with experimentally determined activation volumes, for the elucidation of unknown TSs or TS ensembles in flexible systems. Although this methodology is ultimately targeted at large biochemical systems, such as the folding of proteins, here we focus primarily on smaller systems where the accuracy of our predictions can be properly verified.

## 1.5. Brief outline of this work

In this work, volume profiles were used to predict and interpret the effects of pressure on different types of reactions, ranging from radical hydrogen abstraction to conformational changes in large flexible systems. A significant portion of this thesis has been published, and revised versions of these publications are presented where noted.

We begin in Chapter 2 by formulating and testing the proposed method for the calculation of elevated pressure Gibbs energy profiles by its application to a small chemical system: radical hydrogen transfer between methane and a methyl radical. We also investigate other associated effects of high pressure on the system, such as the pressure-induced changes in the enthalpy profile and the kinetic effect of pressure-induced viscosity. Portions of the work presented in this chapter are adapted from three publications: Spooner, J.; Wiebe, H.; Boon, N.; Deglint, E.; Edwards, E.; Yanciw, B.; Patton, B.; Thiele, L.; Dance, P.; Weinberg, N. "Molecular dynamics calculation of molecular volumes and volumes of activation" *Phys. Chem. Chem. Phys.* **2012**, *14* (7), 2264–2277; Wiebe, H.; Spooner, J.; Boon, N.; Deglint, E.; Edwards, E.; Dance, P.; Weinberg, N. *J. Phys. Chem. C* "Calculation of molecular volumes and volumes of activation using molecular dynamics simulations" **2012**, *116* (3), 2240–2245 and Wiebe, H.; Prachnau, M.; Weinberg, N. "Hydrogen transfer reactions in viscous media – Potential and free energy surfaces in solvent-solute coordinates and their kinetic implications" *Can. J. Chem.* **2013**, *91* (9), 787–794.

In Chapter 3, the ability of volume profiles to predict the structural properties of systems at elevated pressure is utilized in a discussion of pressure-controlled molecular machines. Using eq. 1.10, we predict pressure-induced conformational changes in model molecular wires, and design a prototype pressure-controlled unidirectional molecular rotor. The work presented here is adapted from Spooner, J.; Wiebe, H.; Louwerse, M.;

Reader, B.; Weinberg, N. "Theoretical analysis of high-pressure effects on conformational equilibria" *Can. J. Chem.* **2017**, *96*(2), 178-189.

In Chapter 4, we formulate and test a novel approach towards the determination of TS ensembles in flexible systems, based on a combination of theoretical volume profiles with experimental activation volume data. We demonstrate its successful application to internal rotation in 1,12-dimethoxy-[12]-paracyclophane (adapted from Wiebe, H.; Louwse, M.; Weinberg, N. "Theoretical volume profiles for conformational changes: Application to internal rotation of benzene ring in 1,12-dimethoxy-[12]-paracyclophane" *J. Chem. Phys.* **2017**, *146* (10), 104107) and the unfolding of a model chain system, serving as a prototype for macromolecular chains (adapted from Wiebe, H.; Weinberg, N. "Theoretical volume profiles as a tool for probing transition states: Folding kinetics" *J. Chem. Phys.* **2014**, *140* (12), 124105).

In order to apply the method developed in Chapter 4 to analysis of conformational changes in highly complex systems such as proteins, a reaction coordinate for the folding process in such systems must be defined. Therefore, in Chapter 5, we investigate the different collective variables traditionally used as reaction coordinates for multidimensional Gibbs energy surfaces, discuss their limitations, and propose a collective coordinate with a new functional form. In Chapter 6, some preliminary results towards the definition of a reaction coordinate for biological systems, such as helical polypeptides and the staphylococcal nuclease protein, are presented and discussed.

## Chapter 2.

# High pressure effect on radical hydrogen abstraction<sup>†</sup>

## 2.1. Radical hydrogen abstraction

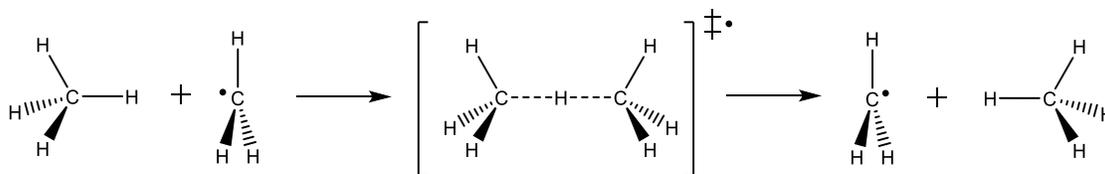
The kinetic effect of pressure on chemical reactions can manifest through two different mechanisms. The first, as detailed in section 1.2.4, is the pressure-induced modification of the Gibbs energy surface, which results in a change in the activation barrier. This can lead to either an increase or a decrease in the reaction rate depending on the volumetric properties of the system in question. The second mechanism occurs through the pressure-induced increase in solvent viscosity, as described in section 1.1.4. When noticeable, this effect always results in a decrease in reaction rate. The interplay between these two effects can result in complex dependences of experimentally observed rate constants.<sup>65–69</sup> To investigate this interplay and to illustrate the utility of MD-calculated volume profiles for high pressure kinetic analysis, we performed an extensive study of these effects on radical hydrogen abstraction reactions in hydrocarbon systems.

Radical hydrogen abstraction is an important elementary step that is found in a wide variety of chemical reactions, ranging from the relatively simple halogenation of saturated hydrocarbons to the vastly complex petrochemical processes of natural oil formation.<sup>130</sup> Since many of such processes occur at elevated pressures and in viscous environments,<sup>131</sup> it is important to understand how pressure affects the kinetics of these reactions. As a model system, we have chosen to study the hydrogen transfer between methane and a methyl radical, shown in Figure 2.1.

---

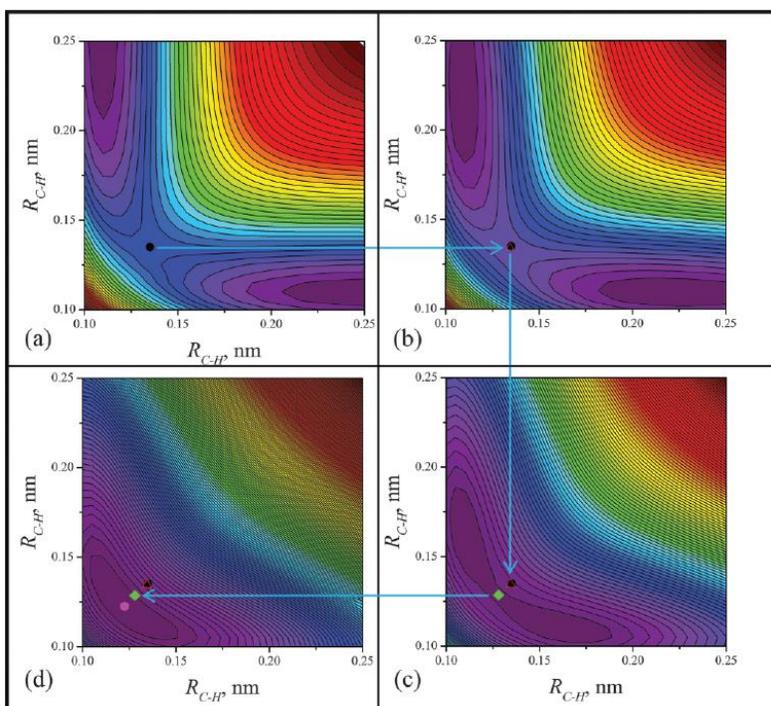
<sup>†</sup> Portions of this chapter are adapted from previously published work. The parameterization scheme of section 2.2.2 and the activation volume calculation presented in section 2.3.1 were originally published in Spooner, J.; Wiebe, H. *et al*, *Phys. Chem. Chem. Phys.* **2012**, *14* (7), 2264–2277 and Wiebe, H. *et al*, *J. Phys. Chem. C* **2012**, *116* (3), 2240–2245. In addition, the viscosity effect discussed in section 2.4 is adapted from Wiebe, H.; Prachnau, M.; Weinberg, N. *Can. J. Chem.* **2013**, *91* (9), 787–794.

All calculations and analysis in this chapter were performed by H. Wiebe apart from the elevated pressure enthalpy profiles of section 2.3.3 and the nonequilibrium solvation surfaces of section 2.4 which were calculated by undergraduate students A. Sherwood and M. Prachnau under the supervision of H. Wiebe.



**Figure 2.1** The hydrogen abstraction reaction between methane and a methyl radical.

A recent *ab initio* study<sup>132</sup> of this system in a compressed medium has shown that with increased pressure, the positions of the reactant and product on the potential energy surface shift towards the TS, with an eventual collapse into a single minimum as illustrated in Figure 2.2. However, this result was obtained using a static approach with pressure imposed by compression of an inert solid matrix, and thus the dynamic effects of temperature and solvent-solute interactions were ignored. We extend this study here using MD simulations to obtain a better understanding of the process by including these dynamic effects.

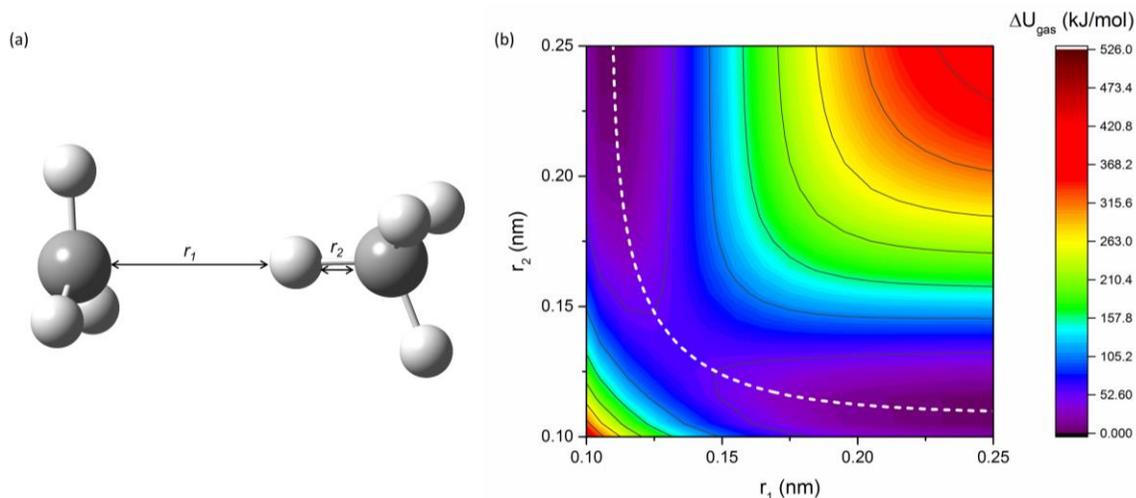


**Figure 2.2** The effect of compression on the PES of the methyl-methane hydrogen abstraction system for (a) ambient conditions, (b) low, (c) medium, and (d) high compression. As compression increases, the reaction barrier decreases and the positions of the reactant and product begin to shift towards the TS. The blue arrows trace the position of the TS. The contour spacing is 10 kJ/mol. Figure reprinted from ref. 132, © 2016 Canadian Science Publishing or its licensors. Reproduced with permission.

## 2.2. Description of the system

### 2.2.1. Reaction coordinate

A B3LYP/6-31++G(d,p) potential energy surface was obtained for this system in  $r_1$  and  $r_2$  coordinates, defined in Figure 2.3(a). This surface is shown in Figure 2.3(b), and the TS was identified at the saddle point with  $r_1 = r_2 = 0.1346$  nm.



**Figure 2.3** (a) The two C-H bond lengths labelled  $r_1$  and  $r_2$ , which are used in the definition of the reaction coordinate for this system. (b) Gas phase B3LYP/6-31++G(d,p) potential energy surface in  $r_1$  and  $r_2$  coordinates. The white dashed line is the reaction path.

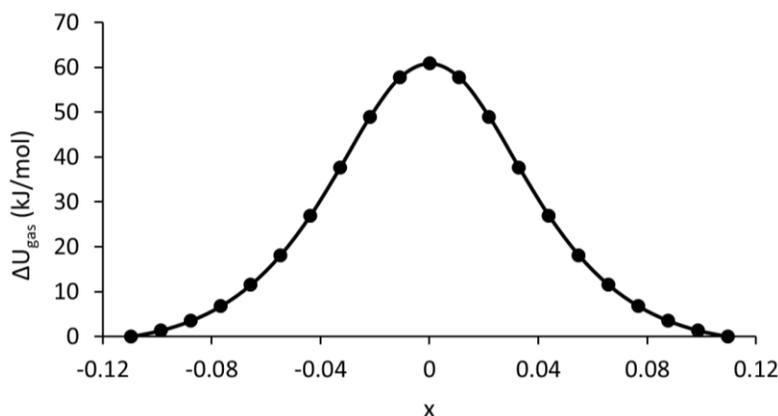
The reaction path (white dashed line) was described parametrically in terms of the C-H bond order  $n$  using the bond energy bond order (BEBO) equations (eq. 2.1), with parameters  $r_e = 0.1091$  nm and  $A = 0.037$  fitted to reproduce the reactant ( $n = 0$ ), TS ( $n = 0.5$ ) and product ( $n = 1$ ) structures.

$$\begin{aligned} r_1 &= r_e - A \ln n \\ r_2 &= r_e - A \ln(1 - n) \end{aligned} \tag{2.1}$$

The bond order  $n$  is non-uniformly scaled along the reaction path. An alternative uniformly scaled reaction coordinate  $x$  can be defined as the signed length along the reaction path, using eq. 2.2.

$$x(n) = r_e \int_{0.5}^n \frac{\sqrt{n^2 + (1-n)^2}}{n(1-n)} dn \quad (2.2)$$

To obtain a one-dimensional potential energy profile along the  $x$  coordinate, 10 evenly spaced structures were selected on either side of the TS ( $x = 0$ ) to yield a total of 21 structures along the reaction path. These structures were chosen such that  $r_1$  and  $r_2$  did not exceed a value of 0.25 nm ( $0.0221 \leq n \leq 0.9779$ ). The resulting gas phase potential energy profile  $\Delta U_{gas}(x)$  calculated at the B3LYP/6-31++G(d,p) level is shown in Figure 2.4.



**Figure 2.4** Gas phase B3LYP/6-31++G(d,p) energy profile  $\Delta U_{gas}(x)$  along reaction coordinate  $x$ .

### 2.2.2. Force field parameterization

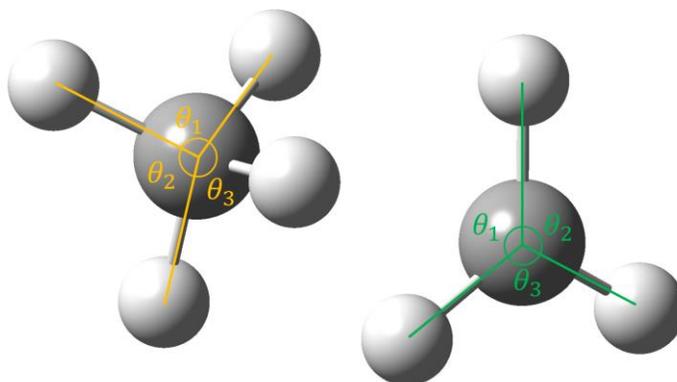
For MD calculations, force field parameters are required for all molecules in the system. The OPLS force field<sup>125,126</sup> was used to describe the  $n$ -hexane solvent and the reactant and product states (i.e. methane and the methyl radical). Since standard force fields like OPLS are only parameterized for stable species, new parameters had to be generated for the transient structures along the reaction coordinate.<sup>83</sup>

For the non-bonded parameters, such as the Lennard Jones  $\sigma$  and  $\varepsilon$  and the charges  $Q_i$  on the atoms, a linear interpolation approach was used based on the hybridization of the carbon atoms. The intermediate degrees of hybridization for the transient structures were assessed via the pyramidity parameter  $\gamma$  defined by eq. 2.3, where  $\theta_{sp^2}$  is the trigonal planar angle of  $120^\circ$ ,  $\theta_{sp^3}$  is the tetrahedral angle of  $109.47^\circ$  and

$\sum \theta$  is the sum of the three adjacent HCH angles in the transient state excluding the transferred hydrogen, as illustrated in Figure 2.5.

$$\gamma = \frac{3\theta_{sp^2} - \sum \theta}{3\theta_{sp^2} - 3\theta_{sp^3}} \quad (2.3)$$

The  $\gamma$  parameter is equal to 0 in a pure  $sp^2$  state, as in the methyl radical, and equal to 1 in a pure  $sp^3$  state, as in methane. The non-reactive hydrogens were assigned the  $\gamma$  value of the carbon they were bonded to.



**Figure 2.5** The pyramidity parameter  $\gamma$  of each carbon atom was assessed using the sum of the three H-C-H bond angles labelled in yellow or green.

The Lennard-Jones parameters  $\sigma$  and  $\varepsilon$  for the atoms in the transient structures were then smoothly interpolated between their respective values  $x_R$  and  $x_P$  in the reactant and product states according to eq. 2.4, where  $\gamma_R$  and  $\gamma_P$  are the pyramidity parameters of the atom in question in the reactant and product states.

$$x(\gamma) = \frac{(\gamma - \gamma_R)x_P + (\gamma_P - \gamma)x_R}{\gamma_P - \gamma_R} \quad (2.4)$$

The charges on the carbon atoms were also smoothly interpolated between their reactant and product values according to eq. 2.4. For the reactive hydrogen, the  $\sigma$ ,  $\varepsilon$  and charge were kept constant at the standard OPLS  $sp^3$  hydrogen value. The charges on the remaining hydrogens were set such that they compensated the charge of the carbon atom to which they were attached plus half the charge of the reactive hydrogen.

For bonded parameters, such as bond stretching and angle bending, the force constants were obtained by fitting to QM deformation energies using the ForceBalance software package.<sup>127</sup> Random conformations of the system were obtained from MD simulations using dummy force field parameters, and the potential energy of these conformations was obtained via QM B3LYP/6-311++G(d,p) single point energy calculations. The force constants  $k_i$  were then fitted to minimize the difference between the QM and MD energy responses according to eq. 2.5, under the assumption that the energy response is harmonic for each geometrical parameter  $x_i$ .

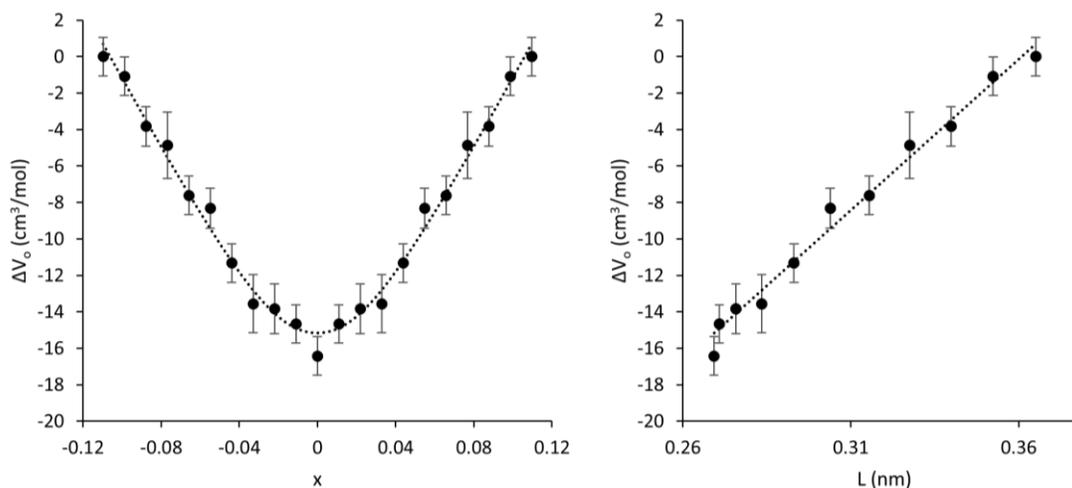
$$U_{QM} - U_{MD} = \frac{1}{2} \sum_i k_i (x_i - x_o)^2 \quad (2.5)$$

In order to maintain a constant value of  $x$  for each structure during the simulation, the  $r_1$  and  $r_2$  bonds were constrained using the LINCS algorithm.<sup>133</sup> The force constant for the intermolecular C-H-C angle was set to a constant value of 276.144 kJ mol<sup>-1</sup> rad<sup>-2</sup>, which is the standard value for the H-C-H bending force constant in the OPLS force field.<sup>125,126</sup>

## 2.3. High pressure reaction profiles

### 2.3.1. Activation volume and volume profile

The volume profile  $\Delta V_o(x)$  for this system was calculated at 1 bar and 298 K using the Archimedean displacement model in a solvent consisting of 256 molecules of *n*-hexane. Due to the symmetry of this system, calculations were performed for the first 11 structures along the  $x$  coordinate, and the remaining 10 were mirrored across  $x = 0$ . Fifty parallel seeds were used for each calculation. The resulting volume profile is shown in Figure 2.6, scaled relative to the structures furthest from the TS ( $x = \pm 0.11$ ). The error bars were calculated as the standard error of the mean using the block averaging technique.<sup>134,135</sup>



**Figure 2.6** The ambient pressure volume profile  $\Delta V_o(x)$  for the methyl-methane system (left). The dotted line is a fit to the profile obtained from a linear plot of volume against the distance  $L$  between carbon atoms,  $L = r_1 + r_2$  (right).

Expectedly, the volume profile reaches its minimum value at  $x = 0$ , representing the TS. Direct calculation of  $\Delta V^\ddagger$  as the difference in volume between the TS ( $x = 0$ ) and the reactant species (eq. 2.6) gives a  $\Delta V^\ddagger$  value of  $-22 \pm 3 \text{ cm}^3/\text{mol}$ .<sup>83,111</sup>

$$\Delta V^\ddagger = V^\ddagger - (V_{\text{methane}} + V_{\text{methyl}}) \quad (2.6)$$

Since the activation volume  $\Delta V^\ddagger$  is negative, this reaction is accelerated by pressure. While an experimentally-determined  $\Delta V^\ddagger$  for this system does not exist, this value is in line with measured values for similar reactions,<sup>4,136</sup> and the obtained computational error is within the errors typically obtained in high pressure experiments.<sup>11–13</sup> The  $5.5 \text{ cm}^3/\text{mol}$  difference between this  $\Delta V^\ddagger$  and the minimum value of  $\Delta V_o(x)$  reflects the difference in the choice of a reference point: the reactant ( $x = -\infty$ ,  $n = 0$ ) for  $\Delta V^\ddagger$  and  $x = -0.11$  ( $n = 0.019$ ) for  $\Delta V_o(x)$ .

### 2.3.2. High pressure Gibbs energy profile

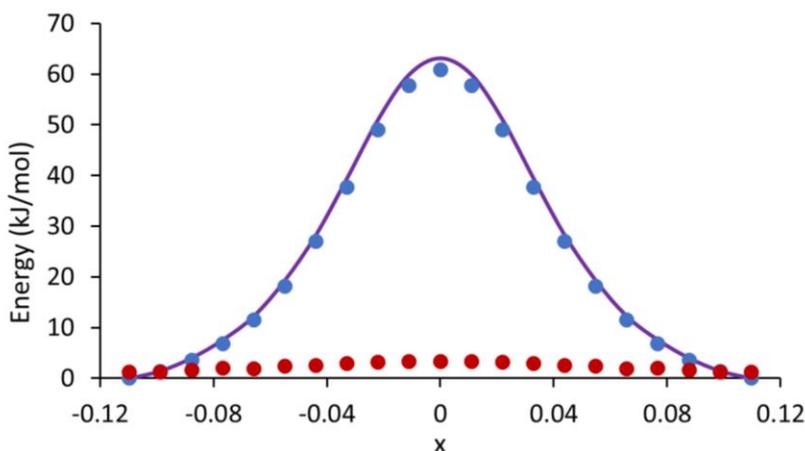
The high pressure Gibbs energy profiles  $\Delta G(x; P)$  for this reaction were calculated using eq. 2.7 by a combination of the ambient pressure Gibbs energy profile  $\Delta G_o(x)$  and volume profile  $\Delta V_o(x)$ .

$$\Delta G(x; P) = \Delta G_o(x) + (P - P_o)\Delta V_o(x) \quad (2.7)$$

The required ambient pressure Gibbs energy profile  $\Delta G_o(x)$ , was obtained by adding the solvation Gibbs energy  $\Delta G_{solv}(x)$  to the gas phase potential energy profile  $\Delta U_{gas}(x)$  (Figure 2.4) according to eq. 2.8.

$$\Delta G_o(x) = \Delta U_{gas}(x) + \Delta G_{solv}(x) \quad (2.8)$$

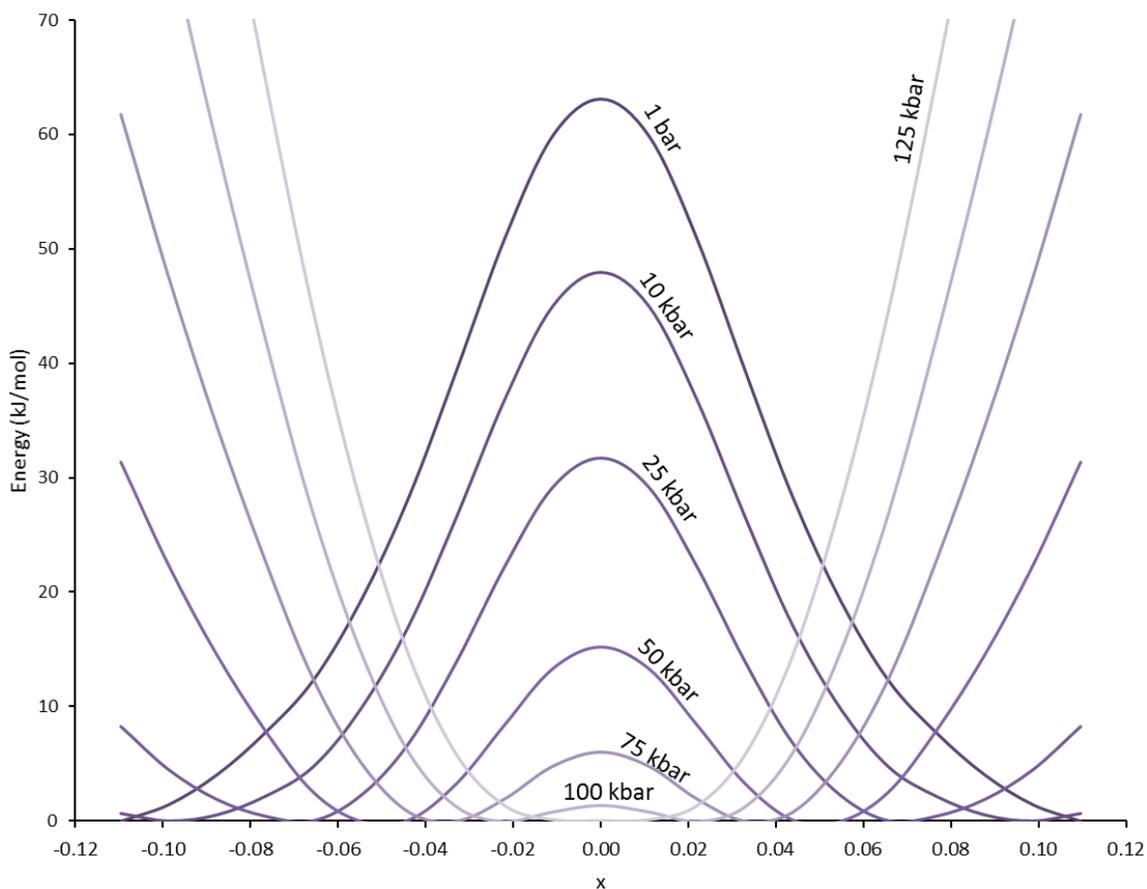
For each of the 21 structures along the reaction coordinate, the solvation Gibbs energy was calculated using thermodynamic integration with respect to solvent-solute interactions scaled between  $\lambda = 0$  (no interaction) and  $\lambda = 1$  (full interaction) over 21 points. For each point, 1.0 ns trajectories were run in triplicate. The resulting  $\Delta G_{solv}(x)$  profile is shown in Figure 2.7 along with the final  $\Delta G_o(x)$ . As this is a nonpolar reaction in a nonpolar solvent, the solvation energy is almost negligible in comparison to the  $\Delta U_{gas}(x)$  term, and the resulting  $\Delta G_o(x)$  profile is nearly indistinguishable from the QM-calculated potential energy profile.



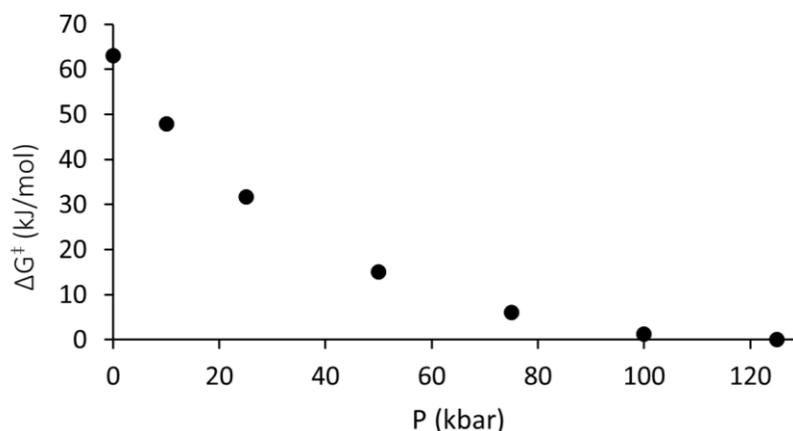
**Figure 2.7** The ambient pressure Gibbs energy profile  $\Delta G_o(x)$  (solid purple line) obtained as a sum of the gas phase potential energy  $\Delta U_{gas}(x)$  (blue data points) and the solvation Gibbs energy  $\Delta G_{solv}(x)$  (red data points).

The resulting high pressure  $\Delta G(x; P)$  profiles were calculated according to eq. 2.7 and are shown in Figure 2.8. As pressure increases, two effects are seen. First, the reaction barrier decreases dramatically as shown in Figure 2.9, requiring a pressure of

125 kbar to eliminate the barrier entirely. Secondly, the reactant and product minima shift towards  $x = 0$  while the TS remains stationary, with an eventual merge into a single minimum occurring at the highest pressure. This is in agreement with the original results obtained using the *ab initio* approach,<sup>132</sup> but now we are able to assess a pressure range for this transition. Similar results have recently been reported for pericyclic reactions.<sup>137</sup>



**Figure 2.8** The calculated  $\Delta G(x; P)$  profiles show that as pressure increases, the reaction barrier decreases. A pressure of 125 kbar was required to eliminate the reaction barrier. Above that pressure, the Gibbs energy profile possesses a single minimum at  $x = 0$  representing a symmetric structure  $\text{CH}_3\text{HCH}_3$  as a single stable state of the system.



**Figure 2.9** The effect of pressure on the calculated  $\Delta G^\ddagger$  values. The reaction barrier becomes practically negligible above 100 kbar.

### 2.3.3. High pressure effect on activation enthalpy

The effect of pressure on the Gibbs energy profile largely depends on the shape of the volume profile. For systems with negative  $\Delta V^\ddagger$  values, like the methyl-methane hydrogen abstraction reaction or Diels-Alder reactions, the activation Gibbs energies  $\Delta G^\ddagger$  decrease with pressure due to the increasingly negative contribution of the  $(P - P_o)\Delta V_o(x)$  term in eq. 2.7. Interestingly, it has been observed that activation enthalpies  $\Delta H^\ddagger$  can increase with pressure for systems with negative  $\Delta V^\ddagger$ , as illustrated by the experimental data in Table 2.1 for the Diels-Alder dimerization of cyclopentadiene.<sup>138</sup>

**Table 2.1** Experimentally measured pressure-induced changes in activation Gibbs energy ( $\Delta\Delta G^\ddagger$ ) and activation enthalpy ( $\Delta\Delta H^\ddagger$ ) relative to ambient pressure for the Diels-Alder dimerization of cyclopentadiene. Data from ref. 138.

Pressure (kbar)	$\Delta\Delta G^\ddagger$ (kJ/mol)	$\Delta\Delta H^\ddagger$ (kJ/mol)
0.001	0	0
1.000	-1.01	1.26
2.000	-1.89	2.09
3.000	-2.77	5.44

El'yanov and Gonikberg,<sup>139</sup> rationalized this in terms of the thermodynamic relationship (eq. 2.9) between pressure derivatives of  $\Delta H$  and  $\Delta G$ .

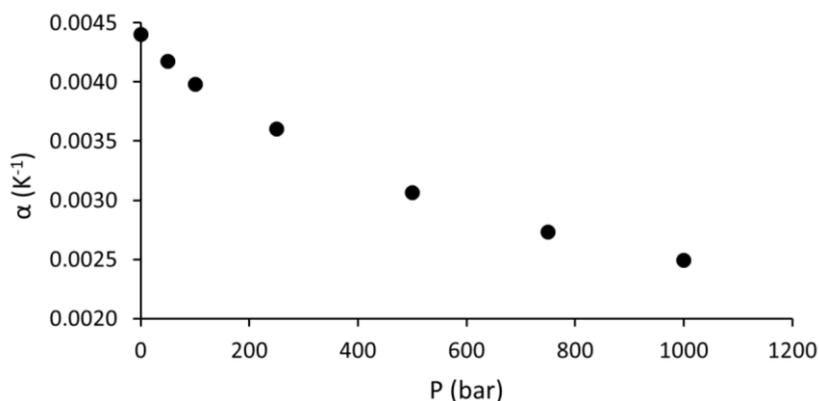
$$\left(\frac{\partial\Delta H}{\partial P}\right)_T = (1 - \kappa T) \left(\frac{\partial\Delta G}{\partial P}\right)_T \quad (2.9)$$

The thermal expansion factor  $\kappa$ , defined in eq. 2.10, describes how the reaction volume  $\Delta V$  for a process is affected by temperature. These authors analyzed  $\Delta V$  data for 34 organic reactions and found the average value for  $\kappa$  to be  $0.0044 \text{ K}^{-1}$  at  $20^\circ\text{C}$  and 1 bar, which makes the  $1 - \kappa T$  factor negative at ambient conditions.

$$\kappa = \frac{1}{\Delta V} \left( \frac{\partial \Delta V}{\partial T} \right)_P \quad (2.10)$$

They further assumed that  $\kappa^\ddagger$  has a similar value, thus explaining the inverse relationship between the behaviour of  $\Delta H^\ddagger(P)$  and  $\Delta G^\ddagger(P)$ .

However, it needs to be stated that thermal expansion coefficients of pure liquids decrease significantly with pressure, as shown in Figure 2.10 for the case of *n*-hexane at 300 K. It can therefore be expected that the value of  $\kappa^\ddagger$  will likewise decrease with pressure, in which case the  $1 - \kappa^\ddagger T$  term may become positive at sufficiently high pressures, and the sign of the pressure derivative of  $\Delta H^\ddagger$  will match the sign of the pressure derivative of  $\Delta G^\ddagger$ . That is, the  $\Delta H^\ddagger$  of a reaction with a negative  $\Delta V^\ddagger$  may be expected to start decreasing with pressure at sufficiently high compressions.

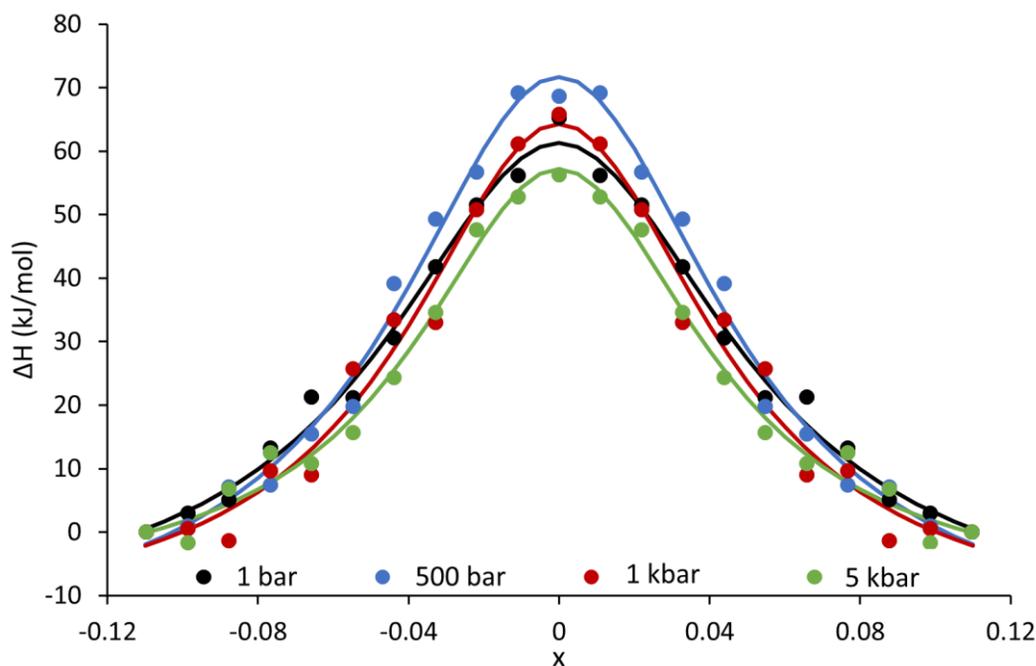


**Figure 2.10** The thermal expansion coefficient  $\alpha = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_P$  for *n*-hexane decreases with increasing pressure, due to the tighter packing of the molecules at higher pressures. The  $\alpha$  values were obtained by numerical differentiation of the equation of state from data in ref. 140.

To investigate this effect for our system, we calculated enthalpy profiles for pressures ranging from 1 bar to 5 kbar, according to eq. 2.11.

$$\Delta H(x) = \Delta U(x) + P\Delta V_o(x) = \Delta U_{gas}(x) + \Delta U_{nb}(x) + P\Delta V_o(x) \quad (2.11)$$

The potential energy profile was broken up into two terms, the QM-calculated gas phase profile  $\Delta U_{gas}(x)$  from Figure 2.4 and the MD-calculated non-bonded potential energy profile  $\Delta U_{nb}(x)$ , which describes the solvent-solute and solvent-solvent interactions obtained by averaging over 1.0 ns trajectories for the first 11 structures, with the remaining 10 obtained via symmetry across  $x = 0$ .



**Figure 2.11** Enthalpy profiles for the methyl-methane system at four different pressures. The MD-calculated data points were fit with a Lorentzian function as a visualization aid.

**Table 2.2** Barrier heights  $\Delta H^\ddagger$  as a function of pressure.

Pressure (bar)	$\Delta H^\ddagger$ (kJ/mol)
1	61
500	72
1000	64
5000	57

The resulting profiles are shown in Figure 2.11 with explicit pressure dependence of the activation enthalpies  $\Delta H^\ddagger$  listed in Table 2.2. An initial increase in  $\Delta H^\ddagger$  up to 500 bar is followed by a decrease at higher pressures, which indicates that the switch in the sign

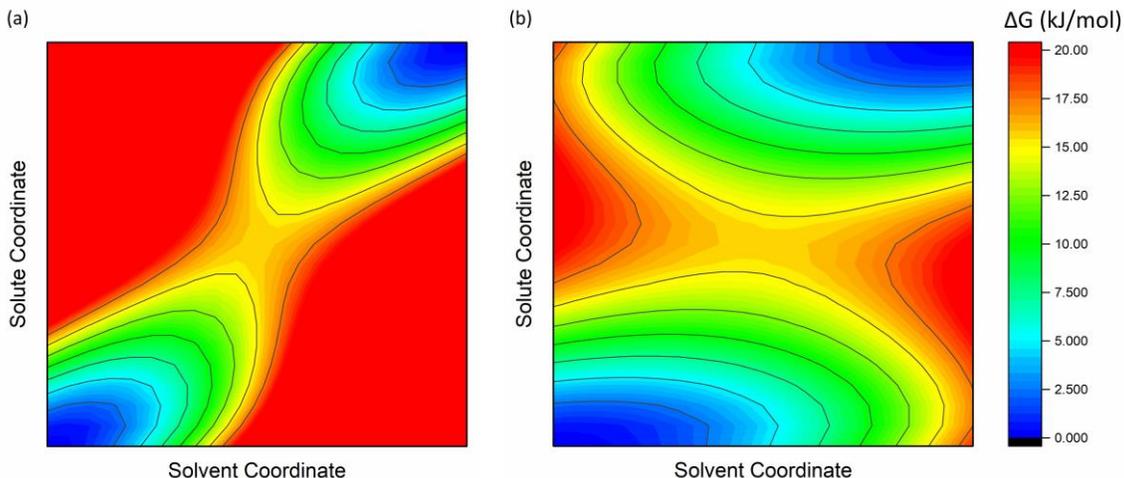
of the  $1 - \kappa^\ddagger T$  term from negative to positive occurs for this system between 500 bar and 1 kbar.

## 2.4. Pressure-induced viscosity effect

In nonviscous media, the solvent is fast enough to restructure synchronously with the reaction system, and the reaction rate is dependent only on the activation barriers for the process. In viscous media, however, the solvent is slower than the reaction system, which can result in a situation where the size and shape of the solvent cavity no longer matches the geometry of the solute. First described by Kramers,<sup>53,54</sup> this viscosity-induced effect, known as nonequilibrium solvation, leads to the interference of the solvent with the restructuring of the reaction system and can result in a significant reduction in reaction rates. Since pressure invariably increases the viscosity of the reaction medium,<sup>141</sup> pressure-induced viscosity effects become significant at sufficiently high compressions.<sup>65–69,142,143</sup> Here, the importance of these effects is assessed for the methyl-methane hydrogen transfer system.<sup>144</sup>

### 2.4.1. Solute and solvent coordinates

The strength of nonequilibrium solvation effects in a particular reaction can be evaluated by analysis of its 2D Gibbs energy surface (GES) in solute-solvent coordinates.<sup>145,146</sup> If the GES possesses a consecutive arrangement of reactant and product valleys as shown in Figure 2.12(a), reaction can occur only through a synchronized restructuring of the solute and solvent.<sup>30</sup> This solvent-solute system is therefore quasi-one-dimensional, and its classical reaction trajectories must pass from the reactant to the product in the vicinity of the TS. As a result, the kinetic effects of viscosity for this system are expected to be strong. A different type of surface, which possesses reactant and product valleys that are parallel, is shown in Figure 2.12(b). Reaction trajectories here can cross the barrier at multiple points representing a wide range of solvent configurations, and thus can deviate far from the TS.<sup>31,32</sup> The effects of viscosity on this type of system will be weaker, as the solute and solvent are less strongly coupled.



**Figure 2.12 Two types of Gibbs energy surfaces in solvent-solute coordinates: (a) strong solute-solvent coupling; (b) weak solute-solvent coupling.**

The reaction coordinate  $x$ , as defined in section 2.2.1, is a natural choice to represent the solute configuration. The solvent coordinate is defined using the approach described in ref. 144 and outlined below.

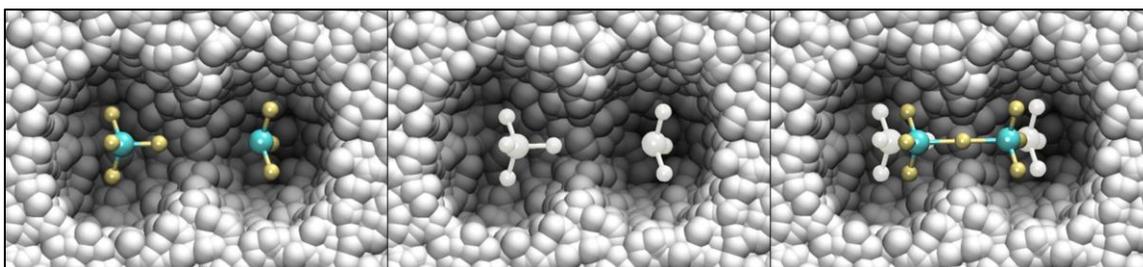
In an MD simulation, the trajectories of the solvent  $Z(t)$  and solute  $z(t)$  are interdependent due to the solvent-solute interactions in the system. If the solute is frozen in some geometry  $z(x)$ , determined by the solute coordinate  $x$ , then the solvent trajectory will become parametrically dependent on  $x$  as well:  $Z(t|x)$ . The total energy of the system will also become  $x$  dependent, as it is averaged over both solute and solvent trajectories (eq. 2.12).

$$E(x) = \langle E(z(x), Z(t|x)) \rangle \quad (2.12)$$

To describe nonequilibrium solvation, a mismatch needs to be introduced between  $Z(t|x)$  and  $z(x)$ . To do so, the solvent-solute system is first run with the solute at a given value of the solute coordinate,  $y$ , referred to as the “pilot” solute. This produces a solvent trajectory that is equilibrated to the pilot solute configuration:  $Z(t|y)$ . The pilot solute is then removed, and the “true” solute is added at a different configuration,  $x$ . The nonequilibrium solvation energy  $E(x, y)$  can then be assessed according to eq. 2.13 as the interaction between solute in one configuration  $x$  and solvent equilibrated to the pilot solute at configuration  $y$ .

$$E(x, y) = \langle E(z(x), Z(t|y)) \rangle \quad (2.13)$$

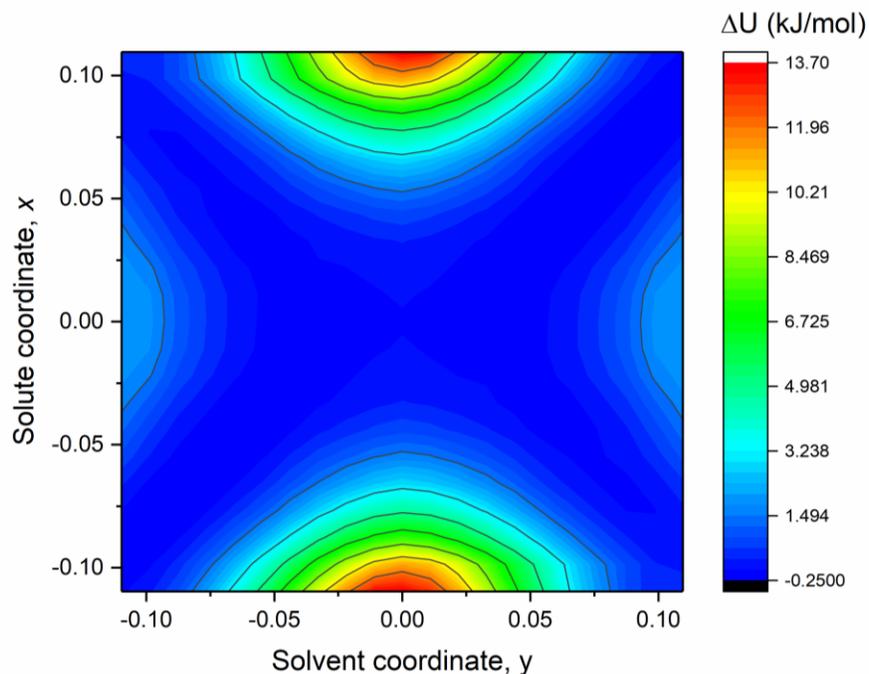
Parameters  $x$  and  $y$  of eq. 2.13 can serve as the solute and solvent coordinates, respectively. As seen in Figure 2.13, the equilibration of the solvent trajectory around the pilot solute results in a solvent cavity that is determined by the size and shape of the solute. Replacement of the pilot solute with the true solute shows the mismatch in both size and shape of the solvent cavity, indicating the nonequilibrium solvation.



**Figure 2.13** Illustration of the solvent coordinate construct. The solvent-solute system is first equilibrated to the pilot solute in some configuration  $y$  (left). The pilot solute is removed, but the solvent trajectory remains equilibrated to it (middle). The pilot solute is then replaced with the solute in a different configuration  $x$  (right).

## 2.4.2. Gibbs energy surfaces

To obtain the GES for the methyl-methane system in solvent-solute coordinates, each of the 21 structures along the solute coordinate  $x$  were frozen in space and inserted as a pilot solute into a cubic MD simulation box containing 256 hexane molecules as solvent. MD runs with a duration of 1.0 ns were then performed for each of these structures and the solvent trajectories were stored. The pilot solutes were then removed from the solvent trajectories and replaced with true solutes in various configurations,  $x$ . To minimize the degree of misfit between the solute and solvent, the position and orientation of the true solute in configuration  $x$  was optimized with respect to the solvent in configuration  $y$  by translating and rotating the solute until the lowest energy state was reached. The energy of the combined system,  $\Delta U_{solv}(x, y)$ , which includes all solvent-solvent and solvent-solute interactions, was then averaged over the entire trajectory to give the surface shown in Figure 2.14. The nonequilibrium solvation energy was highest at the points where the reactant- and product-like configurations of the solute ( $x \sim \pm 0.1$ ) were combined with a solvent trajectory equilibrated to a TS-like structure ( $y \sim 0$ ).

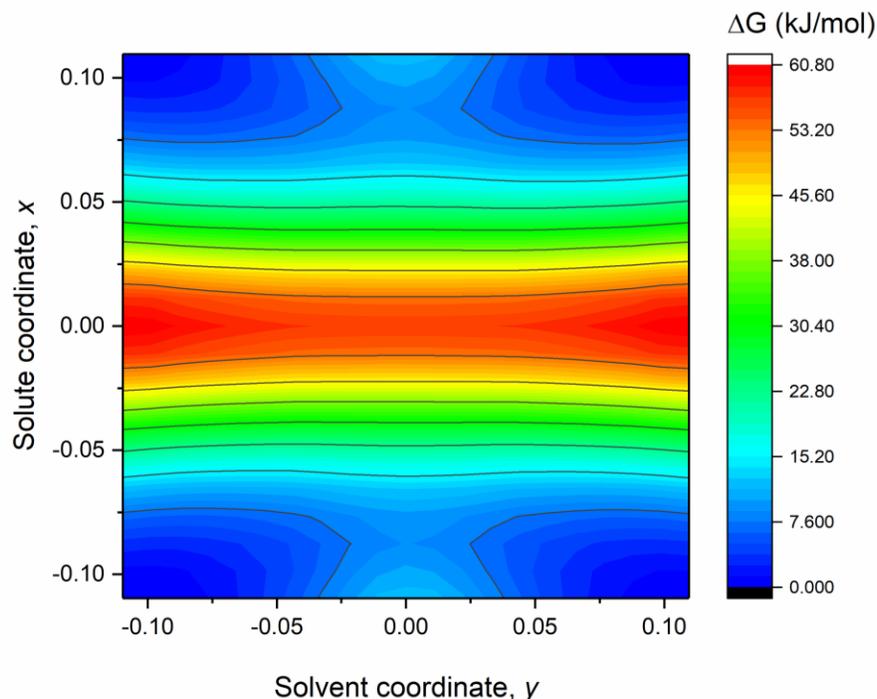


**Figure 2.14** Nonequilibrium solvation energy for the methyl-methane system in solvent-solute coordinates. The properly solvated reactant is in the bottom left corner, and the properly solvated product in the top right.

The GES was obtained according to eq. 2.14, by combining the nonequilibrium solvation energy  $U_{solv}(x, y)$  with the gas phase potential energy profile of the solute  $U_{gas}(x)$  and the Gibbs energy of the solvent configuration  $\Delta G_{solv}(y)$  of Figure 2.7.

$$\Delta G(x, y) = \Delta U_{solv}(x, y) + \Delta U_{gas}(x) + \Delta G_{solv}(y) \quad (2.14)$$

To minimize random errors, the energy values were smoothed out by averaging over four adjacent points and by using the symmetry condition  $\Delta G(-x, -y) = \Delta G(x, y)$ . The resultant GES is shown in Figure 2.15, with the properly solvated reactant in the bottom left corner and the properly solvated product in the top right corner. The surface displays a parallel arrangement of reactant and product valleys, which indicates that barrier crossing along solute coordinate  $x$  may occur at multiple configurations of the solvent  $y$ .



**Figure 2.15** Gibbs energy surface for the methyl-methane system in solvent-solute coordinates. The properly solvated reactant is in the bottom left and the properly solvated product is in the top right. This surface has two parallel valleys separated by the nearly uniform activation barrier, which implies that the kinetics of this reaction is nearly independent of solvent viscosity.

The pressure-induced viscosity effect for this system is therefore described by the Agmon-Hopfield model,<sup>57,58</sup> according to which the rate constant  $k$  can be found as a weighted average of the partial rate constants  $k_i(y)$  over an ensemble of solvent configurations described by the normalized viscosity-dependent probability distribution  $P(y)$ , per eq. 2.15.

$$k = \int k_i(y)P(y)dy \quad (2.15)$$

The partial rate constants  $k_i(y)$  can be described by eq. 2.16, with constant frequency factor  $\nu$  and solvent-configuration-dependent activation energy,  $\Delta G^\ddagger(y)$ .

$$k_i(y) = \nu e^{\frac{-\Delta G^\ddagger(y)}{RT}} \quad (2.16)$$

For this system, the GES possesses an energy ridge of nearly uniform altitude, and therefore the activation energies  $\Delta G^\ddagger(y)$  are practically independent of  $y$ , which according to eq. 2.16, makes partial rate constants  $k_i(y)$  also  $y$ -independent. As a result, the overall rate constant,  $k$ , becomes insensitive to the solvent distribution function,  $P(y)$ , and is thus practically independent of the solvent viscosity. Therefore, the net results of application of pressure to this system consists only of an increase in reaction rate due to the deformation of the Gibbs energy profile, as described in section 2.3, with the pressure-induced increase in viscosity having no significant effect on the reaction rate.

## 2.5. Conclusions

The utility of the MD displacement volume method has been demonstrated in a complete analysis of the various effects of pressure on a model reaction system: the hydrogen transfer between methane and a methyl radical in hexane solvent.

The activation volume for the system was calculated to be  $-22 \pm 3 \text{ cm}^3/\text{mol}$ , indicating that the reaction system decreases in volume as it approaches the TS, and that pressure accelerates the reaction. The volume profile for the system was also calculated and was used in conjunction with the ambient-pressure Gibbs energy profile to calculate high pressure Gibbs energy profiles. The Gibbs activation energy was found to decrease with pressure, accompanied by a shift in reactant and product minima towards the TS structure. A pressure of 125 kbar was required to completely eliminate the barrier, resulting in stabilization of the TS and destabilization of all other species. The activation enthalpy, on the other hand, was found to initially increase with pressure followed by a decrease between 500 bar and 1 kbar. The manipulation of Gibbs energy barriers and stabilization of transient species at high pressure has many interesting potential applications, and will be discussed further in Chapter 3.

An increase in pressure has the secondary effect of increasing the viscosity of the solvent, and this effect was investigated by construction of a 2D Gibbs energy surface in solvent-solute coordinates. The observed parallel orientation of reactant and product valleys along with the uniform energy barrier lead to the conclusion that viscosity does not significantly affect the kinetics of this reaction at high pressure. This is likely a result of rather weak solvent-solute interactions. Stronger solvent-solute interactions, perhaps due

to a change in polarity or the size of the system, would lead to more significant viscosity effects.

## Chapter 3.

# Pressure-induced conformational changes: Application to molecular machines<sup>‡</sup>

### 3.1. Molecular conformations and molecular machines

Crowned with the 2016 Nobel Prize in Chemistry,<sup>147</sup> molecular machines have been a subject of interest in experimental and theoretical research for quite some time.<sup>148–151</sup> Molecular machines are a class of molecules that perform an output function for a given input stimulus. These machines can be biological in origin, such as enzymes, ribosomes or motor proteins,<sup>152</sup> or they can be synthetically designed.<sup>148–151</sup> Recent years have seen the development of a wide variety of synthetic molecular machines including molecular switches,<sup>153,154</sup> wires<sup>155,156</sup> and rotors<sup>157</sup> for potential use in nanodevices.

The output function of a molecular machine is accomplished through a conformational change which results in a useful change in the properties of the molecule. This can involve the electronic properties of the molecule, such as its conductivity or luminescence, or it can be more mechanical in nature, such as the walking motion exhibited by the motor proteins kinesin and dynein.<sup>158,159</sup> Some examples of input stimuli used to activate molecular machines are the application of electromagnetic fields, a change in temperature or pH, chemical modification, or mechanical force.<sup>148–152</sup> One input stimulus that has, thus far, not been applied to molecular machines is pressure.

Since different conformations of molecules are likely to have different volumes, pressure is an ideal tool to drive these conformational changes. Here, the use of pressure to control two different types of molecular machines is demonstrated.<sup>160</sup> The first are polythiophenes, molecular wires whose  $\pi$  electron conjugation is dependent on their conformation, and thus their electrical and luminescent properties are available for

---

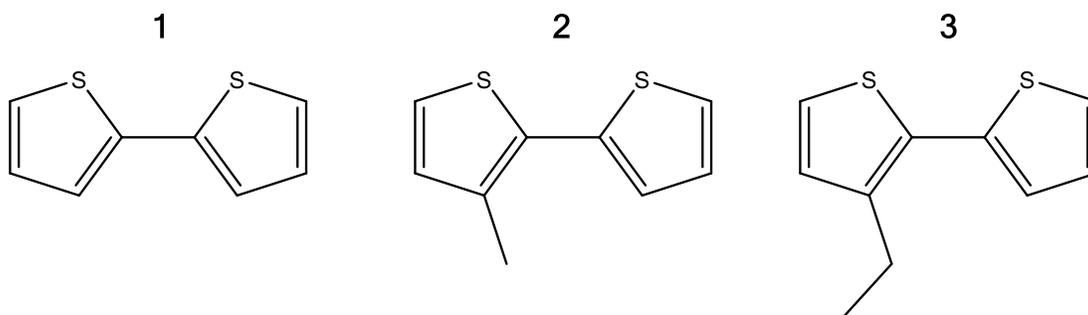
<sup>‡</sup> The work presented in this chapter was published as part of a study of the effects of pressure on different conformational equilibria: Spooner, J.; Wiebe, H.; Louwse, M.; Reader, B.; Weinberg, N. *Can. J. Chem.* **2017**, *Ahead of print*, DOI: 10.1139/cjc-2017-0411. All calculations and analysis presented in this chapter were performed by H. Wiebe apart from the volume profiles for the proposed rotor in section 3.3.2, which were calculated by undergraduate student M. Louwse under the supervision of H. Wiebe.

pressure control. The second is a preliminary design for a pressure-driven unidirectional molecular rotor.

## 3.2. Bithiophene as a prototype for molecular wires

Polythiophenes belong to a class of molecular machines known as molecular wires. The  $\pi$ -electron conjugation between the thiophene rings in these polymers makes them capable of conducting electricity, but the level of conjugation in the polymer is dependent on the relative alignment of the thiophene rings. A fully planar conformation would have the highest  $p$  orbital overlap and thus the highest conductivity. However, polythiophenes have been observed to adopt a twisted conformation in solution.<sup>161–163</sup> The ability to drive the thiophene rings in and out of a planar conformation would allow for control of electrical conductivity in the system.

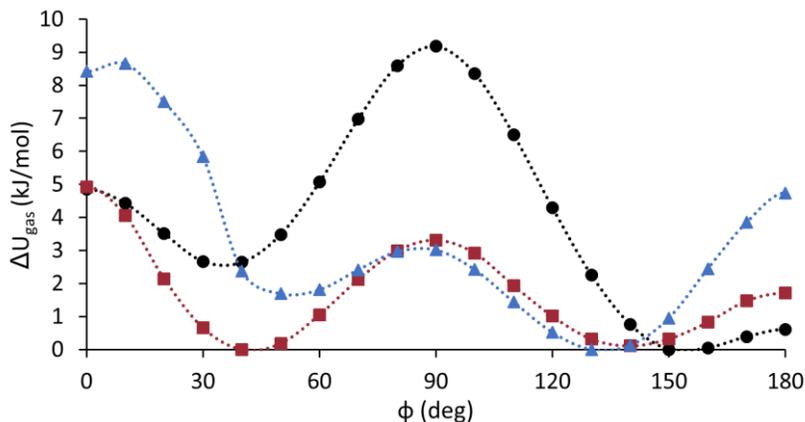
To investigate the effect of pressure on the conformational dynamics of polythiophenes, we have chosen to use a series of 2,2'-bithiophenes **1-3** shown in Figure 3.1 as models. The substitution at the 3 position has been demonstrated to increase the solubility of these compounds.<sup>164–167</sup>



**Figure 3.1** The three bithiophene systems used as models for polythiophene molecular wires.

### 3.2.1. Reference pressure torsion profiles

For these systems, the mutual orientation of the two thiophene rings is described by the S-C-C-S torsion angle,  $\phi$ , which serves as a reaction coordinate for the conformational change. The QM B3LYP/6-31++G(d,p) potential energy profiles  $\Delta U_{gas}(\phi)$  for rotation around  $\phi$  are shown in Figure 3.2.



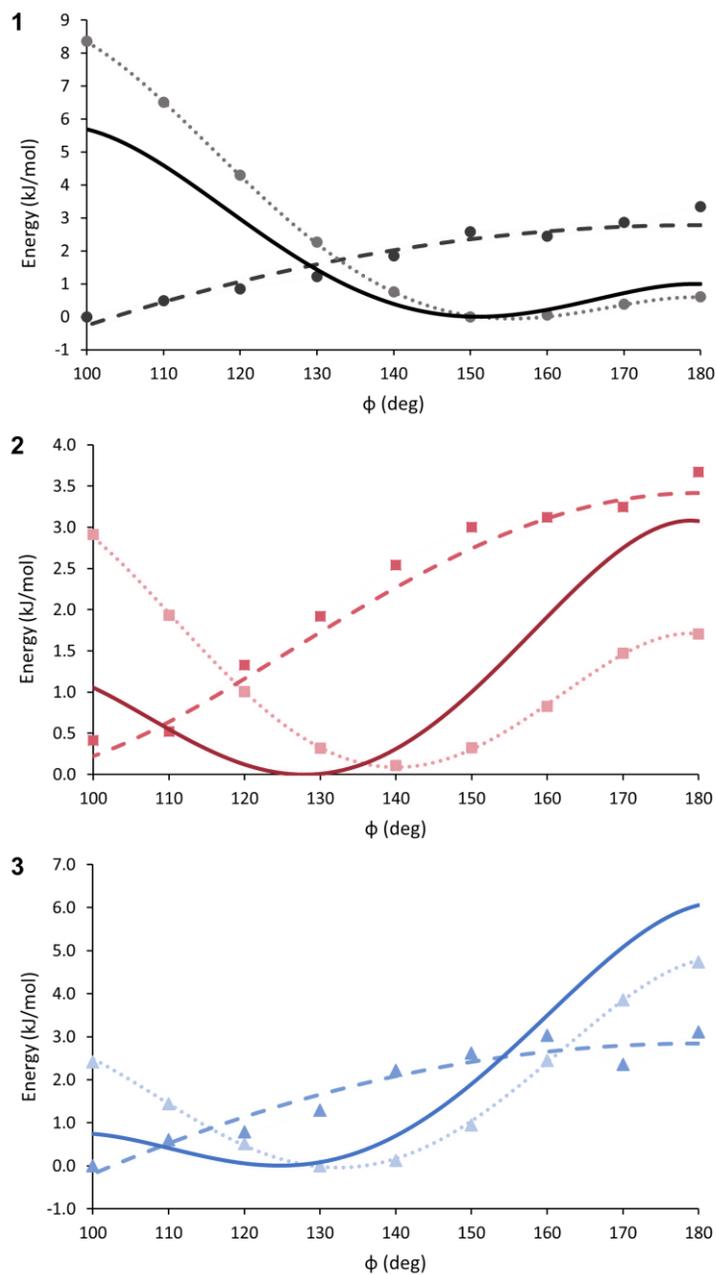
**Figure 3.2** Gas phase QM potential energy profiles for the three bithiophene models, **1** (black circles), **2** (red squares) and **3** (blue triangles). Lines are fit to smooth computational errors.

The relative stability of various conformations of these systems is controlled by two competing effects:  $\pi$ -electron conjugation and the steric repulsion between the two rings. The stabilizing effect of conjugation is greatest at the two planar conformations  $\phi = 0^\circ$ ,  $180^\circ$ , but the destabilizing effect of steric repulsion is highest at those same conformations. Therefore, the stable minima are twisted conformations that fall between  $120$ - $150^\circ$  for the *trans* state and between  $30$ - $60^\circ$  for the *cis* state. These minima drift further from the planar conformations with increasing substitution at the 3 position. The dramatic decrease in barrier height at  $\phi = 90^\circ$  between model **1** and models **2** and **3** can also be understood as a result of the increasing contribution of steric repulsion. At  $\phi = 90^\circ$  the inter-ring steric repulsion is at its minimum. Therefore, as the size of the substituent group grows, the conformation at  $\phi = 90^\circ$  is stabilized despite the loss of  $\pi$ -electron conjugation, leading to an overall decrease in the barrier height.

Since the potential energy barrier at  $180^\circ$  is lower than the barrier at  $0^\circ$  and the *trans* minimum is lower in energy and is therefore more highly populated than the *cis* minimum at ambient conditions, its movement as a function of pressure is of greater interest. To obtain the reference pressure Gibbs energy profile  $\Delta G_o(\phi)$ , around  $\phi = 180^\circ$  the solvation Gibbs energy  $\Delta G_{solv}(\phi)$  must be added to the  $\Delta U_{gas}(\phi)$  profile according to eq. 3.1.

$$\Delta G_o(\phi) = \Delta U_{gas}(\phi) + \Delta G_{solv}(\phi) \quad (3.1)$$

The  $\Delta G_{solv}(\phi)$  term was obtained using thermodynamic integration in a solvent of 256 toluene molecules at 300 K and 1 bar pressure. The simulations were run in triplicate for each of the 21  $\lambda$  points, and data were collected for 2.5 ns preceded by 100 ps of equilibration. The results are shown in Figure 3.3 for the three models.



**Figure 3.3** Gas phase potential energy (light data points, dotted lines) and solvation Gibbs energy (dark data points, dashed lines) summed to yield the reference pressure Gibbs energy profiles (solid lines) for the three bithiophene models. Dashed and dotted lines are fit to smooth computational errors.

In all cases, the  $\Delta G_{solv}(\phi)$  term is on the same scale as the  $\Delta U_{gas}(\phi)$  term, which means that the contributions from solvation are significant in this system. The minimum in  $\Delta G_{solv}(\phi)$  does not match the minimum in  $\Delta U_{gas}(\phi)$ , and so when the two terms are added together to produce  $\Delta G_o(\phi)$  the *trans* conformation becomes shifted even further from 180°.

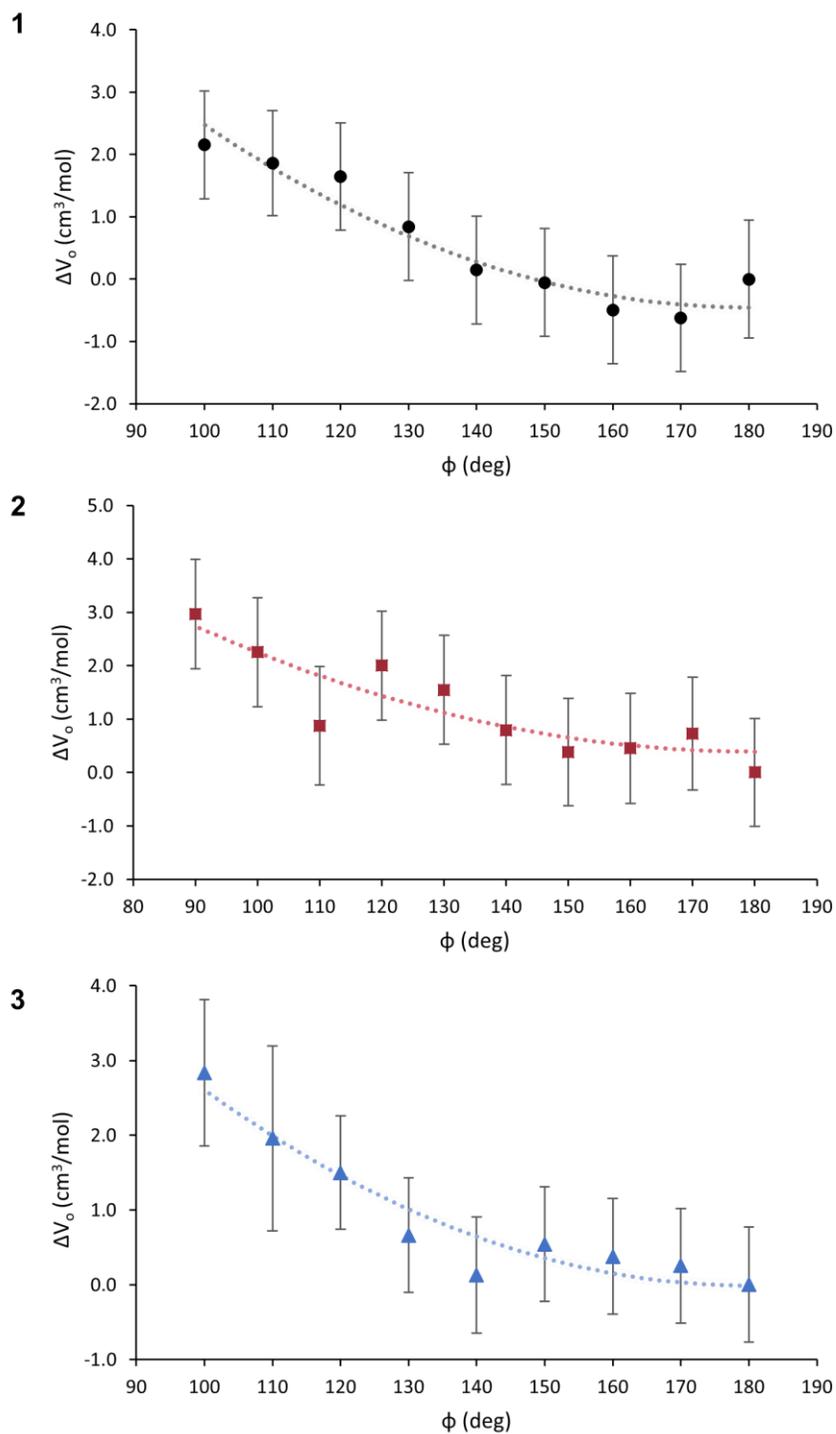
### 3.2.2. High pressure profiles and pressure controlled conformational changes

The effect of pressure on these models was obtained by combination of the reference pressure Gibbs energy profile  $\Delta G_o(\phi)$  with the pressure-dependent contribution from the volume profile  $\Delta V_o(\phi)$  via eq. 3.2.

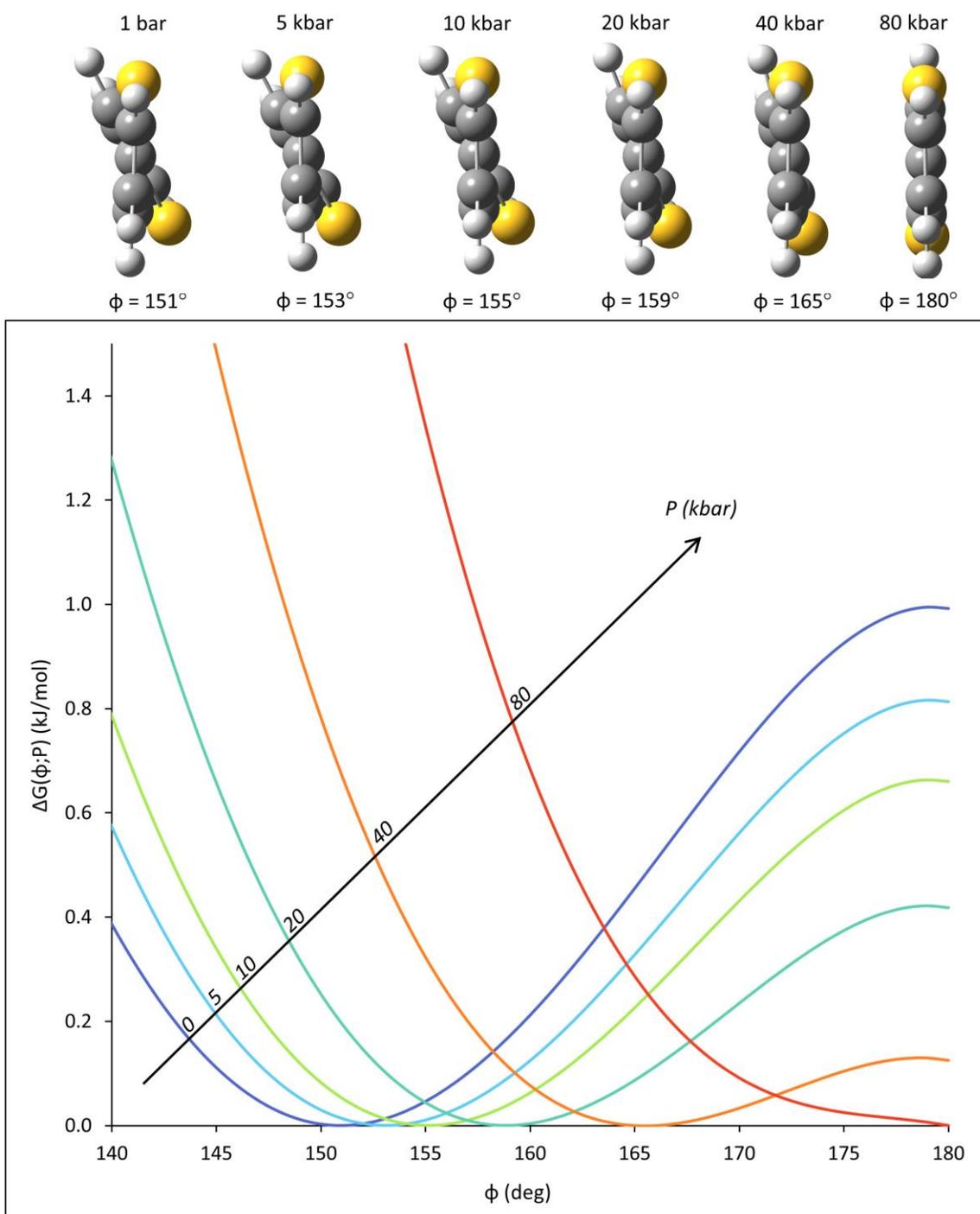
$$\Delta G(\phi; P) = \Delta G_o(\phi) + (P - P_o)\Delta V_o(\phi) \quad (3.2)$$

The  $\Delta V_o(\phi)$  for these systems were obtained using the Archimedean displacement model. The calculations were performed for systems of 1 solute and 256 toluene solvent molecules, and were run in parallel with 50 seeds for a total simulation time of 2-4  $\mu$ s. During the simulations, the  $\phi$  angle was held constant in a steep harmonic potential with a force constant of 10,000 kJ mol<sup>-1</sup> rad<sup>-2</sup>. This produced the volume profiles shown in Figure 3.4, with a quadratic function fitted to the data in order to smooth out computational errors. The three volume profiles are similar in shape and magnitude, indicating that the substituents do not have a major contribution to the volume change.

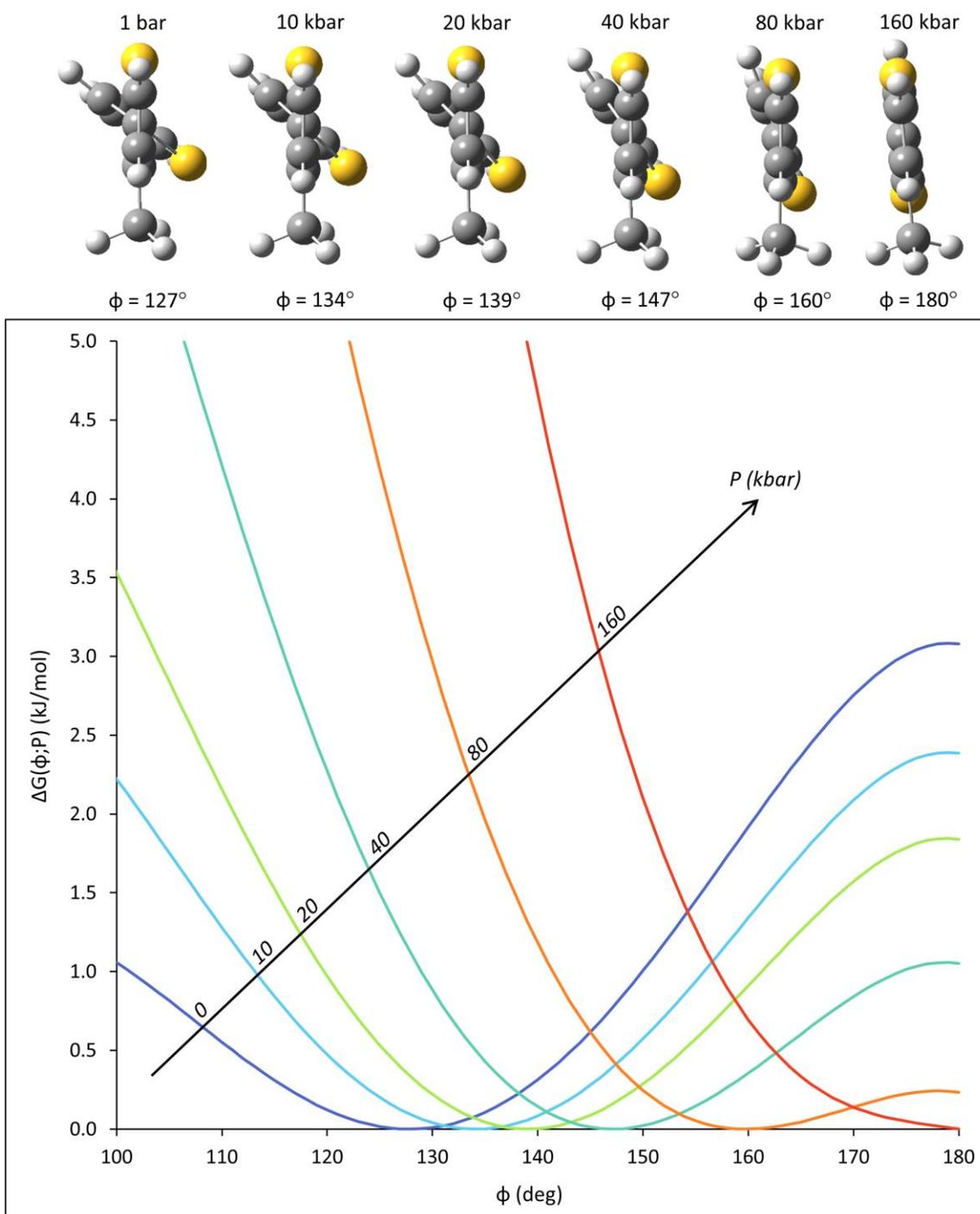
The fitted  $\Delta V_o(\phi)$  segments were combined with the corresponding  $\Delta G_o(\phi)$  of Figure 3.3 according to eq. 3.2 at a series of different pressures to obtain the high pressure  $\Delta G(\phi; P)$  shown in Figure 3.5, Figure 3.6 and Figure 3.7 for models **1**, **2** and **3** respectively. In all cases, as the pressure increases, the *trans* minimum shifts closer to 180°, eventually becoming fully planar at the highest pressure.



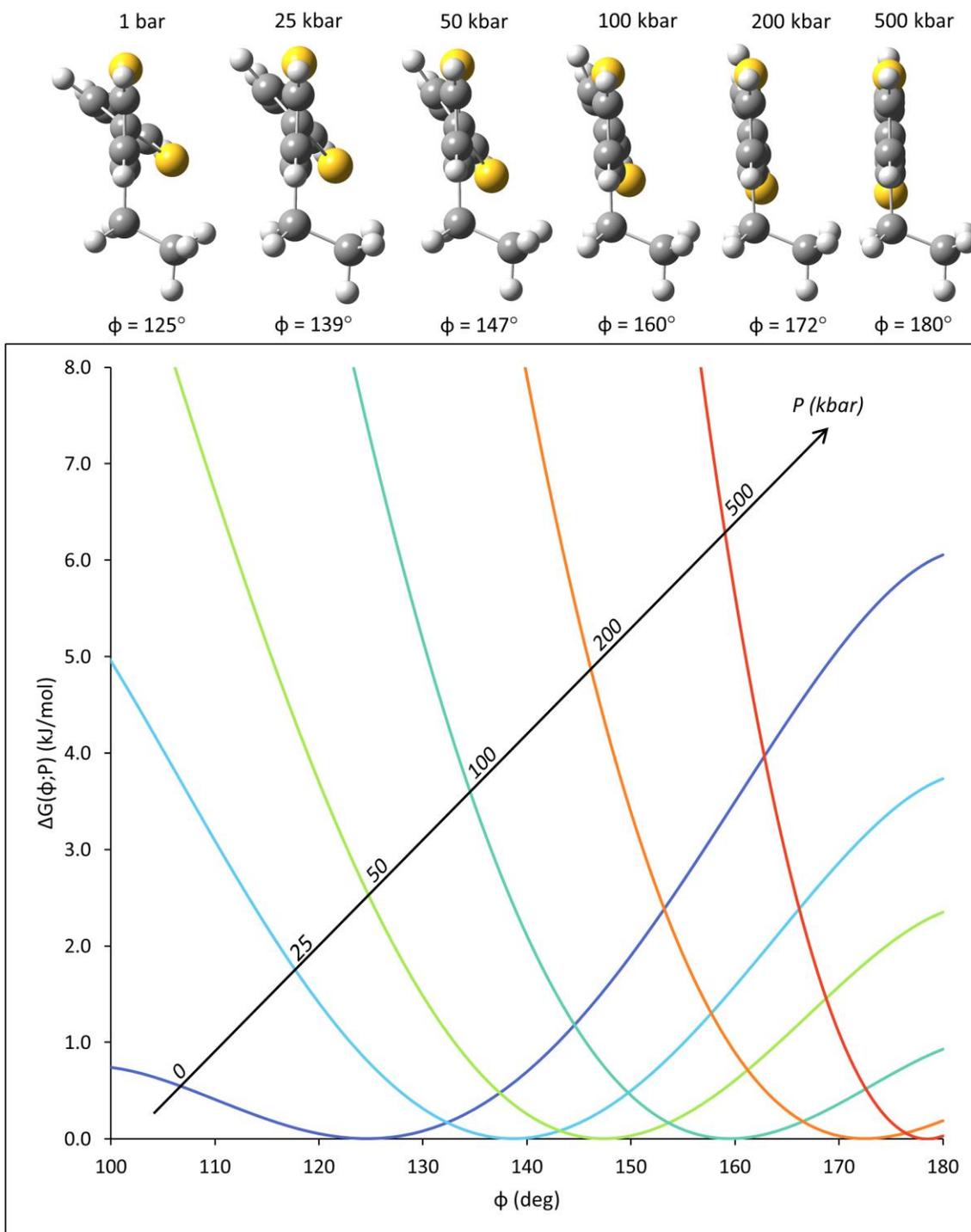
**Figure 3.4** Volume profiles  $\Delta V_o(\phi)$  for rotation around  $\phi$  in bithiophene models 1, 2, and 3. Error bars were calculated using block averaging. The portion of the volume profile around  $180^\circ$  was fitted with a quadratic function to smooth out computational errors (dotted line).



**Figure 3.5** High pressure Gibbs energy profiles  $\Delta G(\phi; P)$  for bithiophene 1. Increasing pressure to 80 kbar causes the minimum of  $\Delta G(\phi; P)$  to shift from  $151^\circ$  to  $180^\circ$ .



**Figure 3.6** High pressure Gibbs energy profiles  $\Delta G(\phi; P)$  for bithiophene 2. A pressure of 160 kbar is required to shift the minimum of  $\Delta G(\phi; P)$  from  $128^\circ$  to  $180^\circ$ .



**Figure 3.7** High pressure Gibbs energy profiles  $\Delta G(\phi; P)$  for bithiophene 3. The increase in the Gibbs energy barrier at  $180^\circ$  in  $\Delta G_o(\phi)$  coupled with the relatively small volume change from model 2 means that model 3 requires a high pressure of 500 kbar to drive the system into its planar conformation.

The change in the electronic properties of the models as a function of pressure was assessed by calculating the change in the HOMO-LUMO gap for the twisted ambient pressure structure and the planar high pressure structure. The results are shown in Table 3.1, with a comparison made for a similar conformational change in butadiene. These data suggest that increasing pressure, which favours planar conformations, enhances the level of  $\pi$ -conjugation and thus modifies the electronic properties of substituted thiophene dimers.

**Table 3.1** The change in the electronic properties of the models assessed by change in HOMO-LUMO energy gap  $\Delta\Delta E_{HOMO/LUMO}$  between the conformation at ambient pressure ( $\phi_o$ ) and the high pressure structure ( $\phi = 180^\circ$ ). For reference, the  $\Delta\Delta E_{HOMO/LUMO}$  for the same structural change in butadiene is provided. Energies were calculated at the B3LYP/6-31++G(d,p) level.

Model	$\phi_o$ (deg)	$\Delta\Delta E_{HOMO/LUMO}$ model (kJ/mol)	$\Delta\Delta E_{HOMO/LUMO}$ butadiene (kJ/mol)
1	151	-20.3	-6.9
2	128	-52.9	-28.6
3	125	-62.0	-33.8

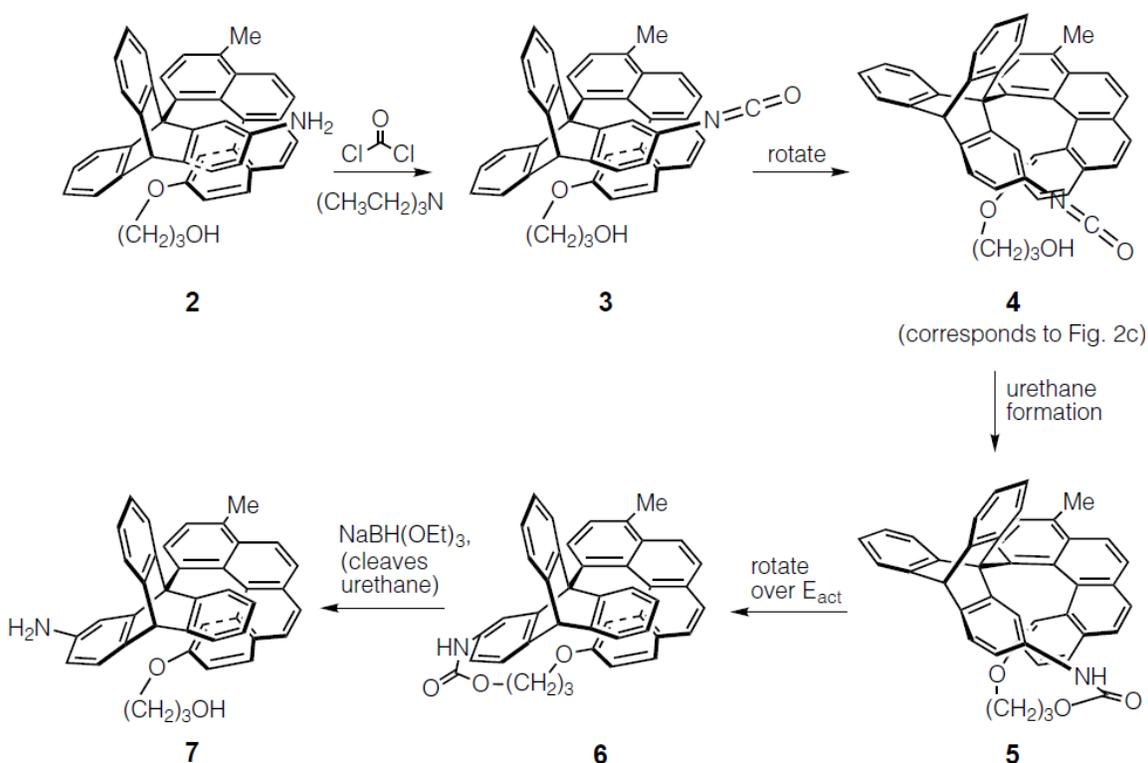
Admittedly, the required pressure of 500 kbar for model **3** is too high for practical use. The pressure required to drive the system into its planar conformation depends both on the magnitude of the reference pressure Gibbs energy barrier, and the difference in volume between the reference pressure minimum and the  $180^\circ$  state. For model **3**, the increase in steric interaction between the larger ethyl substituent and the neighbouring ring resulted in the barrier height doubling relative to model **2** with almost no change in the volume profile. Therefore, the required pressure to overcome the barrier was greatly increased. To reduce the pressure requirement, the system must be modified to either reduce the barrier, or increase the difference in volume between the planar and non-planar states.

### 3.3. Molecular rotors

Molecular rotors are another type of molecular machine which have received considerable attention in the literature.<sup>157</sup> Unidirectional rotors, which sustain rotation in a single direction, are of particular interest due to their potential application as synthetic motors for nanoscale machines.<sup>148–151,157</sup> These molecules typically consist of a bulky aromatic “rotor” which revolves around a stationary axle. Current prototypes for

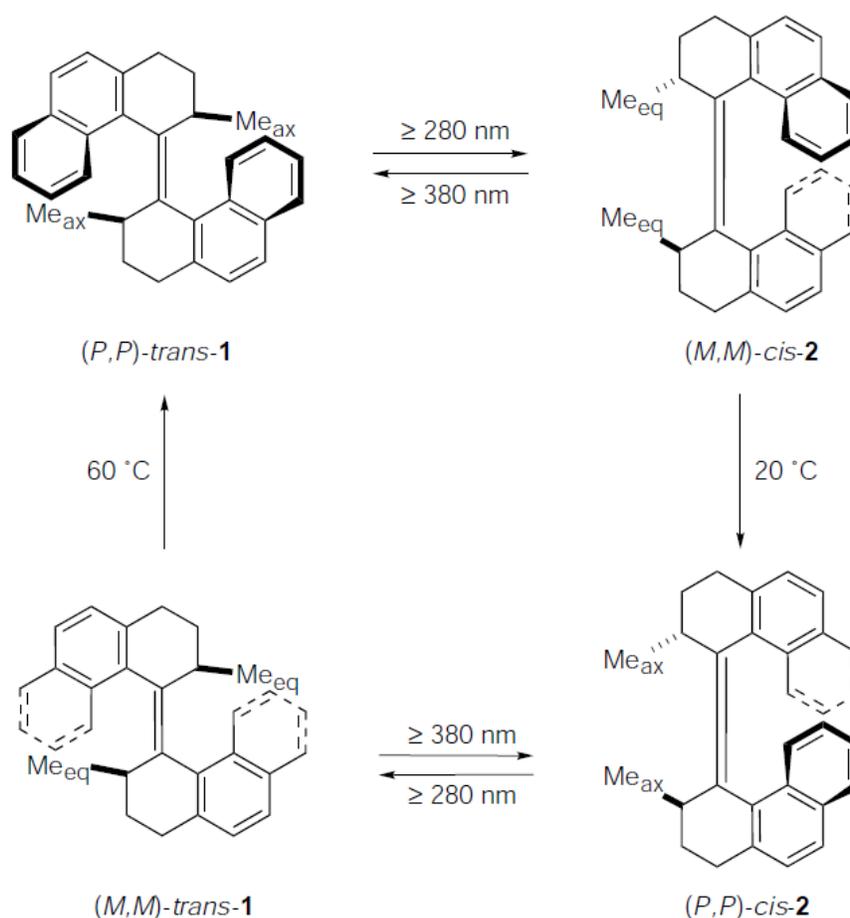
unidirectional molecular rotors fall into two classes: chemically-driven and photo-driven. The first two examples in each category were published in the same issue of Nature in 1999.<sup>168,169</sup>

Drawing inspiration from biological rotors such as the  $F_1$  subunit of ATP synthase<sup>170</sup> and bacterial flagella,<sup>171</sup> Kelly and coworkers proposed and synthesized the first chemically-driven rotor.<sup>168,172</sup> In this system, a series of chemical modifications provided the driving force for one 120° rotation of the three-bladed rotor (Figure 3.8). Unfortunately, the high reaction barriers (~100 kJ/mol) and the need for a different chemical environment for each step resulted in very slow rotation, and repeated rotation beyond the initial 120° was not possible. Improvements to this initial design<sup>172</sup> along with alternative models<sup>173–175</sup> have been proposed, but all suffer from similar drawbacks of long reaction times or reliance on the chemical environment.



**Figure 3.8** The first synthetic chemically-driven rotor.<sup>168</sup> This system relies on a series of chemical modifications to the amine group to promote rotation of the three bladed propeller 120° in the clockwise direction. Reprinted by permission from Springer Nature: Nature, Unidirectional rotary motion in a molecular system, T. Ross Kelly, Harshani De Silva and Richard A. Silva. © 1999

Photo-driven rotors, on the other hand, promise a much faster rotation than chemically-driven rotors. Feringa and coworkers proposed the first photo-driven rotor in 1999<sup>169</sup> (Figure 3.9). This rotor consists of two bulky aromatic groups connected by a carbon-carbon double bond. Crucially, this structure also possesses axial chirality as well as two stereogenic centers. Irradiation of the rotor with UV light allows for isomerization around the central bond, and the fast relaxation of the methyl groups from the sterically-hindered equatorial position into the more favourable axial position prevents rotation in the reverse direction. This initial model has since been significantly improved upon, and more complex rotors have been designed that follow the same basic cycle.<sup>176,177</sup> However, repeated exposure to high energy UV light can lead to photodegradation, and the long term stability of these rotors has not been studied.<sup>157</sup>

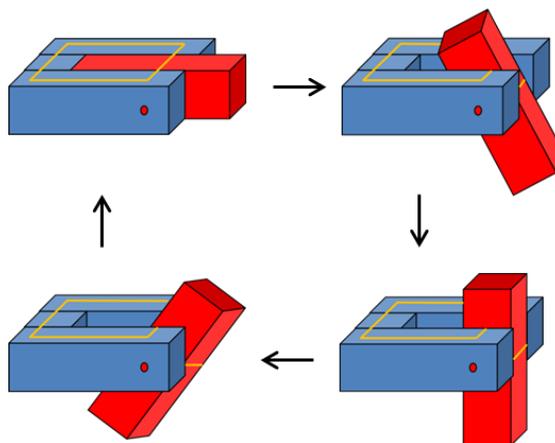


**Figure 3.9** The first photo-driven rotor.<sup>169</sup> Irradiation results in isomerization around the central double bond, and steric interactions promote rotation in a single direction. Reprinted by permission from Springer Nature: Nature, Light-driven monodirectional molecular rotor, Nagatoshi Koumura, Robert W. J. Zijlstra, Richard A. van Delden, Nobuyuki Harada & Ben L. Feringa. © 1999

Through its ability to continuously tune molecular conformations, pressure seems to offer an alternative mechanism of sustaining perpetual unidirectional motion in a molecular rotor system. Pressure has the possible advantage of providing a faster rotation than the chemically-driven systems, as the pressure effect on the reaction profiles is almost instantaneous and a change in chemical environment is not necessary. In addition, since pressure is sensitive only to the change in volume of the system, the use of pressure may allow for the design of rotors possessing a wider variety of molecular structures and functional groups, as the requirement of specific chemical interactions or photosensitivity is removed.

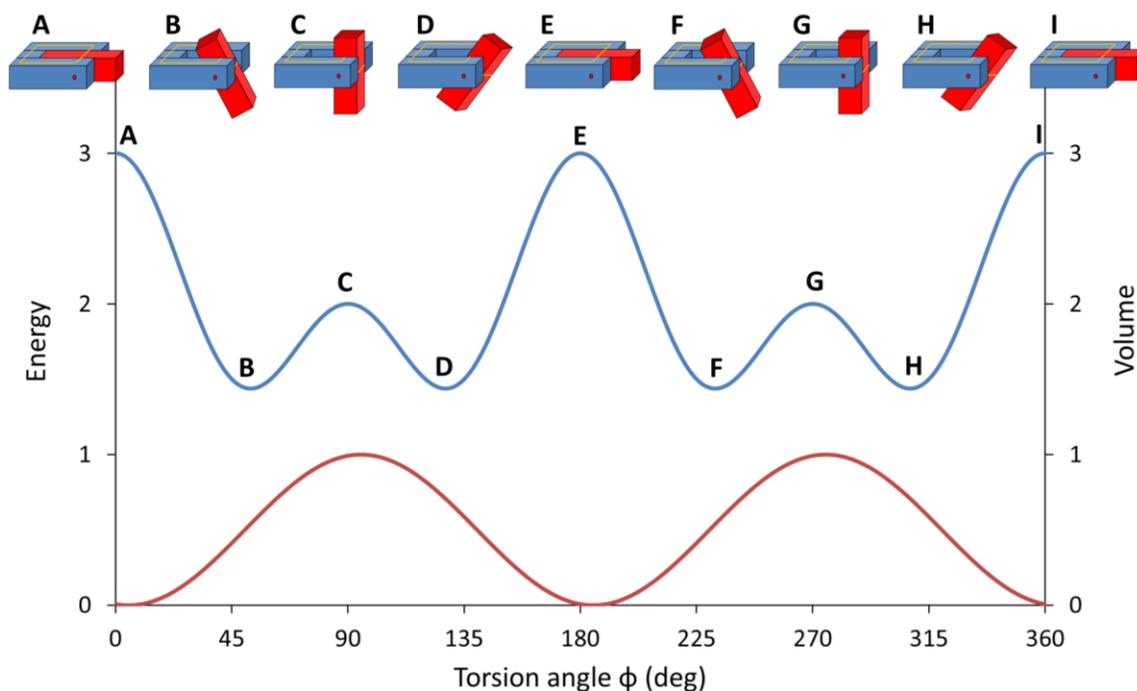
### 3.3.1. Mathematical model of pressure-controlled rotor

To demonstrate the feasibility of a pressure-controlled molecular rotor, we first considered a hypothetical prototype system. The ability of pressure to control molecular rotors is based on eq. 3.2, and thus the shapes of the  $\Delta G_o(\phi)$  and  $\Delta V_o(\phi)$  profiles will determine the behaviour of the rotor. As pressure increases, the  $(P - P_o)\Delta V_o(\phi)$  term will eventually dominate over  $\Delta G_o(\phi)$  and ultimately the minimum of  $\Delta G(\phi; P)$  will correspond to the minimum of volume. Alternatively, if high negative pressure is applied instead of positive pressure, the maximum of volume will be preferred as the negative  $(P - P_o)\Delta V_o(\phi)$  term overtakes the  $\Delta G_o(\phi)$  term. Therefore, based on the arrangement of extrema in  $\Delta G_o(\phi)$  and  $\Delta V_o(\phi)$ , a properly designed pressure pulse should be able to bias the movement of a  $\Delta G(\phi; P)$  minimum in a single direction.

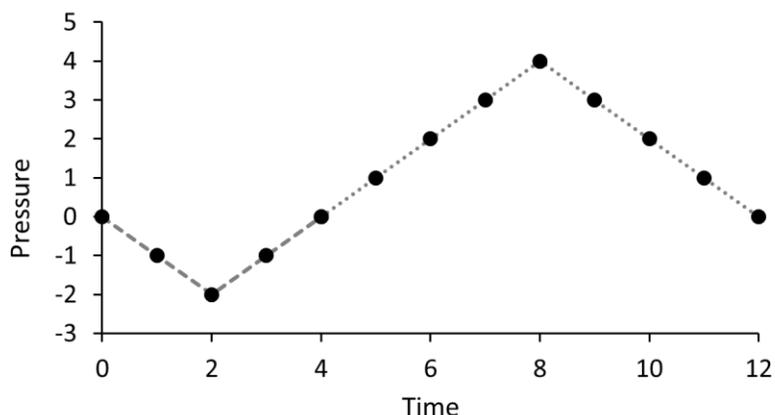


**Figure 3.10** A schematic of a hypothetical pressure-driven molecular rotor with energy and volume profiles shown in Figure 3.11. The mobile part (red) rotates within the stationary strap (blue) by a properly designed pressure pulse (see text).

One possible design for a unidirectional rotor is shown in Figure 3.10, with  $\Delta V_o(\phi)$  and  $\Delta G_o(\phi)$  profiles in Figure 3.11. The rotor consists of a mobile part (red) which rotates inside of a stationary strap (blue). The angle  $\phi$  between the mobile and stationary parts of the rotor serves as a reaction coordinate for the process. The system is designed so that it is the most compact when it is flat, thus giving  $\Delta V_o(\phi)$  minima near  $\phi$  of  $0^\circ$  (**A**),  $180^\circ$  (**E**), and  $360^\circ$  (**I**). Due to steric repulsion between the static and moving parts of the rotor, the  $\Delta G_o(\phi)$  function has maxima near those points. Its additional maxima at  $90^\circ$  (**C**), and  $270^\circ$  (**G**) are due to disruption of a different type of interaction favoring planar conformations, for example,  $\pi$ -electron conjugation or specific chemical interactions of properly positioned functional groups. As a result, the minimum energy conformations (**B**, **D**, **F**, **H**) are twisted, as in the case of the bithiophene models above.



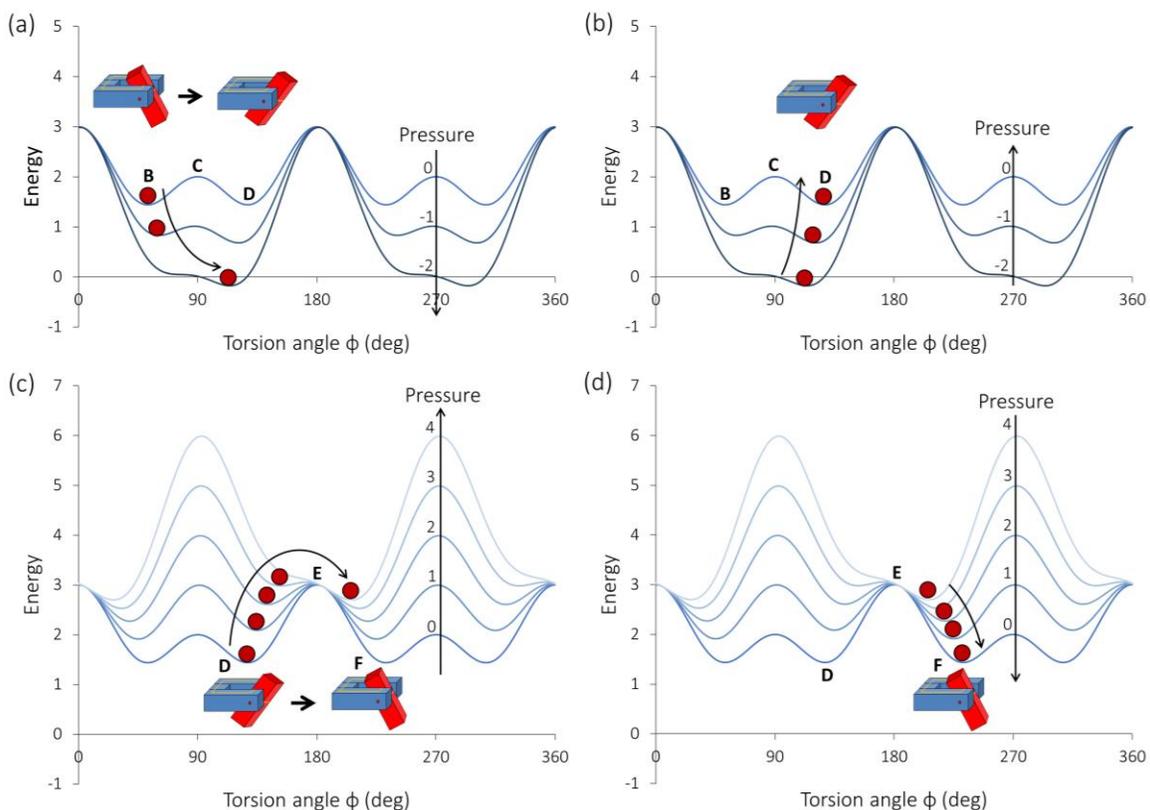
**Figure 3.11** A hypothetical molecular rotor system described by a four-fold ambient Gibbs energy profile  $\Delta G_o(\phi)$  (blue) and a two-fold volume profile  $\Delta V_o(\phi)$  (red). For convenience, both  $\Delta G_o(\phi)$  and  $\Delta V_o(\phi)$  are dimensionless. Labels A to I serve to relate conformations of the rotor to the respective maxima and minima of  $\Delta G_o(\phi)$ .



**Figure 3.12** The periodic two stage pressure pulse sequence. The first part of the pulse (dashed line) is in the negative direction, followed by a positive pressure pulse (dotted line).

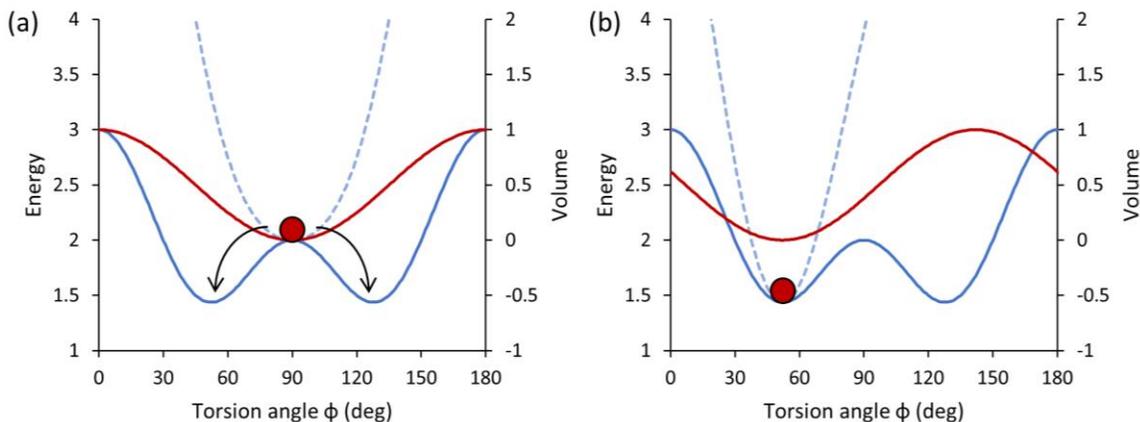
A periodic pressure pulse sequence designed to propel this molecular rotor system in the clockwise direction is shown in Figure 3.12. This pulse sequence can be separated into two stages: a negative downward pressure pulse followed by restoration back to ambient conditions, and a positive upward pressure pulse followed by restoration of ambient pressure.

The general schematic explaining the mechanism of the propelling effect is presented in Figure 3.13, where the red circles label the current configuration of the system at a given time. Initially, the system is assumed to be in its leftmost minimum **B**, separated from the next minimum **D** by maximum **C** at  $\phi = 90^\circ$ . With application of negative part of the pressure pulse, the increasingly negative  $(P - P_o)\Delta V_o(\phi)$  term in eq. 3.2 biases  $\Delta G(\phi; P)$  towards the **D** state, eventually eliminating barrier **C** and leaving **D** the only minimum between  $0^\circ$  and  $180^\circ$  (Figure 3.13a). The system is now in minimum **D** and the remaining part of the negative pressure pulse restores the ambient pressure without any further significant conformational changes (Figure 3.13b). The subsequent positive pressure pulse pushes the system to minimum **F** by destabilizing minimum **D** to push it over barrier **E**. (Figure 3.13c). As with the negative part of the pressure pulse, the restoration of ambient pressure does not result in significant conformational changes and the system remains in minimum **F** (Figure 3.13d). After this two-pulse pressure cycle, the system has been shifted continuously and in a clockwise manner from minimum **B** to its periodically related minimum **F**. Further two pulses of the pressure cycle will rotate the system a full  $360^\circ$ .



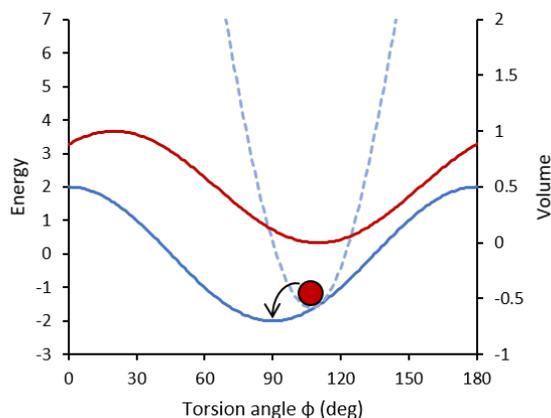
**Figure 3.13** Progression of the pressure-dependent Gibbs energy profiles  $\Delta G(\phi; P)$  for the hypothetical molecular rotor utilising the two stage pressure pulse of Figure 3.12. The red circles trace the shift in the position of the system with pressure. The conformational changes of the rotor for each pressure pulse are illustrated near their respective minima.

For this design, the extrema of  $\Delta G_o(\phi)$  and  $\Delta V_o(\phi)$  were slightly shifted relative to each other. If the extrema of  $\Delta G_o(\phi)$  and  $\Delta V_o(\phi)$  were exactly matching, application of the pressure pulse would result in either no movement of  $\Delta G(\phi; P)$  or a 50% chance of motion in the opposite direction. As illustrated in Figure 3.14(a), if the maximum of  $\Delta G_o(\phi)$  was aligned with the minimum of  $\Delta V_o(\phi)$ , it would result in the high pressure  $\Delta G(\phi; P)$  minimum corresponding with the ambient pressure  $\Delta G_o(\phi)$  maximum and so upon release of pressure the system has a 50% chance of moving forward and a 50% chance of moving backwards. In the other case where the minimum of  $\Delta G_o(\phi)$  is aligned with the minimum of  $\Delta V_o(\phi)$ , there is no change in the conformation of the system at high pressure as shown in Figure 3.14(b).



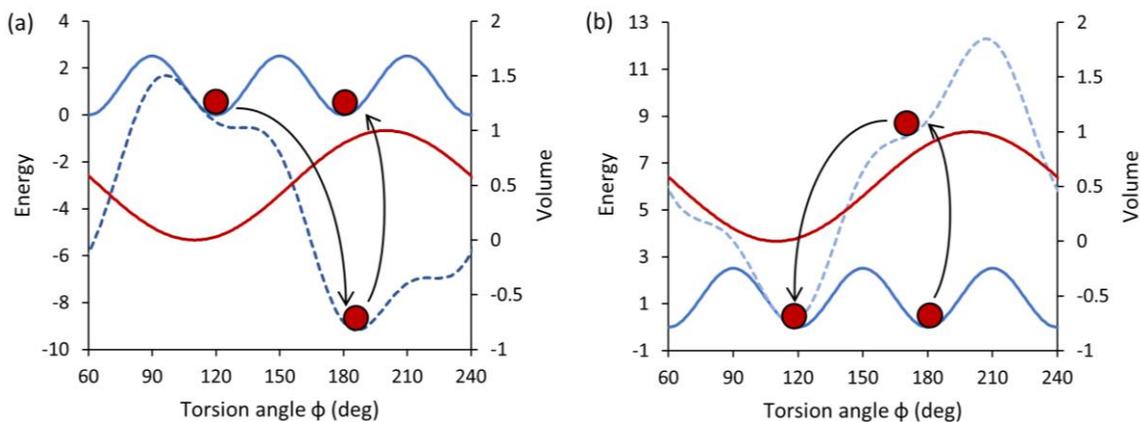
**Figure 3.14** An exact alignment of extrema in  $\Delta V_o(\phi)$  (red) and  $\Delta G_o(\phi)$  (blue solid) produces unwanted behaviour in the high pressure  $\Delta G(\phi; P)$  (blue dashed). (a) The maximum of  $\Delta G_o(\phi)$  aligned with the minimum of  $\Delta V_o(\phi)$  results in a 50% chance of reverse rotation. (b) The alignment of the minima of both profiles results in no change in the conformation of the system at high pressure.

In addition, in this model the  $\Delta G_o(\phi)$  is a four-fold profile (two maxima and two minima), while  $\Delta V_o(\phi)$  is two-fold (one maximum and one minimum), *i.e.* a 2:1 ratio. To test if there are other ratios that could also result in sustained rotation, we applied the pressure pulse of Figure 3.12 to two other models. The first possessed a 1:1 ratio, and the result is shown in Figure 3.15. Application of pressure to this model resulted in no net rotation of the system, as the high pressure minimum is located before the  $\Delta G_o(\phi)$  maximum.



**Figure 3.15** A 1:1 ratio between  $\Delta G_o(\phi)$  (blue solid) and  $\Delta V_o(\phi)$  (red) results in insufficient movement of the high pressure  $\Delta G(\phi; P)$  minimum (blue dashed). Therefore, upon release of pressure, the system returns to the previous minimum of  $\Delta G_o(\phi)$ .

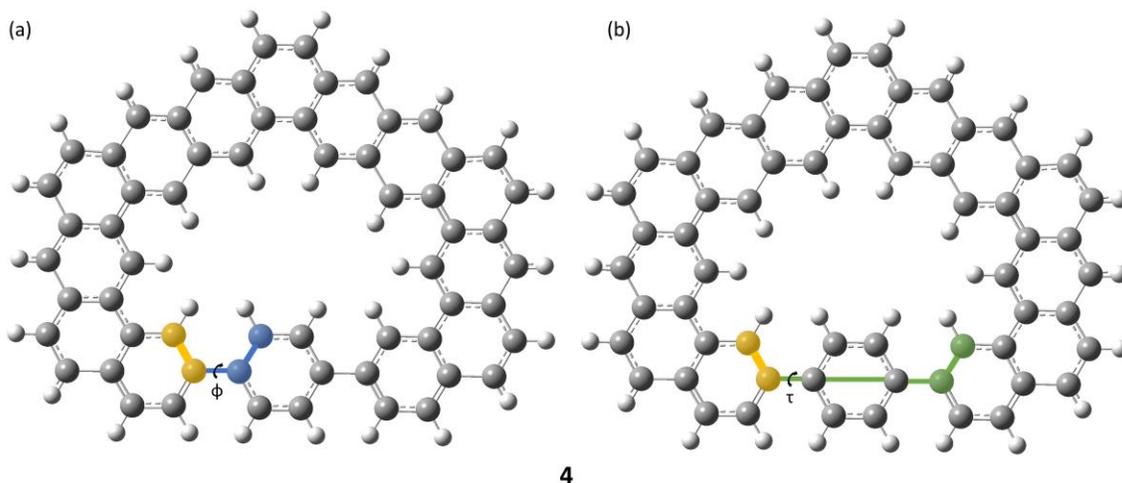
We also tested a 3:1 ratio (Figure 3.16), which resulted in the system becoming trapped in a loop upon application of the pressure pulse, cycling between the minima at 120° and 180°. In both of these cases, the failure of the model is due to the system encountering an extremum of  $\Delta V_o(\phi)$  before passing the maximum of  $\Delta G_o(\phi)$ . Therefore, we can conclude that for unidirectional pressure-driven motion, the extrema of  $\Delta V_o(\phi)$  must be slightly “ahead”, or at a larger value of  $\phi$  than the maxima of  $\Delta G_o(\phi)$



**Figure 3.16** A 3:1 ratio between  $\Delta G_o(\phi)$  (blue solid) and  $\Delta V_o(\phi)$  (red) results in the system becoming stuck between 120° and 180°. (a) A negative pressure pulse (dark blue dashed) eliminates the barrier at 150° and causes the system to fall into the 180° minimum, where it remains. (b) The next positive stage of the pressure pulse (light blue dashed) results in the 180° minimum becoming destabilized, and the system falls back into the 120° minimum.

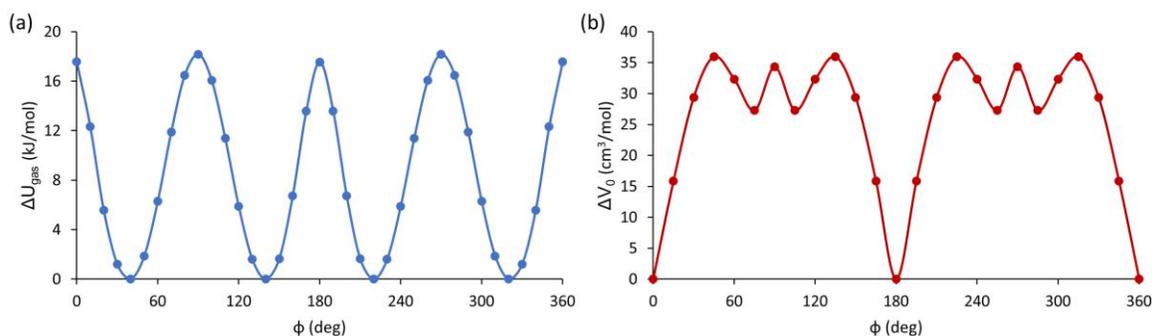
### 3.3.2. Energy and volume profiles of possible molecular candidates

As a next step, we attempted to design a real pressure-controlled unidirectional rotor, with properties based on the model system described above. The rotor **4**, shown in Figure 3.17, has a design directly inspired by the hypothetical model and consists of a benzene ring rotating inside of a pericondensed strap. For rotation, the reaction coordinate is considered to be the angle  $\phi$  between the benzene ring and the edge of the strap as shown in Figure 3.17(a), while the mutual orientation of the two sides of the strap is fixed by restraining the angle  $\tau$  to 0°, as illustrated in Figure 3.17(b).



**Figure 3.17** A real molecular rotor **4** designed based on the hypothetical system of section 3.3.1. Torsion angles  $\phi$  and  $\tau$  are shown in blue and green respectively. Both angles use the two yellow carbons as reference.

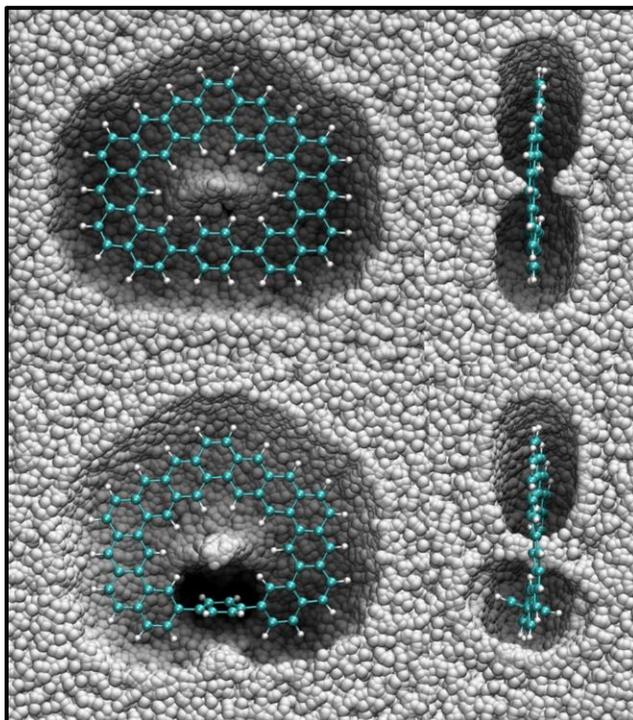
The B3LYP/6-31++G(d,p) potential energy profile for rotation in **4** (Figure 3.18a) shows 180° periodicity and barriers at  $\phi = 0^\circ$  and  $90^\circ$ , matching the hypothetical model. Steric interaction between the hydrogens of the benzene ring and those of the strap are the cause of the barrier at  $0^\circ$ , and the barrier at  $90^\circ$  is due to loss of conjugation between the benzene ring and the pericondensed strap. The minima are at the twisted conformations of  $\phi = 40^\circ$  and  $140^\circ$ . Since this system is nonpolar with relatively large energy barriers, the contribution of solvation to  $\Delta G_o(\phi)$  was assumed to be negligible.



**Figure 3.18** (a) The gas phase potential energy profile  $\Delta U_{gas}(\phi)$  of **4**, showing 2-fold periodicity as with the hypothetical rotor. (b) Volume profile  $\Delta V_o(\phi)$  for **4** obtained using 1000 Lennard-Jones particles as solvent with  $\sigma = 0.35$  nm. Error bars are too small to be seen on this scale.

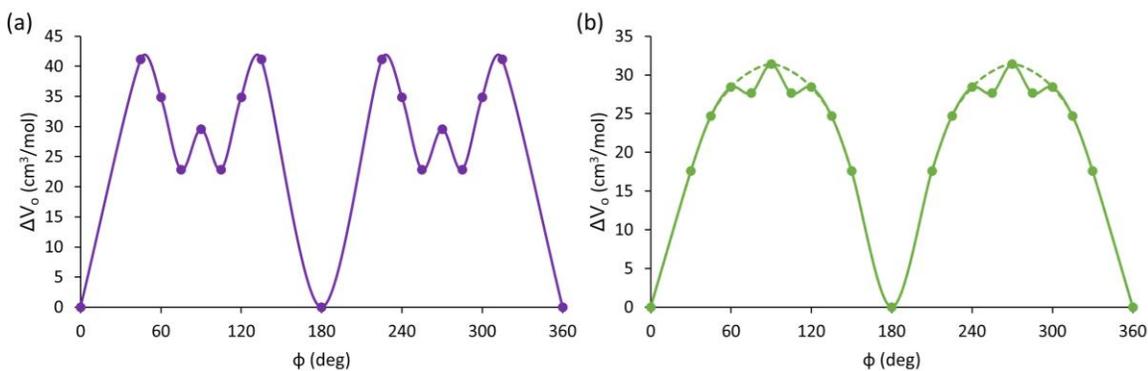
The volume profile for **4** was obtained using the Archimedean displacement model with a simple Lennard-Jones solvent ( $\sigma = 0.35$  nm,  $\varepsilon = 2.0$  kJ/mol, 1000 solvent:1 solute). The profile is shown in Figure 3.18(b), and displays an increase in volume as the ring

becomes orthogonal to the strap, as in the hypothetical model. However, this system also demonstrates a small decrease in volume immediately prior to and after  $90^\circ$ . The cavity images in Figure 3.19 provide an explanation for this decrease in volume – as the ring rotates, the void volume between the ring and the strap becomes large enough for solvent to penetrate, which results in an overall reduction in the volume of the rotor. Due to these dips in volume, application of pressure pulses like one shown in Figure 3.12 will not result in unidirectional motion.



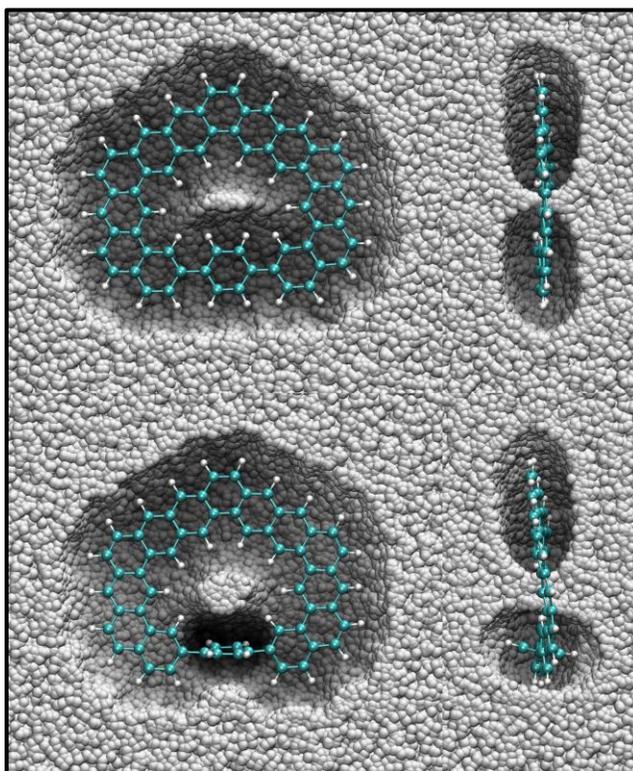
**Figure 3.19** Cavity images for rotor 4 with a  $\sigma = 0.35$  nm solvent: Top – flat conformation ( $\phi = 0^\circ$ ); bottom – orthogonal conformation ( $\phi = 90^\circ$ ); left – front view; right – side view.

To further investigate this unwanted behavior of the volume profile around  $\phi = 90^\circ$ , the  $\sigma$  parameter of the Lennard-Jones potential was varied to imitate smaller- or larger-size solvent. An increase in solvent size to  $\sigma = 0.45$  nm (Figure 3.20a) exacerbated the void volume effect, producing an even larger decrease in volume around  $\phi = 90^\circ$ . On the other hand, a decrease in solvent size to  $\sigma = 0.25$  nm (Figure 3.20b) resulted in the desired smoothing of the volume profile toward a single maximum.



**Figure 3.20** Volume profiles  $\Delta V_o(\phi)$  for **4** obtained with (a)  $\sigma = 0.45$  nm solvent and (b) with  $\sigma = 0.25$  nm solvent. The larger solvent exacerbates the unwanted volume dip near  $\phi = 90^\circ$  whereas the smaller solvent smooths  $\Delta V_o(\phi)$  towards a single maximum at  $\phi = 90^\circ$  (the dashed line, showing the general shape of the “ideal” profile of Figure 3.11).

Examination of the cavity images for  $\sigma = 0.25$  nm in Figure 3.21 shows that this may be due to the fact that the solvent particles are now small enough to penetrate between the strap and the benzene ring at all conformations.



**Figure 3.21** Cavity images for rotor **4** with a  $\sigma = 0.25$  nm solvent: Top – flat conformation ( $\phi = 0^\circ$ ); bottom – orthogonal conformation ( $\phi = 90^\circ$ ); left – front view; right – side view.

As the extrema of  $\Delta V_o(\phi)$  and  $\Delta G_o(\phi)$  are currently aligned, further modification must be made to the rotor prototype in order to shift these profiles relative to each other, as discussed in section 3.3.1. This could be achieved by the addition of chemical interactions between the ring and the strap that would stabilize or destabilize certain conformations, or by a change in size or shape of the rotating ring to shift the extrema of  $\Delta V_o(\phi)$ . In addition, a  $\Delta G_o(\phi)$  energy barrier of 18 kJ/mol with a  $\Delta V_o(\phi)$  maximum of 31 cm<sup>3</sup>/mol would require that  $P - P_o = -5.8$  kbar in order to overcome the first barrier. From a practical point of view, this magnitude of negative pressure cannot be produced due to the inevitable formation of vapour bubbles within the liquid solvent. Therefore, the  $P - P_o$  requirement must be reduced in order to produce a functioning unidirectional pressure-driven rotor. This could be accomplished by either reducing the  $\Delta G_o(\phi)$  energy barrier or by increasing the volume change in  $\Delta V_o(\phi)$ . While this rotor is a good proof of concept, it is clear that further investigations are necessary to construct a functioning pressure-driven unidirectional rotor.

### 3.4. Conclusions

The power of pressure to control molecular machines has been demonstrated for the examples of bithiophene molecular wires and for a model unidirectional molecular rotor. The key to this approach is the fact that applied pressure modifies the Gibbs energy profile to favour the lowest volume conformation at high pressures, or the highest volume conformation at negative pressures. With this in mind, systems can be designed with useful properties that can be tuned with pressure.

The application of high pressures to the three bithiophene models resulted in a change in all models from a twisted conformation to a planar conformation, which in turn resulted in an increase in the  $\pi$ -electron conjugation, as evidenced by the decrease in the HOMO-LUMO gaps. Even though the three systems tested are small models, they show that pressure control of electronic properties is possible, and a larger system could be designed which would change their conductivity, colour or fluorescence with the application of pressure.

The idea of pressure-controlled conformation was taken one step further, and a sequence of positive and negative pressure pulses was designed to continuously propel a properly constructed molecular system through a series of conformational changes

adding up to a continuous unidirectional rotation. A model rotor was proposed and tested, and further work must be undertaken to introduce a proper shift between extrema of the profiles.

## Chapter 4.

# Theoretical volume profiles as tools for probing transition states and transition state ensembles<sup>§</sup>

### 4.1. The concept

Transition states (TSs) play a central role in theories of rate processes<sup>9,53,178,179</sup> as gatekeeper species, controlling the saddle point passages between the reactant and product domains. For elementary unimolecular processes, TS theory relates the rate constant  $k$  to the activation Gibbs energy  $\Delta G^\ddagger$  by the Eyring equation (eq. 4.1), where  $T$  is the temperature and  $k_B$ ,  $h$ , and  $R$  are the Boltzmann, Planck, and gas constants, respectively.

$$k = \frac{k_B T}{h} e^{-\frac{\Delta G^\ddagger}{RT}} \quad (4.1)$$

For conformationally flexible and other non-rigid systems, where reactive transformations can be achieved through a variety of parallel pathways, eq. 4.1 remains valid if the overall rate constant  $k$  is taken to be the sum of partial rate constants  $k_i$ , enumerated over all possible reaction pathways. The  $\Delta G^\ddagger$  value must therefore be re-interpreted as a weighted average of Gibbs activation energies  $\Delta G_i^\ddagger$  of individual TSs constituting a TS ensemble (TSE).<sup>180–182</sup> Although identification of an isolated TS is a routine task easily accomplished by standard QM methods,<sup>183–187</sup> enumeration of all important TSs constituting a TSE for a non-rigid system is a far less trivial problem requiring a substantial computational effort,<sup>188–193</sup> and quickly becoming a daunting problem even for relatively small systems.<sup>194–200</sup>

---

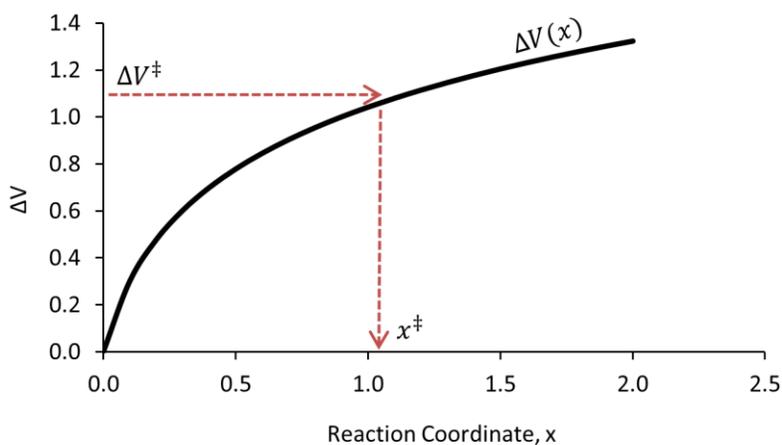
<sup>§</sup> The work presented in this chapter is adapted from two publications: Wiebe, H.; Weinberg, N. *J. Chem. Phys.* **2014**, *140* (12), 124105, wherein the volume profile method is formulated and applied to the model chain system of section 4.2, and Wiebe, H.; Louwse, M.; Weinberg, N. *J. Chem. Phys.* **2017**, *146* (10), 104107, wherein the method is applied to the paracyclophane system of section 4.3. All calculations and analysis are contributed by H. Wiebe, except for the volume and Gibbs energy profiles of section 4.3 which were obtained by undergraduate student M. Louwse under the supervision of H. Wiebe.

Volume provides us with an alternative way to identify the TS, as the activation volume  $\Delta V^\ddagger$  is an indirect measurement of the volume of the TS. As discussed in section 1.2.3, the Archimedean displacement model allows us to calculate the volume of any transient structure and thus a full volume profile,  $\Delta V(x)$ , which completely describes how the volume of a reaction system varies along reaction coordinate  $x$  (eq. 4.2).

$$\Delta V(x) = V(x) - V_R \quad (4.2)$$

If the  $\Delta V^\ddagger$  has been experimentally determined for a reaction system with an unknown TS, then the position  $x^\ddagger$  of the TS along the reaction coordinate can be determined by locating  $\Delta V^\ddagger$  on the theoretically calculated profile using eq. 4.3. This concept is illustrated in Figure 4.1 for a hypothetical volume profile.

$$\Delta V(x^\ddagger) = \Delta V^\ddagger \quad (4.3)$$



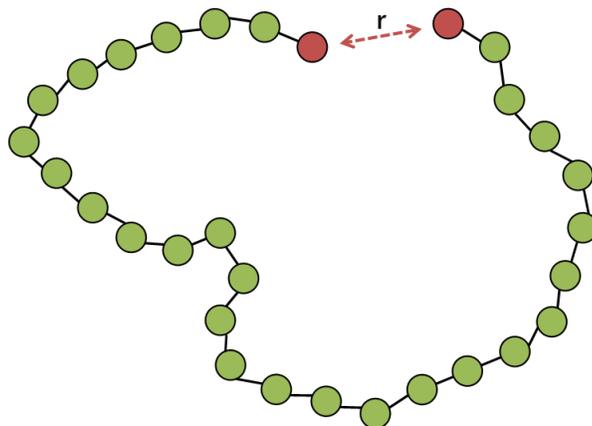
**Figure 4.1** An illustration of the use of eq. 4.3. Combination of the experimentally determined  $\Delta V^\ddagger$  and the MD-calculated  $\Delta V(x)$  yields the location of the unknown TS.

This approach promises an exceptional computational efficiency, as it only requires simulation of individual structures, not full reactive dynamics, and thus the timescale of the reaction is unimportant. As a proof of concept, the volume profile method is used here to identify the TSE for unfolding of a model flexible chain<sup>201</sup> and for internal rotation of benzene ring in 1,12-dimethoxy-[12]-paracyclophane.<sup>202</sup>

## 4.2. Model flexible chain as a prototype for protein folding

### 4.2.1. System

As an initial test of the volume profile method, we required a system with well controlled kinetic parameters, an easily defined reaction coordinate, and a predictable TSE. As such, we chose to use a model system consisting of a 30-unit chain with strongly interacting terminal atoms. Similar systems have previously been used to model conformational dynamics in proteins and other polymers.<sup>203–207</sup> Since the terminal atoms were chosen to be the only atoms to display a strong through-space interaction, the choice of reaction coordinate for the folding process was straight-forward and it could clearly be identified as the distance  $r$  between the terminal atoms (Figure 4.2).



**Figure 4.2.** 30 unit model chain. The distance,  $r$ , between terminal atoms serves as the reaction coordinate for the unfolding process.

The interaction between terminal atoms was described by a Morse potential with  $D = 10 \text{ kJ mol}^{-1}$ ,  $\beta = 20 \text{ nm}^{-1}$ ,  $r_o = 0.440 \text{ nm}$ , modified to reach zero at a finite separation  $r = \rho < 0.9 \text{ nm}$  between termini (eq. 4.4). Parameters  $\rho = 0.800 \text{ nm}$  and  $k = 10.94 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  were fitted to provide a smooth transition at the junction point of  $0.7 \text{ nm}$ .

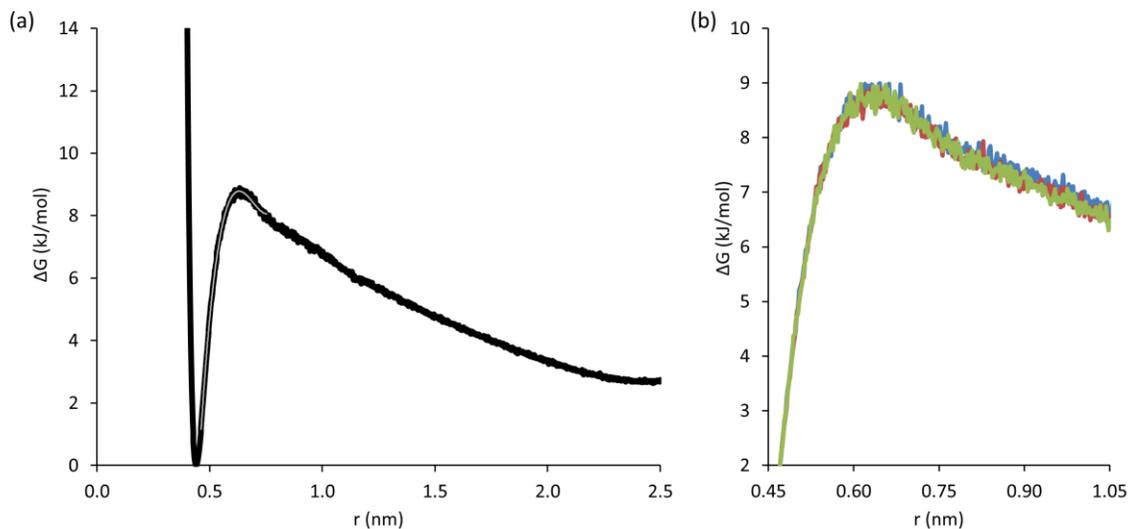
$$U(r) = \begin{cases} D[1 - \exp(-\beta(r - r_o))]^2 - D & \text{for } r < 0.7 \\ -k(r - \rho)^2 & \text{for } 0.7 \leq r \leq \rho \\ 0 & \text{for } r > \rho \end{cases} \quad (4.4)$$

For all other interactions within the model chain, the bonded and non-bonded force field parameters were taken from the literature, wherein these parameters were successfully used to model the kinetics of conformational changes in *n*-butane.<sup>208</sup> The solvent was a Lennard-Jones liquid consisting of 1500 neutral monatomic particles with  $\sigma = 0.35$  nm and  $\varepsilon = 2.0$  kJ mol<sup>-1</sup>. Unless otherwise stated, all simulations were run at 300 K and 1 kbar pressure required for the solvent to remain liquid.

#### 4.2.2. Gibbs energy profile and expected TSE

Since this system is simple, it is feasible to directly obtain the exact location  $r^\ddagger$  of the TSE along the unfolding reaction coordinate  $r$  as the position of the maximum on the Gibbs energy profile,  $\Delta G(r)$ . This profile was calculated using eq. 4.5 from the probability distribution function,  $P(r)$ , accumulated over a sufficiently long MD trajectory. These computationally expensive calculations were performed solely to provide a benchmark against which to test our volume approach and are therefore not a required step in the process.

$$\Delta G(r) = -RT \ln(P(r)) \quad (4.5)$$



**Figure 4.3** (a) The averaged Gibbs energy profile for unfolding of the model chain (black), fitted with a cubic parabola (grey) to obtain the maximum. (b) The quality of the profile is assessed by the level of convergence between data from three MD trajectories with the initial  $r$  values of 0.44 nm (blue), 0.9 nm (red) and 1.5 nm (green).

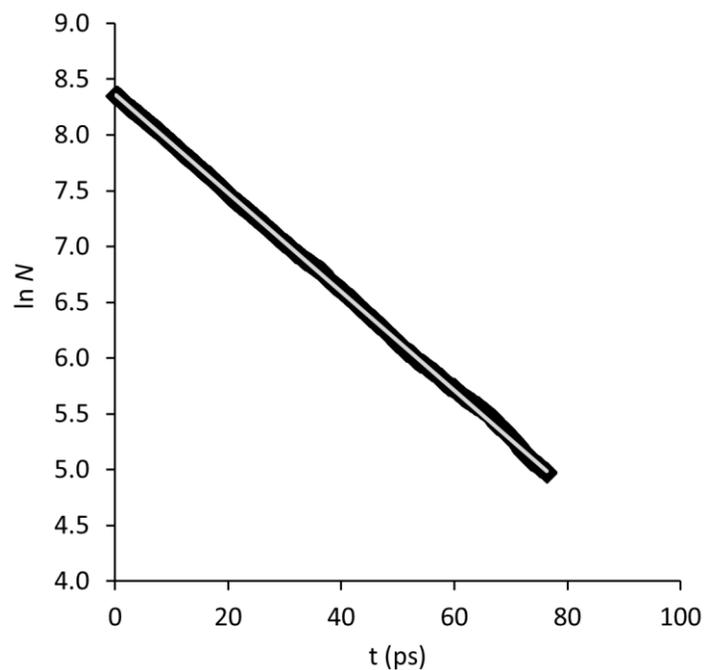
To ensure that  $\Delta G(r)$  was free of memory effects from the starting configurations, calculations were performed with three different initial values of  $r$  until the distribution functions obtained from these three trajectories converged. This required 3  $\mu$ s of simulation time for each trajectory and gave the Gibbs energy profiles shown in Figure 4.3. The final profile, obtained by averaging the three converged trajectories, had two stable minima (the folded state at  $r = 0.44$  nm and the unfolded state at  $r \approx 2.5$  nm). To remove the residual noise, the top of the barrier was approximated by a cubic parabola, locating the TSE at  $r^\ddagger = 0.638$  nm.

### 4.2.3. MD high pressure kinetics

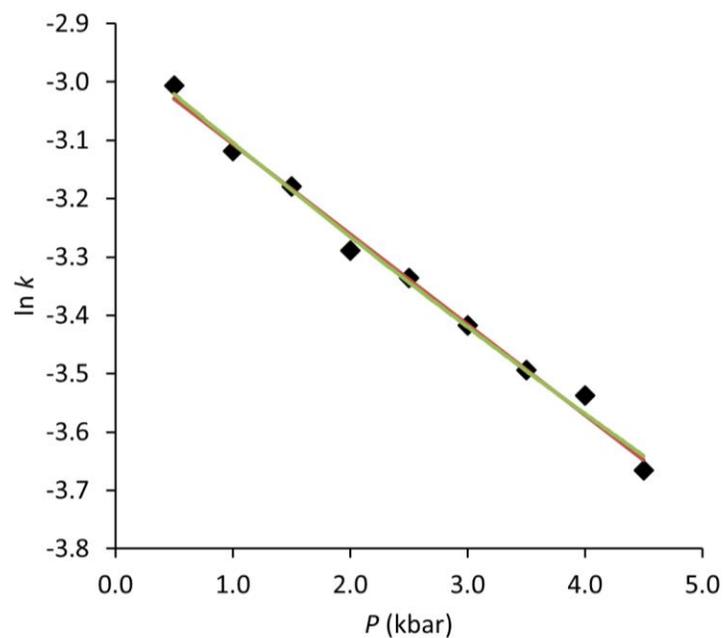
The experimental activation volume  $\Delta V^\ddagger$  is an essential component of this method. However, due to the artificial nature of our model system, an experimental value is unavailable. Therefore, we used MD simulations to mimic the kinetic experiment<sup>209</sup> and obtained the pressure-dependent rate constants  $k(P)$  necessary for determination of  $\Delta V^\ddagger$ .

In these simulations, the model chain was equilibrated with the solvent for 100 ps at a given pressure  $P$ , while confined to the folded state with a harmonic restraint. The restraint was then removed and the system was allowed to evolve in time for 1.0 ns. The reaction events were identified as the instance of the first exit from the reactant well. The simulation was repeated 3600-4200 times (depending on  $P$ ) in order to obtain good statistics. Plotting the natural logarithm of the number of unreacted trajectories  $N$  vs. time yields a first order survival probability curve, a typical example of which is shown in Figure 4.4. The pressure-dependent rate constant  $k(P)$  was obtained as the negative slope of this plot.

The process was repeated at different pressures to produce a series of pressure-dependent rate constants,  $k(P)$ . Due to the slightly curved nature of this plot, the data were approximated using both linear and quadratic regression, with fits shown in Figure 4.5. The activation volumes were found by differentiation of the fitted equations, which gave 3.83 cm<sup>3</sup>/mol for linear and 4.06 cm<sup>3</sup>/mol for quadratic fit.



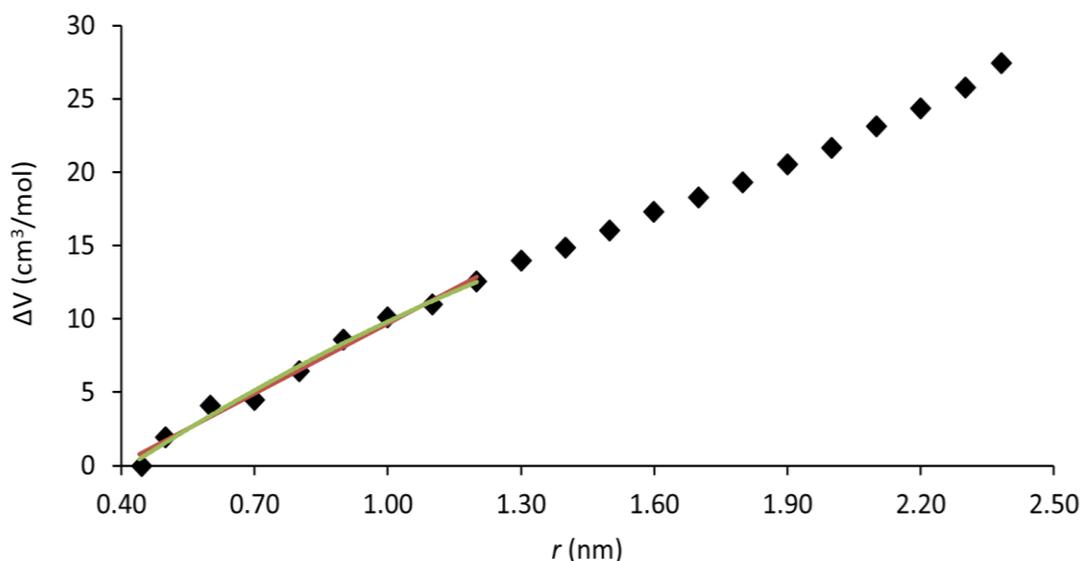
**Figure 4.4** A sample survival probability plot,  $\ln N$ , for  $r^\ddagger = 0.638$  nm and  $P = 1$  kbar.



**Figure 4.5** Fit to MD simulated high pressure kinetic data with linear (red) and quadratic (green) regression functions.

#### 4.2.4. Volume profile and predicted TSE

The volume profile  $\Delta V(r)$  for the model chain was obtained from a series of conformational ensembles along the unfolding coordinate  $r$ . These ensembles were chosen to have  $r$  values ranging from 0.5 nm to 2.4 nm with an increment of 0.1 nm. The inclusion of the reactant ensemble at  $r = 0.44$  nm yielded a total of 21 points. To maintain constant  $r$  for a given ensemble, we used a steep harmonic constraint imposed on the distance between terminal atoms. All remaining degrees of freedom of the chain were unconstrained, allowing for conformational flexibility. The volume of each ensemble was calculated using the Archimedean displacement model with 50 parallel seeds, and sufficiently long MD trajectories were generated to ensure convergence of the average volume. The resulting volume profile is displayed in Figure 4.6 and shows that the volume of the system increases along the unfolding coordinate.



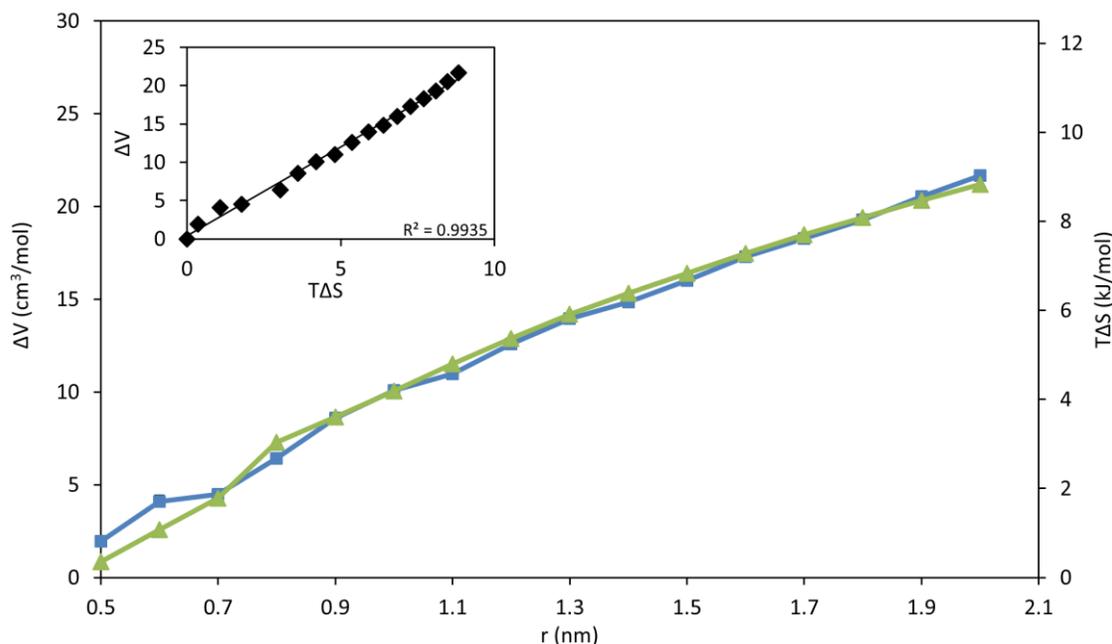
**Figure 4.6** Volume profile for the unfolding of the model chain system. The region of interest (between 0.44 nm and 1.20 nm) was fitted with both a linear (red) and quadratic (green) regression function.

The specific shape of a volume profile depends on the nature of the process as well as on strength and type of the solute-solvent interactions. For nonpolar reactions in nonpolar solvents, the volume of a reaction system correlates with its linear dimension,<sup>118</sup> as was the case for the methyl-methane system of Chapter 2. As a result, the typical values of activation volumes for bimolecular reactions, such as Diels-Alder or hydrogen transfer, are negative, whereas the activation volumes for homolytic bond cleavage

processes are positive.<sup>1,3,11–13,210,211</sup> These simple trends can be reversed by strong solvation effects, as in heterolytic bond cleavage, where the observed activation and reaction volumes are negative.<sup>1,3,11–13,210,211</sup> In the more complex cases of conformational changes in large flexible systems, such as protein unfolding, the activation and reaction volumes are negative<sup>40,41,43,44,46–51,212,213</sup> primarily due to opening of the solvent-excluded void volume and, to a lesser degree, due to solvent restructuring.<sup>34,35,37</sup> Neither of these effects is present in this relatively short nonpolar model chain immersed in a nonpolar model solvent. As a result, the volume changes associated with the unfolding of our model chain are primarily due to the geometric and entropic factors. The entropy profile for our system can be approximated by eq. 4.6, utilizing the bond potential  $U(r)$  (eq. 4.4), the calculated volume profile  $\Delta V(r)$  and Gibbs energy profile  $\Delta G(r)$ .

$$T\Delta S(r) \approx U(r) + P\Delta V(r) - \Delta G(r) \quad (4.6)$$

A strong correlation between the calculated  $\Delta V(r)$  and  $T\Delta S(r)$  shown in Figure 4.7 supports the hypothesis that the volume change of unfolding is mostly due to entropic effects.



**Figure 4.7** Correlation between volume profile  $\Delta V(r)$  (blue) and the entropy profile  $T\Delta S(r)$  (green) for the unfolding of the model chain. The inset shows a direct linear correlation between the two.

To locate the “experimental” activation volume  $\Delta V^\ddagger$  on the theoretical volume profile of Figure 4.6, we used linear and quadratic approximation to its segment between 0.44 and 1.2 nm. The values of the TSE parameter  $r^\ddagger$  thus obtained are listed in Table 4.1 and agree well with the benchmark  $r^\ddagger = 0.638$  nm obtained from the Gibbs energy profile of Figure 4.3. The results are generally better when the same level of approximation (linear or quadratic) is used to describe the volume profile  $\Delta V(r)$  and to obtain activation volumes  $\Delta V^\ddagger$  from the pressure dependences of rate constants  $k(P)$ , likely due to cancellation of matching errors. In the best case scenario, the error in recovery of  $r^\ddagger$  was as small as 2 pm, which is good even by the high standards of *ab initio* TS calculations for small systems.<sup>214,215</sup>

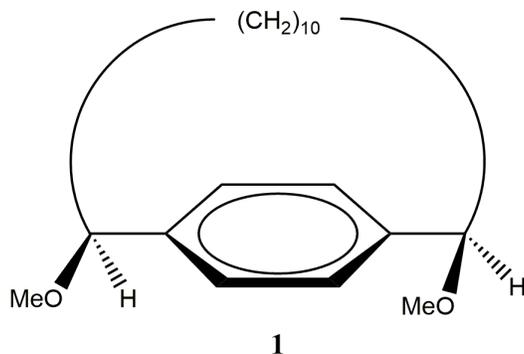
**Table 4.1** TSE positions  $r^\ddagger$  (in nm) for the unfolding of the model chain with different fitting methods. Deviations from the TSE predicted by the Gibbs energy profile are listed in parentheses.

	Linear $k(P)$	Quadratic $k(P)$
Linear $\Delta V(r)$	0.633 (-0.005)	0.650 (+0.012)
Quadratic $\Delta V(r)$	0.625 (-0.013)	0.640 (+0.002)

## 4.3. Conformational changes in paracyclophane

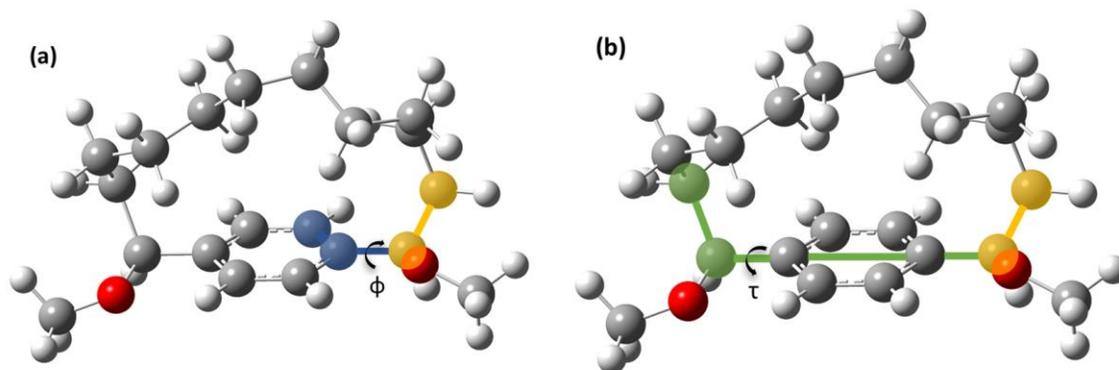
### 4.3.1. System

After a successful proof of concept in a test with a model system with a simulated  $\Delta V^\ddagger$ , we proceed here to apply this method to a real system with an experimentally-determined  $\Delta V^\ddagger$ . The chosen system is 1,12-dimethoxy-[12]-paracyclophane **1** (Figure 4.8). This molecule consists of a benzene ring, connected at the *para* positions to a flexible 12-unit methylene tether. The benzene ring can rotate within the molecule along its C2 axis, and the high pressure kinetics of this internal rotation has been studied by Yamada and coworkers.<sup>216,217</sup>

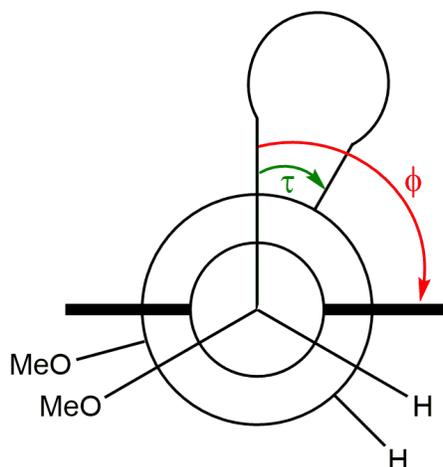


**Figure 4.8** The structure of 1,12-dimethoxy-[12]-paracyclophane.

A natural choice for a coordinate to describe the rotation of the benzene ring is the torsion angle  $\phi$ , shown in Figure 4.9(a), which describes the mutual orientation of the ring and one of the terminal segments of the tether. In addition to the movement of the ring itself, the two terminal segments may move relative to each other to accommodate the rotating ring, which is described by a second torsion angle  $\tau$ , shown in Figure 4.9(b) with a Newman projection of  $\phi$  and  $\tau$  in Figure 4.10.

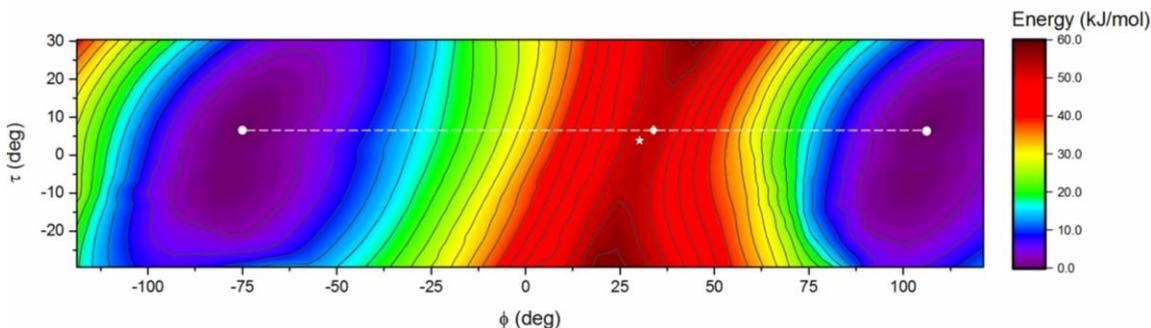


**Figure 4.9** The torsion angles that control the rotation of the benzene ring: (a)  $\phi$ , in blue, and (b)  $\tau$ , in green. Both angles are defined relative to one terminal segment of the tether, shown in yellow.



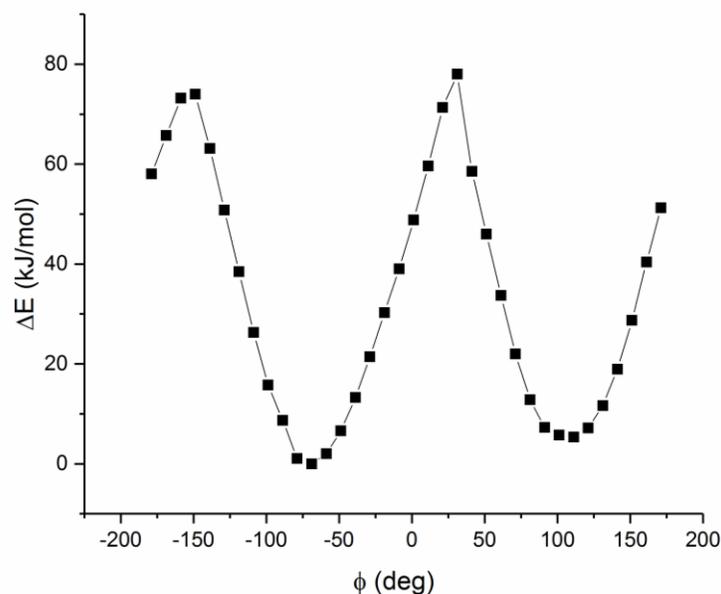
**Figure 4.10** A Newman projection showing the relationship between  $\phi$  and  $\tau$ . Benzene ring is represented as the bold line.

A scan of the potential energy surface in  $\phi$  and  $\tau$  coordinates at the semi-empirical AM1 level (Figure 4.11) shows that the reaction path is mainly controlled by  $\phi$ , with the  $\tau$  contributions being negligible. In fact, the value of  $\tau = 3^\circ$  at the TS is quite close to the value of  $\tau = 6.5^\circ$  at the minima. Therefore, we decided to restrain the value of  $\tau$  to its energy minimized value of  $6.5^\circ$  and use  $\phi$  as the sole contributor to the reaction coordinate. This reaction path, shown as the white dotted line in Figure 4.11, is a reasonable approximation to the minimum energy path.



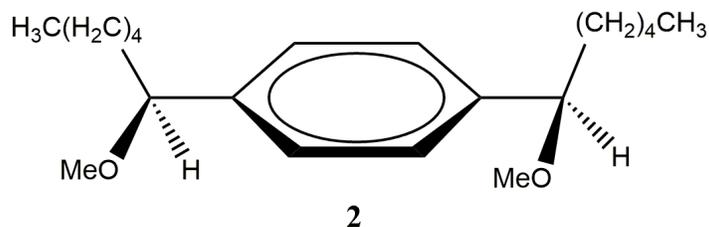
**Figure 4.11** Gas phase AM1 potential energy surface of **1** in  $\phi$  and  $\tau$  coordinates. The chosen constant  $\tau$  reaction path (dotted line) has its highest point (white diamond) at  $\phi = 34.5^\circ$  and  $\tau = 6.5^\circ$ , which is a reasonable approximation of the true TS (white star) at  $\phi = 30^\circ$  and  $\tau = 3^\circ$ .

A gas phase B3LYP/6-31++G(d,p) potential energy profile along the chosen reaction path (Figure 4.12) shows a reaction barrier of roughly 80 kJ/mol, with minima at  $-69^\circ$  and  $111^\circ$  matching those of the lower level AM1 surface.



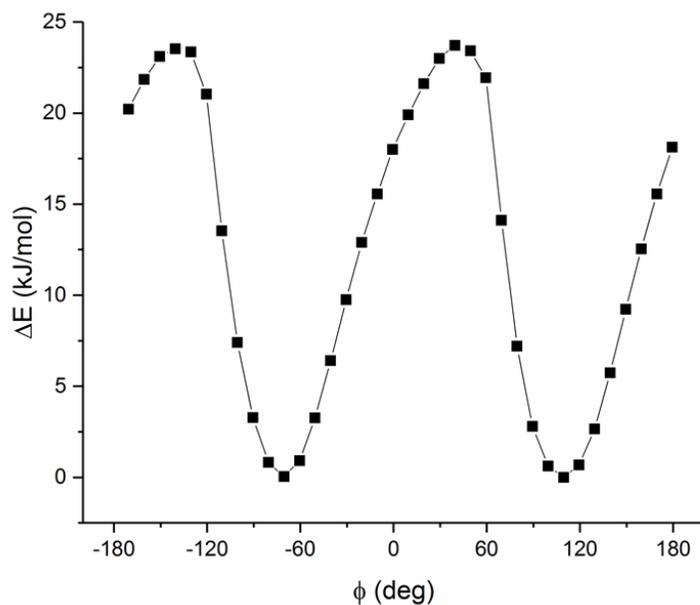
**Figure 4.12** The gas phase B3LYP/6-31++G(d,p) potential energy profile for ring rotation along  $\phi$ . The minima are at  $-70^\circ$  and  $110^\circ$ , and maxima at  $-150^\circ$  and  $30^\circ$ .

The slight deviations from perfect  $180^\circ$  periodicity can be understood as being the effect of the flexible methylene tether. As the ring rotates, it begins to interfere with the tether, causing a change in the conformation of the latter. As proof of this, we recalculated the potential energy profile for *para*-di(1'-methoxyhexyl)benzene **2** (Figure 4.13), where the tether has been severed into two independent 1'-methoxyhexyl chains.

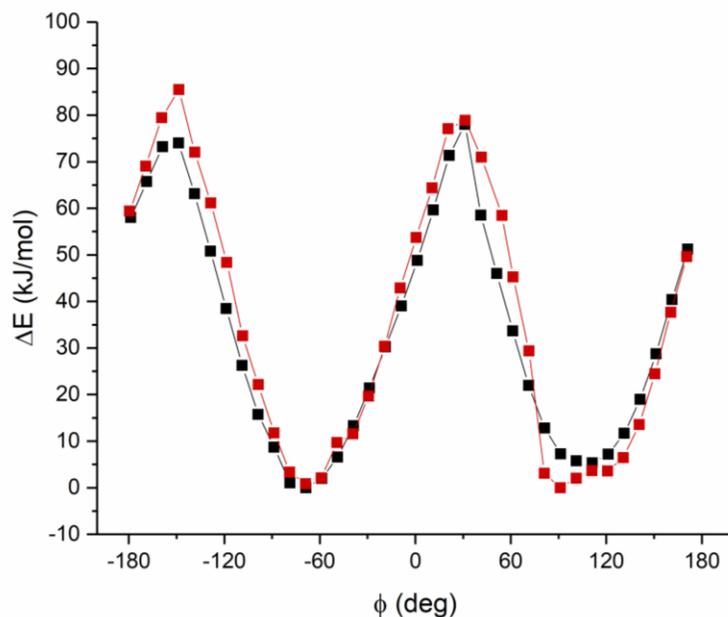


**Figure 4.13** The structure of *para*-di(1'-methoxyhexyl)benzene

As shown in Figure 4.14, the removal of the tether has the effect of both significantly reducing the reaction barrier to 25 kJ/mol and restoring perfect periodicity. This interference from the tether is consistent with the experimental observation that reduction in the size of the tether by a single  $\text{CH}_2$  unit results in a reduction of the rate by a factor of 10-1000.<sup>217</sup>



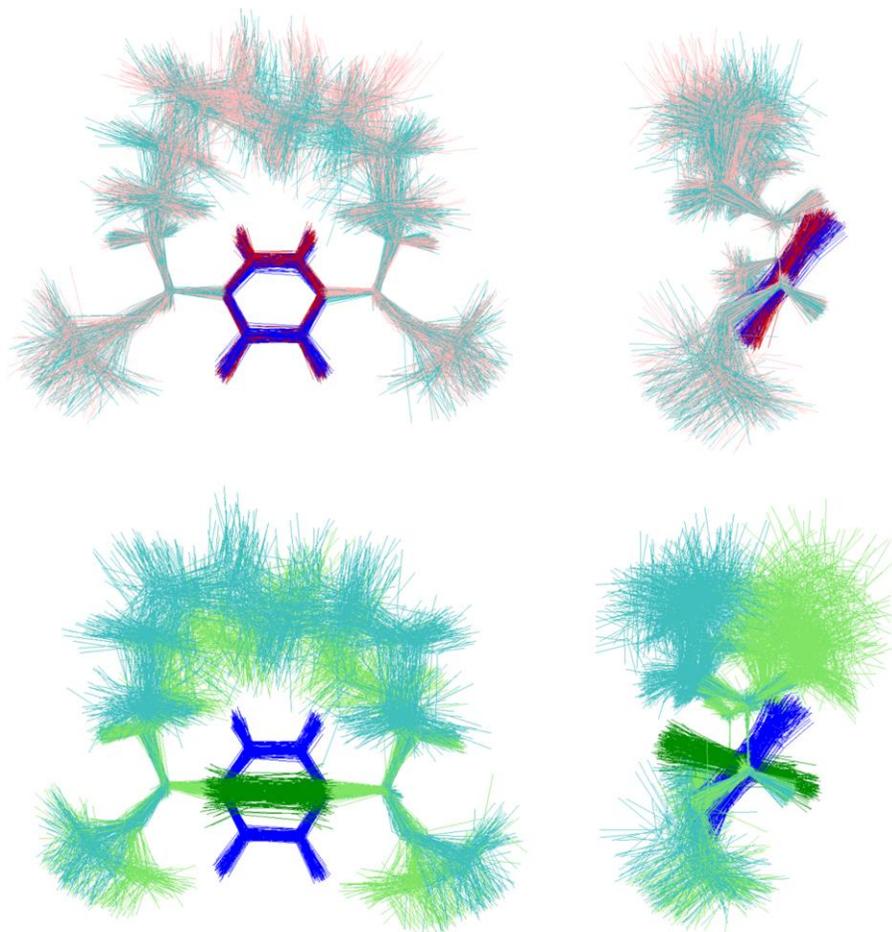
**Figure 4.14** The gas phase B3LYP/6-31++G(d,p) potential energy profile for rotation in **2**. The severing of the tether into two independent chains has restored the perfect periodicity of the profile.



**Figure 4.15** The MD-calculated potential energy profile (red) agrees well with the QM-calculated profile of Figure 4.12 (black).

To capture the effects of the dynamics of the flexible methylene tether on the potential energy profile, we calculated an MD potential energy profile in 1,1,2,2-tetrachloroethane-*d*<sub>2</sub> solvent (the same solvent as was used in the experimental work<sup>217</sup>). The resulting profile, shown in Figure 4.15, has minima, maxima and barrier heights that

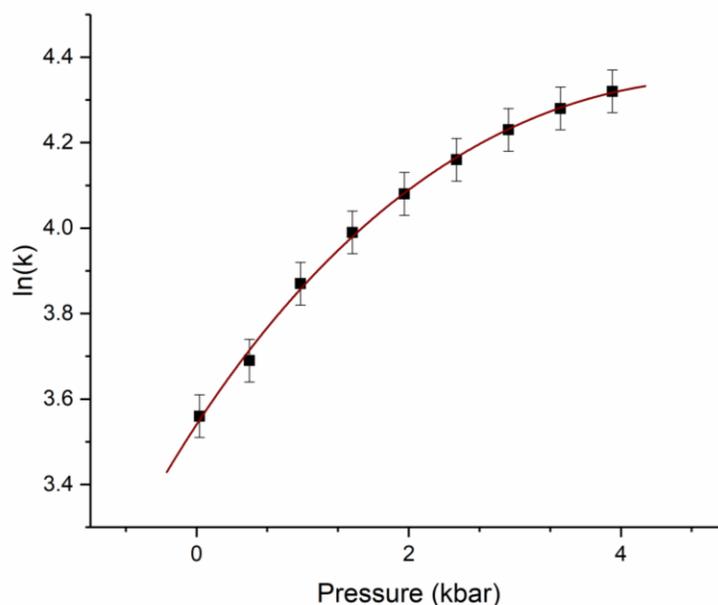
match closely with the QM profile of Figure 4.12. The periodicity is much more imperfect in the MD profile, due to the dynamics of the flexible tether, illustrated in Figure 4.16 by overlays of the instantaneous conformations of paracyclophane **1** sampled along the MD trajectories generated for constant  $\tau = 6.5^\circ$  and  $\phi = -150^\circ$  and  $30^\circ$ , two periodically separated maxima, and  $\phi = -70^\circ$ , a minimum. These conformational ensembles display different behavior of the methylene tether in response to the position of the benzene ring. As expected, the pattern is repeated with a period of  $180^\circ$ .



**Figure 4.16** The effect of the position of the ring on the conformational ensemble of the tether illustrated here for  $\phi = -150^\circ$  (red),  $-70^\circ$  (green) and  $30^\circ$  (red). The front (left) and side (right) views were generated as overlays of instantaneous conformations of **1** sampled along MD trajectories generated for constant  $\tau = 6.5^\circ$ .

### 4.3.2. Experimental high pressure kinetic data

The high pressure kinetics of internal rotation in **1** was studied by Yamada and coworkers at 406.5 K in 1,1,2,2-tetrachloroethane-*d*<sub>2</sub> solvent.<sup>217</sup> Their data, in Figure 4.17, shows a non-linear dependence of  $\ln k$  on pressure, which is quite typical of experimental high pressure kinetic data. The authors assessed a series of average activation volumes  $\Delta V_{ave}^\ddagger$  over 1 kbar ranges and obtained a value of  $\Delta V^\ddagger = -11 \text{ cm}^3/\text{mol}$  for the 0-1 kbar range. However, this value is not appropriate for our purposes, as it more adequately describes the value of  $\Delta V^\ddagger$  closer to the midpoint of 0.5 kbar, rather than the desired  $\Delta V^\ddagger$  at zero pressure.



**Figure 4.17** Experimental high pressure kinetic data (black data points) from ref. 217, fitted with eq. 4.7.

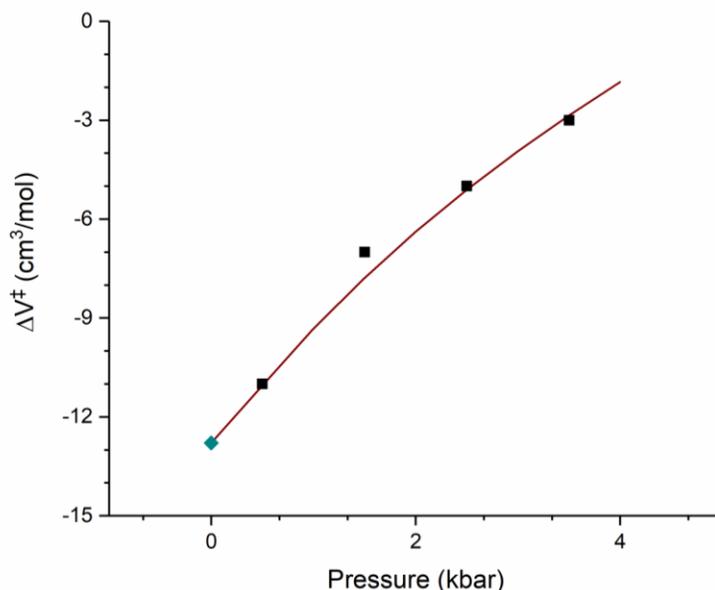
To obtain the value of  $\Delta V^\ddagger$ , it is a common practice to fit the kinetic data with one of many proposed empirical correlation functions.<sup>218,219</sup> This allows for  $\ln k(P)$  to be extrapolated to  $P = 0$  bar and the derivative can then be calculated analytically. We chose to use the El'yanov-Gonikberg correlation function<sup>139,220</sup> (eq. 4.7) with adjustable parameters  $a = 1.128$  and  $b = 0.285 \text{ kbar}^{-1}$  optimized to fit the experimental data. This equation provides a good fit to the experimental data as shown in Figure 4.17, and differentiation at  $P = 0$  bar results in a  $\Delta V^\ddagger$  value of  $-12.8 \text{ cm}^3/\text{mol}$ .

$$\ln k = \ln k_o - \frac{\Delta V^\ddagger}{RT} \left[ (1 + a)P - \left(\frac{a}{b}\right) (1 + bP) \ln(1 + bP) \right] \quad (4.7)$$

The relationship between our  $\Delta V^\ddagger$  and the  $\Delta V_{ave}^\ddagger$  calculated by the authors can be understood by differentiation of eq. 4.7 with respect to pressure, which yields an expression for  $\Delta V^\ddagger$  as a function of pressure (eq. 4.8).

$$\Delta V^\ddagger(P) = \Delta V_0^\ddagger [1 - a \ln(1 + bP)] \quad (4.8)$$

As can be seen from Figure 4.18, there is a close agreement between eq. 4.8 and the literature  $\Delta V_{ave}^\ddagger$ . This indicates that our assessment of  $\Delta V^\ddagger$  is consistent with the experimental data.

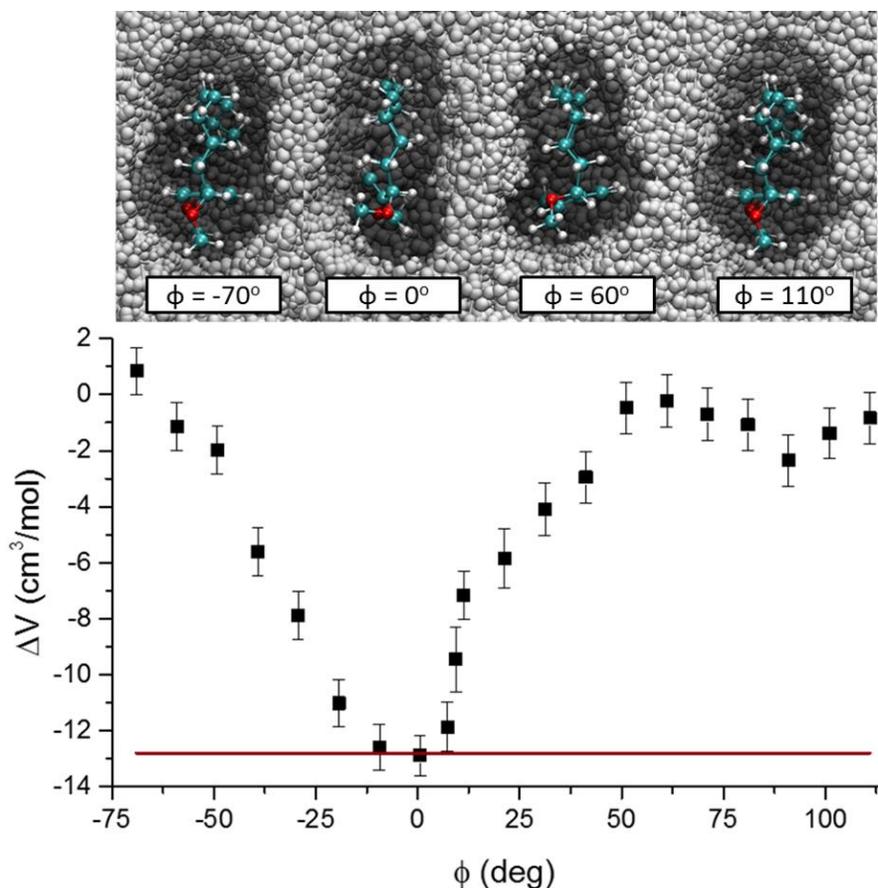


**Figure 4.18** The pressure dependence of  $\Delta V^\ddagger$  as described by eq. 4.8 (red line). Our calculated zero pressure value  $\Delta V^\ddagger = -12.8$  cm<sup>3</sup>/mol (blue diamond) constitutes a starting point. The experimentally-determined  $\Delta V_{ave}^\ddagger$  values (black squares) are consistent with the graph.

### 4.3.3. MD volume profile and predicted TSE

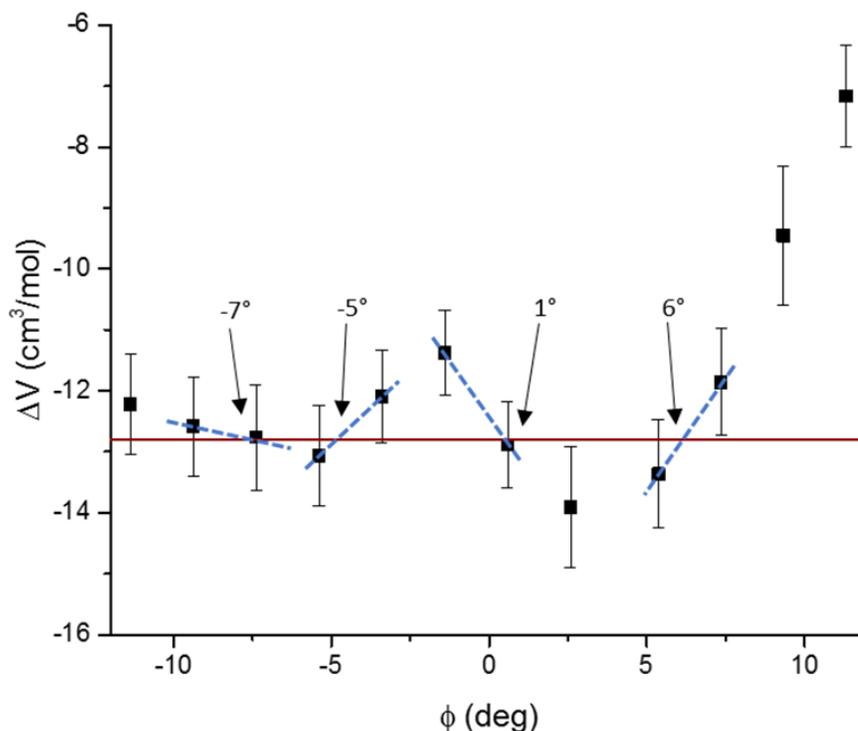
The volume profile  $\Delta V(\phi)$  for **1** was calculated using the Archimedean displacement model. To maintain constant  $\phi$  during the calculation, a harmonic restraint was imposed with force constant 10,000 kJ mol<sup>-1</sup> rad<sup>-2</sup>. A cumulative simulation time of

150 ns over 25 parallel trajectories was required to obtain sufficiently converged average volumes for each point. The resulting volume profile is shown in Figure 4.19, with MD-calculated data points in black and the experimental value of  $\Delta V^\ddagger$  shown in red.



**Figure 4.19** Volume profile  $\Delta V(\phi)$  (black data points) for one  $180^\circ$  rotation of the ring along the  $\phi$  reaction coordinate. The experimental  $\Delta V^\ddagger$  value is shown in red. Top: solvent cavities corresponding to selected values of  $\phi$ .

The intersection point is at approximately  $\phi = 0^\circ$ , but a more detailed  $\Delta V(\phi)$  is needed to locate the TSE more accurately. Therefore, more points were added to the volume profile in the transition state region between  $-11^\circ$  and  $+11^\circ$  (Figure 4.20).



**Figure 4.20** Further detail in the TSE region of the volume profile. There are four points, indicated by arrows, at which  $\Delta V(\phi)$  (black data points) appears to cross the experimental  $\Delta V^\ddagger$ . The value of  $\phi$  at the intersection points was determined using linear interpolation (blue dashed lines).

According to Figure 4.20, there are 4 points at which  $\Delta V(\phi)$  crosses  $\Delta V^\ddagger$ . Only one of these points is the true TSE, the one that possesses the highest Gibbs energy. To assist in this determination, single point Gibbs energy estimates were obtained for these points using eq. 4.9 as the sum of the average potential energy  $U_{MD}(\phi)$  and the solvation Gibbs energy  $\Delta G_{solv}(\phi)$ .  $\Delta G_{solv}(\phi)$  was calculated using thermodynamic integration. The results of this calculation are summarized in Table 4.2. From these results, we concluded that the TSE is located at  $\phi^\ddagger \approx 6^\circ$ .

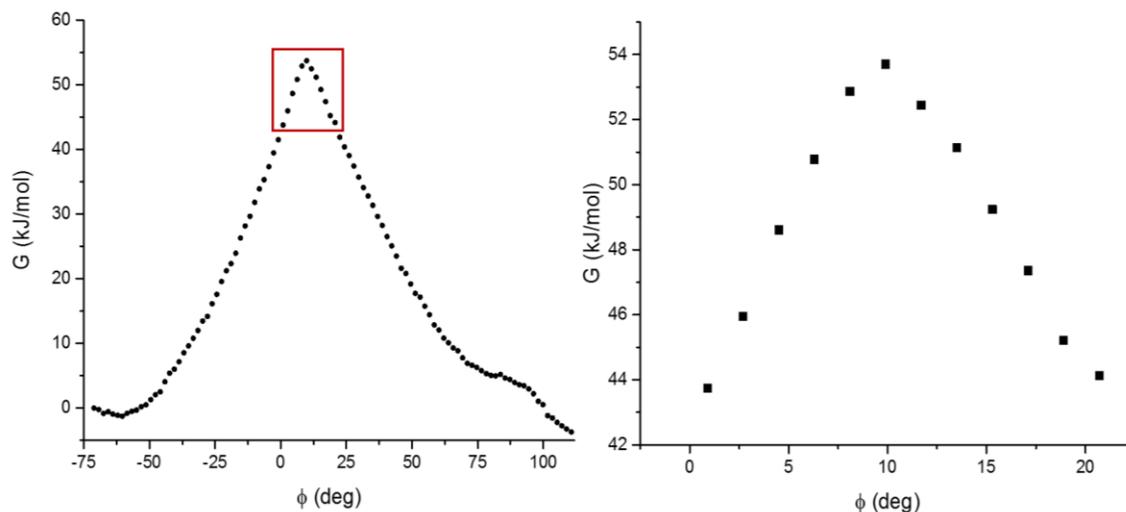
$$G(\phi) \cong U_{MD}(\phi) + \Delta G_{solv}(\phi) \quad (4.9)$$

**Table 4.2** Gibbs energies for four TSE candidates calculated using eq. 4.9.

$\phi$ (deg)	$G(\phi)$ (kJ/mol)
-7	222.9
-5	225.6
1	228.4
6	237.9

#### 4.3.4. Validation from Gibbs energy profile

For comparison, we also found the TSE for internal rotation in **1** via the Gibbs energy profile calculated using umbrella sampling along the reaction coordinate  $\phi$ . These simulations were run in triplicate for 1.0 ns with a 200 ps equilibration period. Sampling windows were placed 10° apart and a force constant of 1000 kJ/mol/rad<sup>2</sup> was used. MD simulations were run within each sampling window for 1.0 ns and the value of torsion angle  $\phi$  collected. The Weighted Histogram Analysis Method (WHAM)<sup>221</sup> was used to align the sampling windows and 100 bootstrap samples were used to determine the errors. To ensure good overlap, two additional sampling windows were placed at -171° and 90°. This gave the profile shown in Figure 4.21. The barrier region was fit with a quadratic function and the TS was determined as the maximum at  $\phi^\ddagger = 10^\circ$ .



**Figure 4.21** Gibbs energy profile for a single 180° rotation of the benzene ring (left) and its expansion (red rectangle) in the vicinity of the maximum (right). The TSE (the maximum of the profile) is located at  $\phi^\ddagger = 10^\circ$ .

The four degree difference between the values of  $\phi^\ddagger$  predicted by our method and obtained from the Gibbs energy profile is very reasonable, considering the “soft” nature of this angular coordinate and the fact that a constant  $\tau$  approximation was used in defining

the reaction path. Indeed, the experimental  $\Delta V^\ddagger$  refers to the true TSE on the Gibbs energy surface, located slightly off the linear reaction path for which the Gibbs energy profile was obtained and, according to Figure 4.11, this may result in a 4-5° difference in  $\phi^\ddagger$  values.

## 4.4. Conclusions

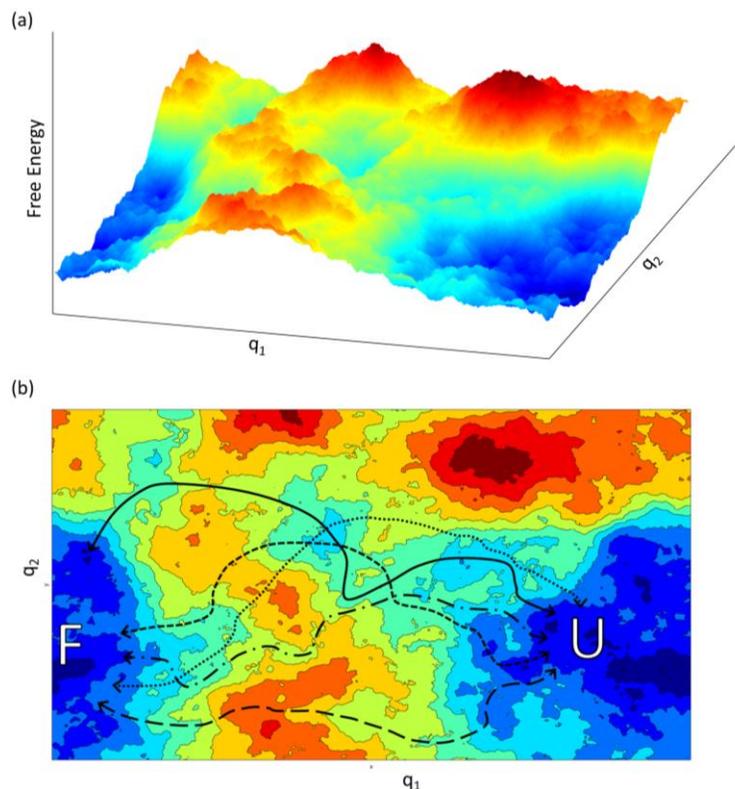
Here, the volume profile method has been shown to accurately locate the position of TSEs along a defined reaction coordinate for both a flexible model chain using simulated high pressure kinetic data, and for internal rotation in 1,12-dimethoxy-[12]-paracyclophane using experimentally-determined high pressure kinetic data. A next step would be the application of this method to larger, more complex systems for which traditional TSE searching methods fail, such as the folding of proteins. In addition to the MD-calculated volume profile  $\Delta V(x)$ , application of this method to protein folding would also require an experimentally determined activation volume  $\Delta V^\ddagger$  and a properly defined reaction coordinate  $x$ . The experimental  $\Delta V^\ddagger$  for protein folding can be obtained relatively easily using established high pressure techniques, and are in fact already available for many proteins.<sup>40,41,43,44,46–51,212,213</sup> Of course, such profiles can be used for the TSE identification only if the experimental values of activation volumes of protein folding can be reproduced sufficiently well by MD simulations. Our experience with calculations of  $\Delta V^\ddagger$  for relatively simple chemical reactions was positive; the calculated values matched experimental activation volumes very closely and the solvent effects were also well reproduced.<sup>83,111</sup> The remaining component of the proposed protocol – identification of a proper reaction coordinate – presents a significant challenge and will be explored further in the next chapter.

## Chapter 5.

# Reaction coordinates for systems with multiple reaction pathways

### 5.1. Introduction

Proteins are large, flexible polymers consisting of several hundred to several thousand atoms. The folding process, by which the protein transitions from a random coil to its characteristic three-dimensional structure, occurs via the breakage and formation of through-space contacts between amino acid residues, such as hydrogen bonds and electrostatic interactions. Due to the high dimensionality of these systems, combined with the relatively low strength of the individual interactions, this transformation occurs through a multiplicity of possible reaction pathways on a highly rugged free energy surface, as illustrated in Figure 5.1.



**Figure 5.1** (a) Schematic free energy surface for the folding of a protein, showing an extremely rugged landscape. (b) Transformations between the folded (F) and unfolded (U) domains can occur through a number of different reaction pathways.

Despite this apparent complexity, the experimentally observed kinetics of protein folding is often found to be first order and the process is frequently described by a two-state model:<sup>28</sup> reactant  $\rightarrow$  TS  $\rightarrow$  product. For this, the reactant and product are considered as domains in conformational space consisting of many local minima separated by relatively low barriers, allowing fast equilibrium between them. In the same way, the TS is reinterpreted to be a TS ensemble (TSE), consisting of various possible TS structures controlling the rates of the partial processes along individual reaction pathways.<sup>180,181,222–224</sup> In this context, a reaction coordinate is a one-dimensional parameter that collectively describes the overall process. Accordingly, the reaction coordinate must satisfy two conditions: it should clearly differentiate between reactant and product domains, and it should also produce a Gibbs energy profile with a kinetically relevant activation barrier at the TSE, where all TS structures are clustered over a narrow range of the coordinate.<sup>8,53,225–228</sup> Since a reaction coordinate is required for the application of the volume profile method of Chapter 4 to the unfolding of proteins, here we explore the reaction coordinates traditionally used to describe this process, discuss their functional forms and limitations thereof, and propose a novel formulation of the reaction coordinate for systems with multiple minima and multiple reaction pathways.

## 5.2. Collective variables

### 5.2.1. Brief overview

Over the past twenty years, many different definitions of reaction coordinates for complex flexible systems have been proposed. One of the earliest and most widely used approaches is that of collective variables, also often referred to as order parameters. A collective variable is a one-dimensional coordinate consisting of a combination of various geometrical parameters of the system, typically through-space distances and torsion angles. These geometrical parameters are readily available in MD simulations, and so collective variables are easy to obtain during protein folding simulations. The choice of collective variable is typically made in an *ad hoc* manner, as there is no generally accepted method for choosing which variable is the most suitable.

Of the commonly used collective variables, the two with the most transparent definition are the root mean square deviation (*RMSD*) from a reference structure<sup>229–236</sup> and radius of gyration ( $R_g$ ).<sup>233,237,238</sup> The *RMSD* coordinate describes the similarity between the

current structure and a reference structure, typically the folded state. It is calculated using eq. 5.1, where  $N$  is the total number of atoms,  $\mathbf{r}_i$  is the position vector for the  $i^{th}$  atom,  $\mathbf{r}_i^o$  is the position vector for this atom in the folded state and  $w_i$  are weighting factors, typically proportionate to the masses  $m_i$  of the atoms. The  $RMSD$  equals zero in the folded state and acquires a large value in the unfolded state.

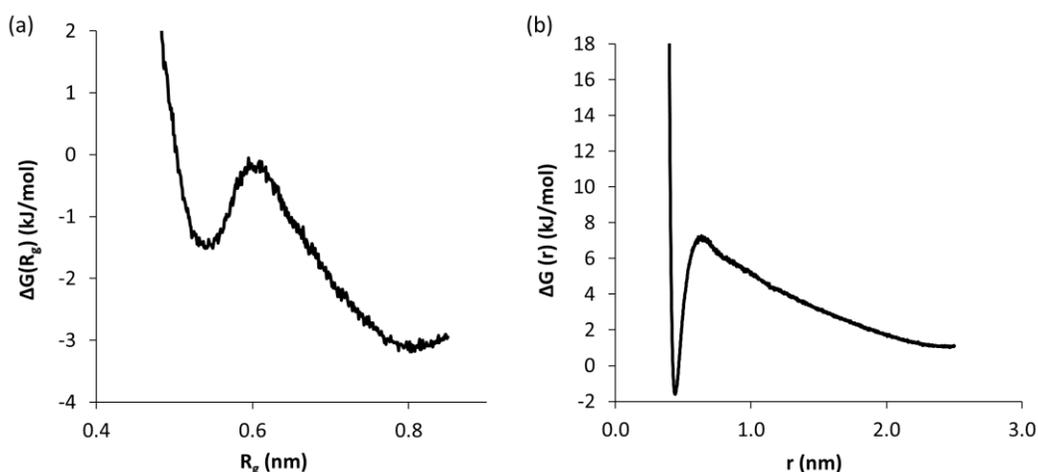
$$RMSD = \sqrt{\sum_{i=1}^N w_i (\mathbf{r}_i - \mathbf{r}_i^o)^2} \quad (5.1)$$

The  $R_g$  coordinate describes the average size of the system, and is calculated according to eq. 5.2 as the mass-weighted  $RMSD$  of each atom from the center of mass of the protein, where,  $N$ , and  $\mathbf{r}_i$  are as in eq. 5.1,  $M$  is the total mass of the protein,  $m_i$  is the mass of the  $i^{th}$  atom, and  $\mathbf{r}_{COM}$  is the position vector for the center of mass of the protein. Since proteins become less compact as they unfold, an increase in  $R_g$  indicates the unfolding of a protein.

$$R_g = \sqrt{\frac{1}{M} \sum_{i=1}^N m_i (\mathbf{r}_i - \mathbf{r}_{COM})^2} \quad (5.2)$$

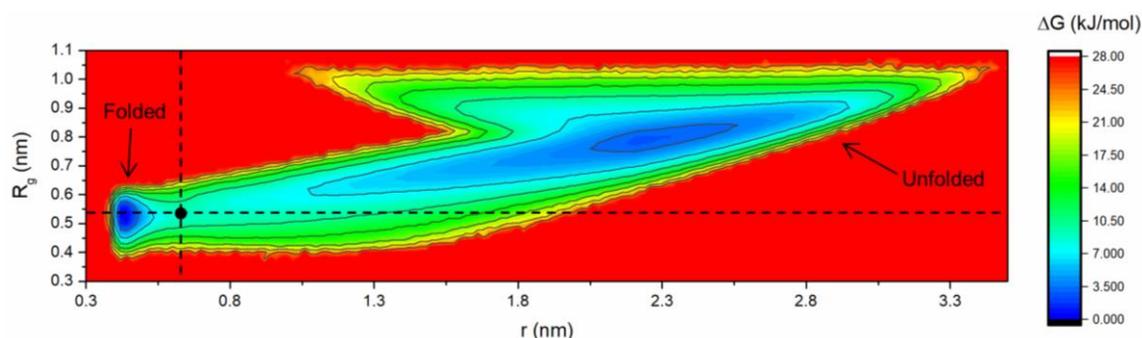
While  $RMSD$  and  $R_g$  are enumerated over all atoms in the protein, there are other collective coordinates that are more focused on specific structural aspects. One example that is very widely used is the fraction of native contacts  $Q$ .<sup>231–235,237,239–244</sup> For this coordinate, only the important contacts in the protein are considered, such as hydrogen bonds between atoms in different amino acid residues.  $Q$  scales from a value of 0 in the unfolded state, where no contacts are taking place, to a value of 1 in the folded state with all contacts properly formed. Helicity<sup>245–247</sup> is an example of a collective variable that focuses on the secondary structure of the protein, specifically the amount of  $\alpha$ -helical content in a given system, and is typically calculated as the average of the protein backbone angles  $\psi$  and  $\phi$ . Another type of collective variable that is occasionally used is the solvent accessible surface area.<sup>233,248</sup> The enthalpy<sup>249</sup> and conformational entropy<sup>250</sup> of the protein backbone have also been explored as potential collective variables.

Although collective variables are often useful at a qualitative level as a means to differentiate between folded and unfolded states, they do not necessarily produce kinetically relevant activation barriers.<sup>251,252</sup> Moreover, the physical picture of the process generated by one coordinate may be quite different from that produced by another. This can be illustrated by an example of the model chain system of Chapter 4. If radius of gyration  $R_g$  is used instead of the distance between terminal atoms  $r$  as the reaction coordinate for this system, the Gibbs energy profiles look quite different as can be seen in Figure 5.2. Not only is the barrier height significantly reduced when  $R_g$  is used, but also the relative stability of the folded and unfolded states is reversed.



**Figure 5.2** Gibbs energy profiles for the model chain system of Chapter 4 as a function of (a)  $R_g$  and (b)  $r$ . A change of coordinate results in dramatic changes in both the barrier height and the relative stability of the reactant and product states.

This situation can be better understood in terms of a two-dimensional Gibbs energy surface  $\Delta G(r, R_g)$  shown in Figure 5.3. This surface displays two stable domains that can be identified as the folded and unfolded states of the chain, separated by a bottleneck region where the TSE (black circle) is located. It can be seen that  $r$  is a good reaction coordinate since it practically coincides with the negative curvature mode of the saddle point at the TSE, and the hypersurface orthogonal to  $r$  (vertical dashed line) properly separates the folded and unfolded domains. The folded minimum and the saddle point (TSE) of  $\Delta G(r, R_g)$  are properly projected into the folded minimum and TS of  $\Delta G(r)$ . On the contrary, there is no separation between folded and unfolded domains by a hypersurface orthogonal to the  $R_g$  coordinate (horizontal dashed line).

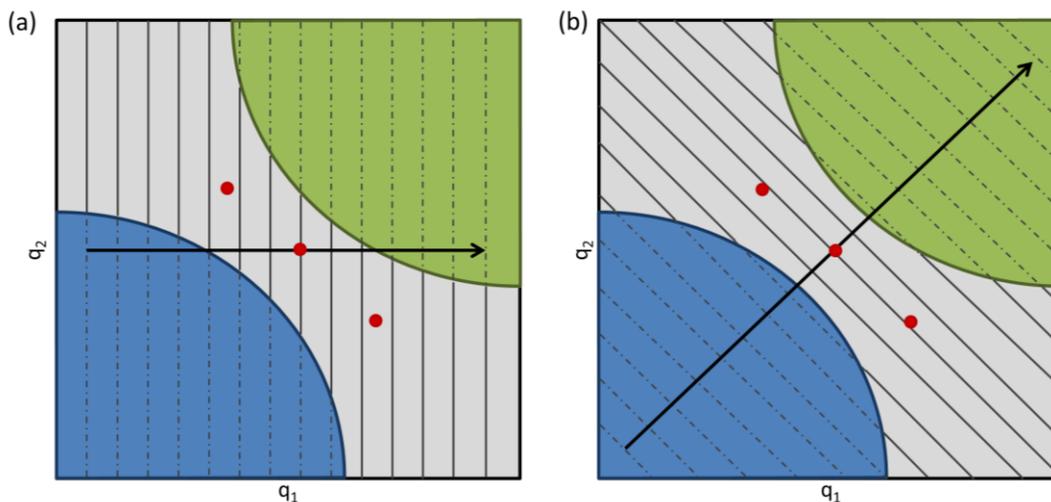


**Figure 5.3** A 2D Gibbs energy surface for the model chain system in  $r$  and  $R_g$  coordinates. The position of the TSE is denoted by the black circle and hypersurfaces orthogonal to  $r$  and  $R_g$  are denoted by the dashed lines.

As a result, when  $\Delta G(r, R_g)$  is projected into  $\Delta G(R_g)$ , the folded and unfolded states become mixed together, and both the position of the TS and the barrier height assessed by  $\Delta G(R_g)$  appear notably different from those assessed from  $\Delta G(r, R_g)$ . Therefore,  $R_g$  is not suitable for describing the unfolding process in this system. Similar artifacts rooted in a wrong choice of collective variable has also been reported for ion pair dissociation,<sup>253</sup> isomerization in an alanine dipeptide,<sup>254</sup> and for more complex systems such as the folding of a small protein.<sup>236</sup>

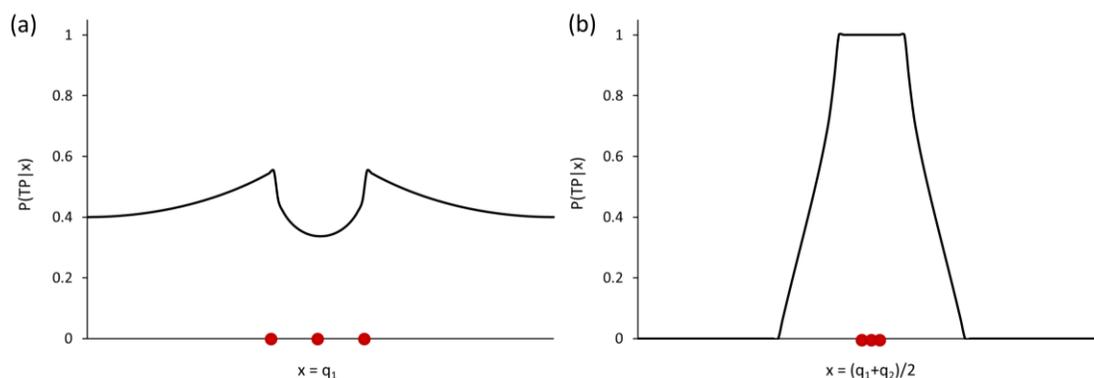
### 5.2.2. Best-Hummer optimization

Several procedures have been proposed to guide the choice of collective variables as reaction coordinates.<sup>225,255–259</sup> One such approach, proposed by Best and Hummer,<sup>225</sup> involves the optimization of a collective variable by minimizing the width of the transition region separating the reactant and product domains. In brief, this approach is illustrated in Figure 5.4. The configurational space of a reaction system is divided into three regions – the reactant domain (bottom left, blue), the product domain (top right, green), and the transition path ( $TP$ ) region (grey). The red circles represent individual TSs which make up the TSE. Conformations of the system are described by structural parameters  $q_1$  and  $q_2$ , which can be important distances, angles, or other geometrical parameters. These parameters can be used as components of different collective variables to describe transformations in the system. Two examples of possible collective variables  $x = q_1$  (Figure 5.4a) and  $x = 1/2 (q_1 + q_2)$  (Figure 5.4b) are shown by the black arrows, accompanied by a system of orthogonal hypersurfaces of constant  $x$ .



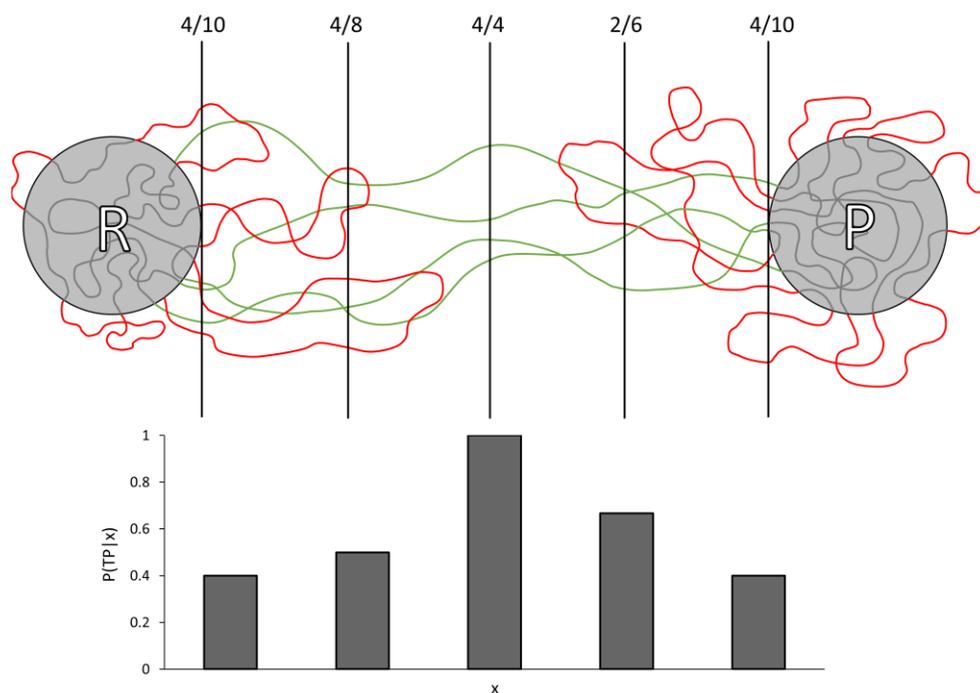
**Figure 5.4** Schematic of the configurational space of a flexible system with two choices of collective variables: (a)  $x = q_1$  and (b)  $x = 1/2 (q_1 + q_2)$ . See text for details.

Depending on the choice of collective variable, these hypersurfaces sample the  $TP$  region differently (solid line segments). Figure 5.5 displays the probability  $P(TP|x)$  that a point on a hypersurface lies in the  $TP$  region for each collective variable. The horizontal coordinate samples the  $TP$  region at all values of  $x$ , and so the distribution is broad and the TSs (red circles) are spread over a wide range of values of  $x$ . The diagonal coordinate, on the other hand, samples the  $TP$  region more efficiently and produces a narrower distribution with the individual TSs of the TSE clustering closer together, as should be expected from a good reaction coordinate.<sup>225</sup>



**Figure 5.5** Probability that a point on a hypersurface of constant  $x$  lies in the transition path region for (a)  $x = q_1$  and (b)  $x = 1/2 (q_1 + q_2)$  as defined in Figure 5.4. Red circles label positions of the individual TSs of the TSE.

Best and Hummer proposed that the reaction coordinate for a given system is defined as the collective variable that produces the narrowest  $P(TP|x)$  distribution and that these distributions could be calculated using sufficiently long equilibrium trajectories, which visit reactant  $R$  and product  $P$  domains multiple times. A schematic depicting one such sampling trajectory is shown in Figure 5.6. The green lines represent the segments of the trajectory that proceed from one state to the other without returning, known as transition paths. The red segments represent non-reactive parts of the trajectory, where the system returns to the state from which it departed. Although the overall density of trajectory segments is higher in the vicinity of the reactant and product domains, the relative density of transition paths is higher in the middle. Therefore, the probability  $P(TP|x)$  is low near the reactant and product domains, and reaches a peak in the middle of the  $TP$  region.



**Figure 5.6** Long equilibrium trajectory that visits reactant ( $R$ ) and product ( $P$ ) domains multiple times. The transition path segments of this trajectory are shown in green, and non-reactive segments are shown in red. The  $P(TP|x)$  distribution (bottom) reaches its highest value in the middle of the  $TP$  region.

From these trajectories, the  $P(TP|x)$  distribution for a given collective variable  $x$ , can be calculated from the Bayesian relationship of eq. 5.3,<sup>24</sup>

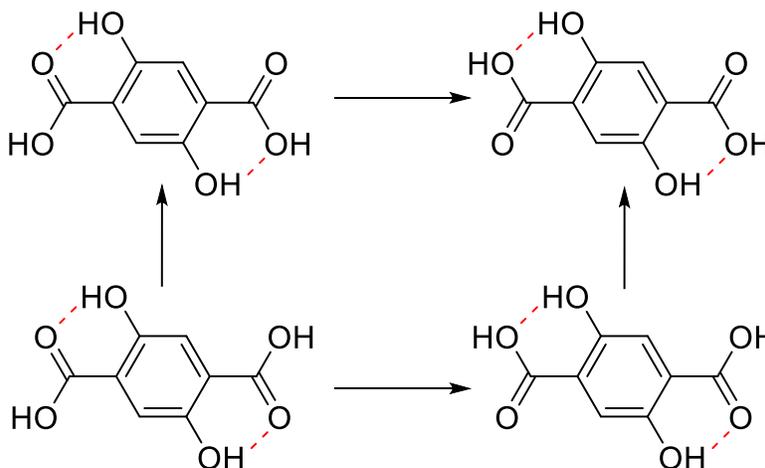
$$P(TP|x) = \frac{P(x|TP)P(TP)}{P(x)} \quad (5.3)$$

where  $P(x)$  is the equilibrium distribution of  $x$  calculated on the entire trajectory,  $P(TP)$  is the probability that a point of the trajectory is on one of its transition path segments, and  $P(x|TP)$  is the distribution of  $x$  calculated only for transition path segments. In the original work,<sup>225</sup> the collective variable  $x$  was chosen to be a weighted average of structural parameters and a Monte Carlo procedure was used to obtain the weights that gave the optimal  $P(TP|x)$  distribution. However, neither of these specific circumstances is central to the approach, which has been used with success for some biological systems.<sup>260–262</sup>

### 5.3. Comparative analysis of reaction coordinates for transformations in a four-state system

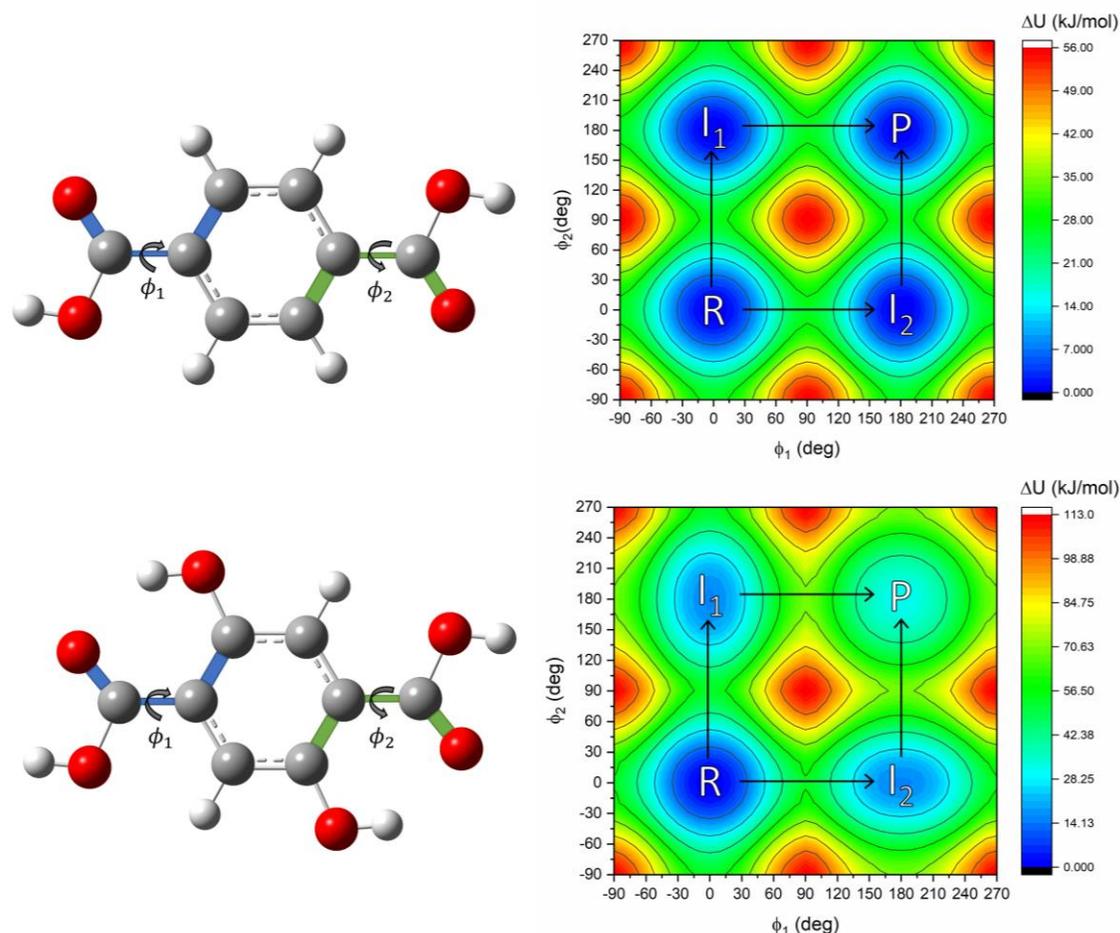
#### 5.3.1. Chemical prototypes

To test different collective variables and to explore how the Best-Hummer process allows for their optimization, we chose to study their application to simple two-dimensional models with multiple-minima potential energy surfaces (PESs). An example of a chemical system with a four-minima PES is terephthalic acid and its derivatives undergoing conformational transformations through rotation of the carboxylic acid groups, as shown in Figure 5.7 for 2,5-dihydroxyterephthalic acid.



**Figure 5.7** Two possible reaction pathways for the rotation of the carboxylic acid groups in 2,5-dihydroxyterephthalic acid resulting in restructuring of the two intramolecular hydrogen bonds (red dashed lines).

Gas phase PESs for terephthalic and 2,5-dihydroxyterephthalic acid as functions of the two torsion angles  $\phi_1$  and  $\phi_2$  obtained using QM calculations at the B3LYP/6-31++G(d,p) level are shown in Figure 5.8. The surfaces possess four stable states matching those shown in Figure 5.7, with the reactant *R* in the bottom left and the product *P* in the top right. The other two states are intermediates, labelled *I*<sub>1</sub> and *I*<sub>2</sub>, and four TSs are located between the adjacent minima. Reactions on this surface can proceed through two reaction paths, indicated by the arrows.

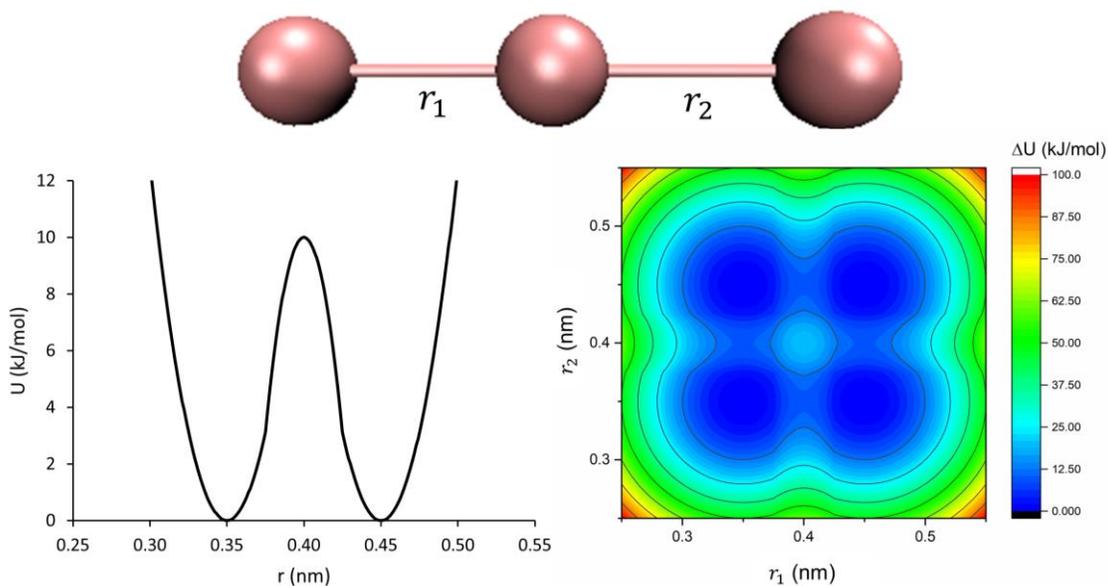


**Figure 5.8** Four-state B3LYP/6-31++G(d,p) PESs for terephthalic acid (top) and 2,5-dihydroxyterephthalic acid (bottom) showing two possible reaction pathways. Torsion angles  $\phi_1$  and  $\phi_2$  are as defined on the left.

Other suitable cases can be found among various examples of chemical systems with two essential torsional degrees of freedom, such as substituted and unsubstituted isophthalic acid, terphenyls and their derivatives.

### 5.3.2. Model system

The simplest model system possessing a similar PES is a linear triatomic, with two bonds  $r_1$  and  $r_2$ , each of which is described by a double well potential (Figure 5.9). For computational convenience, we selected this model to study the suitability of various collective variables as reaction coordinates describing transformations in a four-state system. The double-well potential was constructed from three parabolic segments, smoothly connected to create a barrier of 10 kJ/mol at  $r = 0.40$  nm between two stable minima – a “short” state at  $r = 0.35$  nm and a “long” state at  $r = 0.45$  nm. The non-bonded parameters were the same as those used for the model chain of Chapter 4, and the angle was kept at  $180^\circ$  using a harmonic potential with force constant of  $500.0 \text{ kJ mol}^{-1} \text{ rad}^{-2}$ . Simulations were performed with the model triatomic in 350 molecules of a monoatomic Lennard-Jones solvent ( $\epsilon = 2 \text{ kJ/mol}$ ,  $\sigma = 0.35 \text{ nm}$ ) at  $T = 300 \text{ K}$  and  $P = 1 \text{ kbar}$  required to keep the solvent liquid.

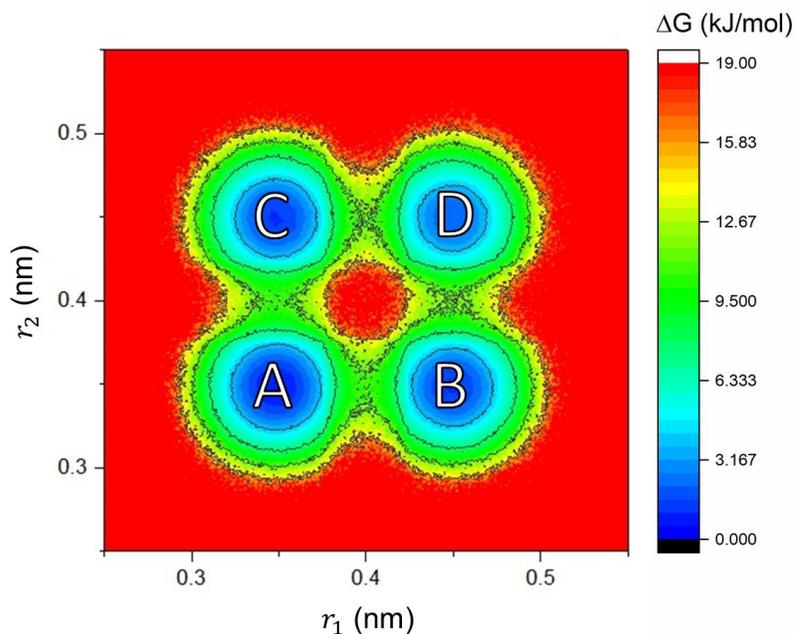


**Figure 5.9** Top: Model linear triatomic system with two bonds labelled  $r_1$  and  $r_2$ . Bottom left: The double well 1D potential used to describe interactions between neighboring (“bonded”) atoms in this system. Bottom right: the four-state 2D PES for the triatomic obtained as a combination of the 1D potentials.

The Gibbs energy surface (GES)  $\Delta G(r_1, r_2)$  of this system in  $r_1$  and  $r_2$  coordinates was calculated according to eq. 5.4 from a probability distribution function  $P(r_1, r_2)$  accumulated over equilibrated MD trajectories with a total length of  $8.9 \mu\text{s}$ .

$$\Delta G(r_1, r_2) = -RT \ln(P(r_1, r_2)) \quad (5.4)$$

The resultant  $\Delta G(r_1, r_2)$ , displayed in Figure 5.10, shows a total of four stable states matching those of the PES of Figure 5.9: state *A* with  $r_1, r_2 < 0.4$  nm, state *D* with  $r_1, r_2 > 0.4$  nm, state *B* with  $r_1 > 0.4$  nm,  $r_2 < 0.4$  nm and state *C* with  $r_1 < 0.4$  nm,  $r_2 > 0.4$  nm. The “flattening” of the GES that occurs in the regions far from the minima is due to insufficient sampling of these high energy states during the MD simulation.



**Figure 5.10** MD-generated Gibbs energy surface  $\Delta G(r_1, r_2)$  for the triatomic system, labelled with the location of states *A*, *B*, *C* and *D*.

### 5.3.3. 2D Kinetics

By analogy with Figure 5.8, the *A* state was considered to be the reactant, and the *D* state was considered to be the product. Kinetics of the depletion of the *A* state and of the accumulation of the *D* state was obtained using the COPASI software<sup>263</sup> for 2D networks shown in Figure 5.11.

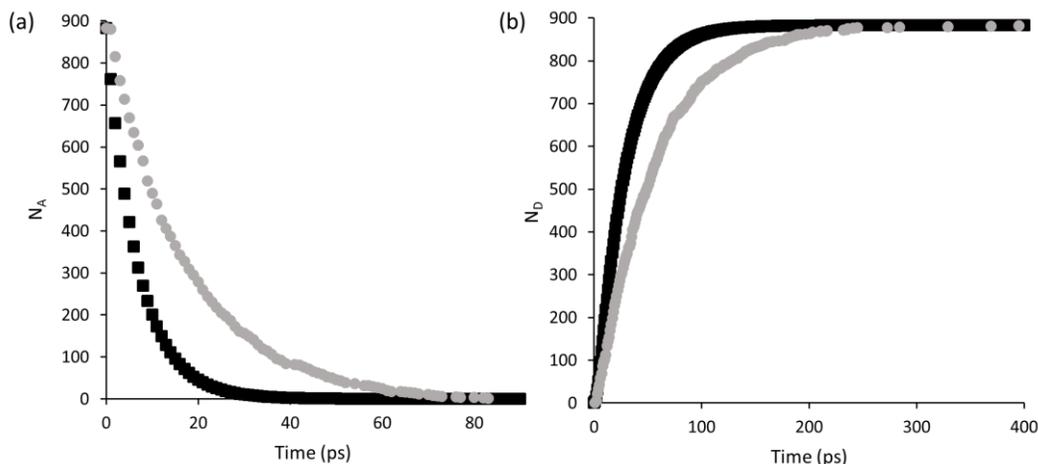


**Figure 5.11** Kinetic schemes for the depletion of the *A* state (left) and for the accumulation of the *D* state (right).

The necessary partial rate constants were calculated from the relevant Gibbs energy barriers  $\Delta G^\ddagger$  assessed on the GES using the Eyring equation (eq. 5.5), where  $k_B$ ,  $h$  and  $R$  are the Boltzmann, Planck and gas constants, respectively.

$$k = \frac{k_B T}{h} \exp\left(\frac{-\Delta G^\ddagger}{RT}\right) \quad (5.5)$$

For comparison, the kinetics was also assessed using MD simulations.<sup>209</sup> To obtain the necessary kinetic data, 885 1.0 ns segments of trajectories initiated from within the *A* state were followed and the amount of time required for the system to exit the *A* state and to enter the *D* state was recorded thus generating the MD kinetic curves. Comparison between these curves and the kinetic curves produced from the GES and the Eyring equation is shown in Figure 5.12.



**Figure 5.12** Kinetic curves for (a) exit from the *A* state and (b) entry into the *D* state, determined using MD simulations (grey circles) and the kinetic schemes of Figure 5.11 (black squares).

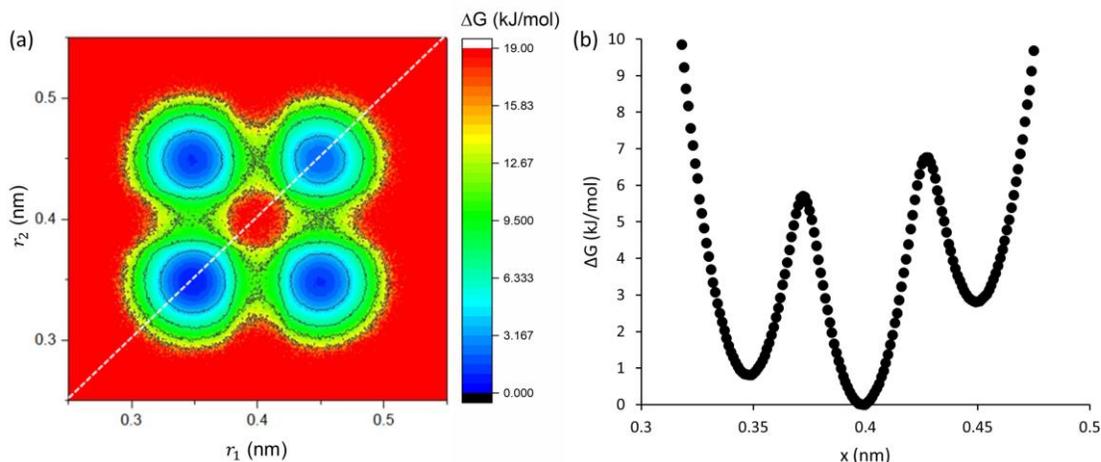
The difference in the kinetic curves, which should probably be attributed to dynamic effects,<sup>264–266</sup> is unimportant in the context of our further discussion since only the TS theory results are needed as a benchmark for assessing the kinetic quality of the Gibbs energy profiles generated by various choices of reaction coordinates.

### 5.3.4. Linear combination coordinate

Following Best and Hummer,<sup>225</sup> we began with the simplest collective variable – a linear combination of the  $r_1$  and  $r_2$  coordinates (eq. 5.6), where  $w_1$  and  $w_2$  are the relative weights of the two coordinates.

$$x = w_1 r_1 + w_2 r_2 \quad (5.6)$$

To find the weighting that produces the optimal reaction coordinate, the  $P(TP|x)$  distribution was assessed via eq. 5.3 using the same MD trajectories that were used to build the GES. The boundary of the reactant state  $A$  and the product state  $D$  were chosen to be  $r_1, r_2 < 0.4$  nm and  $r_1, r_2 > 0.4$  nm respectively, consistent with the GES of Figure 5.10. Monte Carlo optimization produced an optimal weighting of  $w_1 = 0.53$ ,  $w_2 = 0.47$ , which is consistent with  $w_1 = w_2 = 0.5$  expected due to the symmetry of the system.



**Figure 5.13** (a) The optimized coordinate (white dashed line) shown on the GES for the triatomic system. (b) The one-dimensional  $\Delta G(x)$  profile along this coordinate.

The one-dimensional Gibbs energy profile  $\Delta G(x)$  along this optimized reaction coordinate, calculated from the probability distribution  $P(x)$  using eq. 5.7, is shown in Figure 5.13. This profile has 3 minima representing reactant, intermediate and product states, respectively.

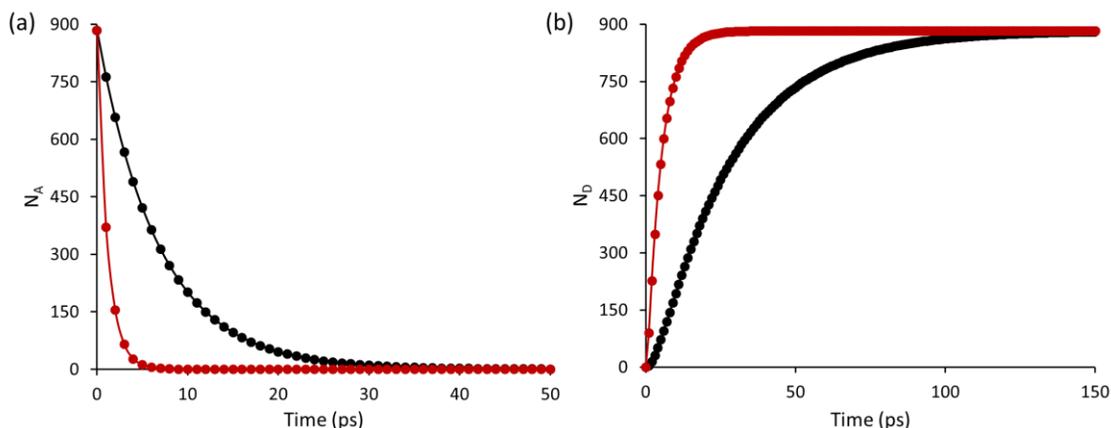
$$\Delta G(x) = -RT \ln P(x) \quad (5.7)$$

The 2D reaction networks of Figure 5.11 project into the 1D kinetic schemes of Figure 5.14. The rate constants  $k_i$  of these 1D kinetic schemes can be calculated using the Eyring equation (eq. 5.5) from the Gibbs energy barriers on the 1D reaction profile of Figure 5.13(b). For this profile to be considered kinetically adequate, the kinetic curves for depletion of reactant and accumulation of product must match those obtained on the underlying 2D kinetic networks.



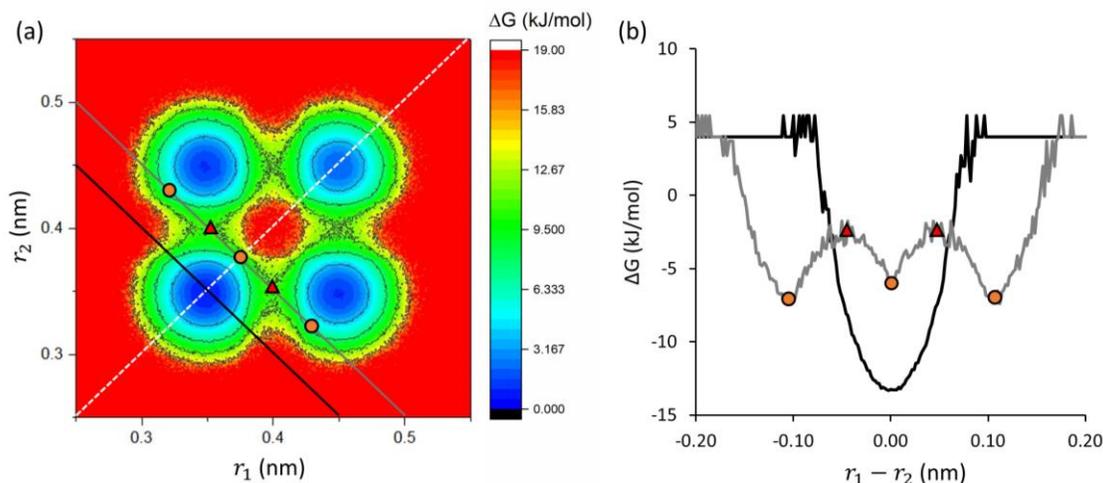
**Figure 5.14** 1D kinetic scheme representing projections of 2D reaction networks of Figure 5.11 onto the optimized reaction coordinate for depletion of the reactant state (left) and accumulation of the product state (right).

A comparison of the kinetic curves from  $\Delta G(x)$  with the curves generated from the full 2D surface (Figure 5.15) shows that kinetics on the 1D profile is significantly faster than on the 2D surface due to the fact that the 1D reaction barriers ( $\sim 5$  kJ/mol) are notably lower than the barriers on the 2D GES ( $\sim 10$  kJ/mol).



**Figure 5.15** Kinetic curves for (a) depletion of the reactant state and for (b) accumulation of the product state for the kinetic network from the 2D surface (black) and for the 1D profile for the optimized linear coordinate (red).

The reason for this becomes clear upon consideration of the cross sections of the surface at constant  $x$ , as illustrated in Figure 5.16. The Gibbs energy profiles along these hypersurfaces of constant  $x$ , which are orthogonal to the optimized  $x$  coordinate, are plotted as a function of  $r_1 - r_2$  for two different values of  $x$ .



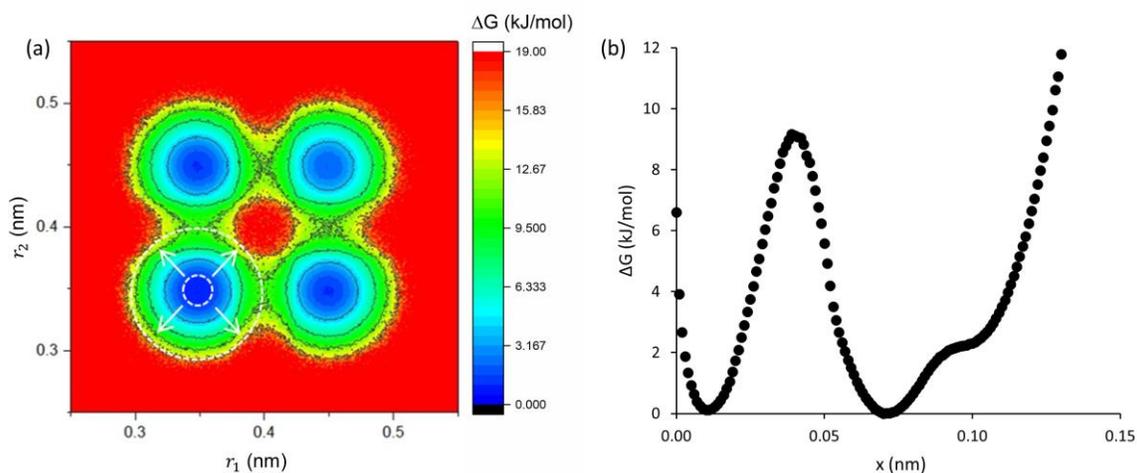
**Figure 5.16** (a) Hypersurfaces of constant  $x$  at the reactant  $x = 0.35$  (black), and the TS  $x = 0.375$  (grey). (b) The Gibbs energy profiles along these hypersurfaces. The maxima and minima of the  $x = 0.375$  profile are labeled by red triangles and orange circles respectively.

The black curve, for constant  $x = 0.35$  nm, corresponds to the reactant state on the 1D  $\Delta G(x)$ . This hypersurface shows a single minimum at  $r_1 - r_2 = 0$ , indicating that the dominant contribution to the energy of this point on the 1D reaction profile comes from the true 2D minimum of state  $A$  at  $r_1 = r_2 = 0.35$  nm. The grey curve, for constant  $x = 0.375$  nm, corresponds to the first TS on the 1D  $\Delta G(x)$ . This hypersurface shows a more complex shape, with two maxima at  $r_1 - r_2 = \pm 0.05$  nm, corresponding to the TSs on the 2D surface (at  $r_1 = 0.40$ ,  $r_2 = 0.35$  nm and at  $r_1 = 0.35$ ,  $r_2 = 0.40$  nm). The Best-Hummer procedure has therefore achieved the goal of producing a reaction coordinate such that all members of the TSE lie on the same hypersurface. However, the TSs are not energy minima on this hypersurface, and therefore do not give the major contribution to the average energy of the corresponding point of the 1D reaction profile. Instead, the major contributions to the average energy of this hypersurface come from the minima located within the  $A$ ,  $B$ , and  $C$  domains at  $r_1 - r_2 = 0$  and  $\pm 0.10$  nm, which artificially reduces the reaction barrier along  $x$  and inflates the value of the rate constant.

### 5.3.5. RMSD coordinate

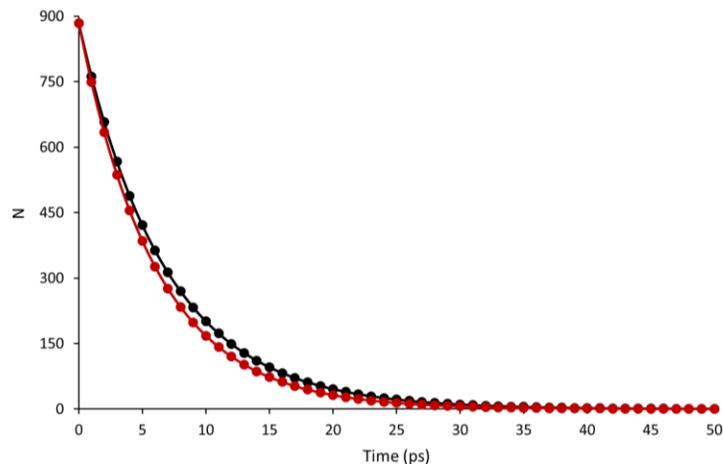
A linear combination of  $r_1$  and  $r_2$  fails to produce a kinetically adequate 1D reaction profile due to the fact that the contribution of the TSE to the reaction barrier is contaminated by contributions from lower-energy structures improperly sampled by the  $x = x^\ddagger$  hypersurface in configurational space. As can be seen from Figure 5.17(a), an RMSD coordinate (eq. 5.8) centered at the bottom of the reactant well ( $r_1 = r_2 = 0.35$  nm) may offer a better sampling.

$$x = \sqrt{w_1(r_1 - 0.35)^2 + w_2(r_2 - 0.35)^2} \quad (5.8)$$



**Figure 5.17** (a) The optimized circular coordinate (white concentric circles) shown on the GES for the triatomic system. (b) The one-dimensional  $\Delta G(x)$  profile along this coordinate.

Monte Carlo optimization gave  $w_1 = 0.52$ ,  $w_2 = 0.48$ , consistent with symmetry-expected values of  $w_1 = w_2 = 0.5$ . This optimized coordinate produces the 1D  $\Delta G(x)$  profile shown in Figure 5.17(b). Unlike the linear coordinate, this profile shows only two minima at  $x = 0.01$  and  $x = 0.07$ , with a shoulder at  $x = 0.1$ , since the hypersurfaces of constant  $x$  for this coordinate do not properly sample the configurational space outside of the reactant domain  $A$ . Yet for the depletion of state  $A$ , this coordinate was able to produce an excellent agreement with the kinetic curve from the 2D surface, as illustrated in Figure 5.18.



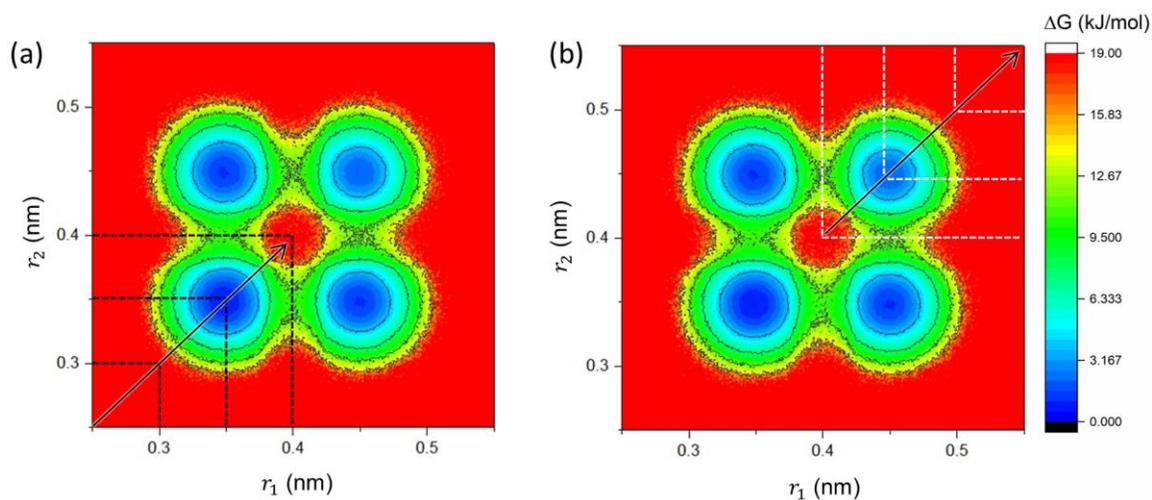
**Figure 5.18** Kinetic curves for the depletion of the reactant state for the kinetic network from the 2D surface (black) and for the 1D profile for the optimized RMSD coordinate (red).

### 5.3.6. Piecewise defined coordinate

The RMSD coordinate is an example of a good reaction coordinate for the  $A$  domain, as it provides good separation between reactant and product states as well as an accurate activation barrier for exit from that region. However, it is a poor coordinate for the rest of the surface due to its inappropriate shape, resulting in considerable overlap for the  $B$ ,  $C$  and  $D$  domains. The situation could have been reversed if the reference point in eq. 5.8 was chosen to be  $(0.45, 0.45)$  representing the minimum  $D$  rather than  $(0.35, 0.35)$  representing minimum  $A$ . In this case, the  $D$  domain would be sampled properly, but at the expense of improper sampling of states  $A$ ,  $B$ , and  $C$ . A coordinate that is appropriate for the entire configurational space could probably be found as an interpolation between these two limiting cases; however such an interpolation does not seem possible for the RMSD coordinate. Therefore, a good coordinate for the entire space will require a different functional form. An example of such coordinate, based on the  $d_\infty$  metric<sup>267,268</sup> is described below.

The reaction coordinate  $x \leq 1$  in the  $A$  domain is defined according to eq. 5.9 with adjustable parameters  $a$  and  $b$ . The hypersurfaces of constant  $x$  defined by this equation are shown in Figure 5.19(a) for the symmetric case of  $a = b$ .

$$x = \max\left(\frac{r_1}{a}, \frac{r_2}{b}\right), \quad \text{if } \frac{r_1}{a} \text{ and } \frac{r_2}{b} \leq 1 \quad (5.9)$$



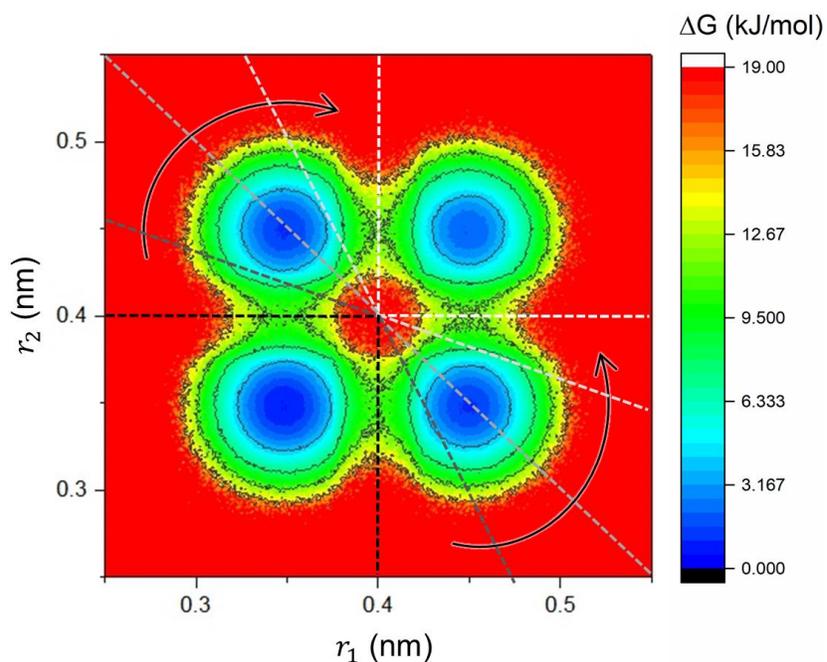
**Figure 5.19** The hypersurfaces of constant  $x$  (dashed lines) for (a)  $x$  defined by eq. 5.9 (the  $A$  domain), and (b)  $x$  defined by eq. 5.10 (the  $D$  domain)

The functional form for  $x$  in the  $D$  domain is similar, but must be inverted for proper sampling. Its values are also shifted by 1 ( $x \geq 2$ ) to make room for  $x$  values describing the intermediates (domains  $B$  and  $C$ ). This coordinate is defined by eq. 5.10, and illustrated in Figure 5.19(b).

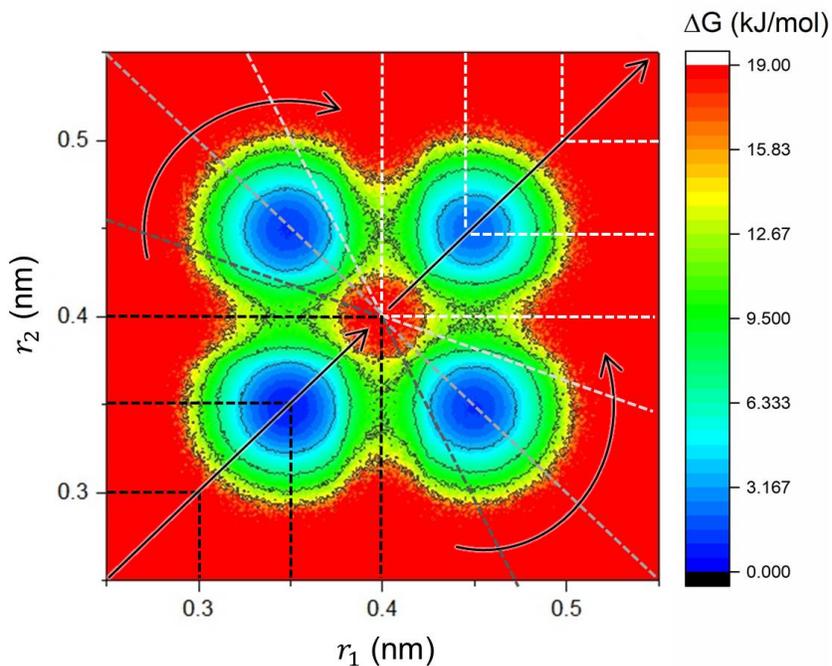
$$x = \min\left(\frac{r_1}{a}, \frac{r_2}{b}\right) + 1, \quad \text{if } \frac{r_1}{a} \text{ and } \frac{r_2}{b} \geq 1 \quad (5.10)$$

In the intermediate region, ( $1 \leq x \leq 2$ ) the constant  $x$  hypersurfaces must transition smoothly between their orientations suitable for domains  $A$  and  $D$  as illustrated in Figure 5.20. This interpolating function for  $x$  is given by eq. 5.11.

$$x = \frac{2}{\pi} \sin^{-1} \left( \frac{\max(r_1 - a, r_2 - b)}{\sqrt{(r_1 - a)^2 + (r_2 - b)^2}} \right) + 1 \quad (5.11)$$



**Figure 5.20** Constant  $x$  hypersurfaces in the intermediate zone (i.e. the  $B$  and  $C$  domains), providing an interpolation between the border of the  $A$  domain (black) and the border of the  $D$  domain (white).



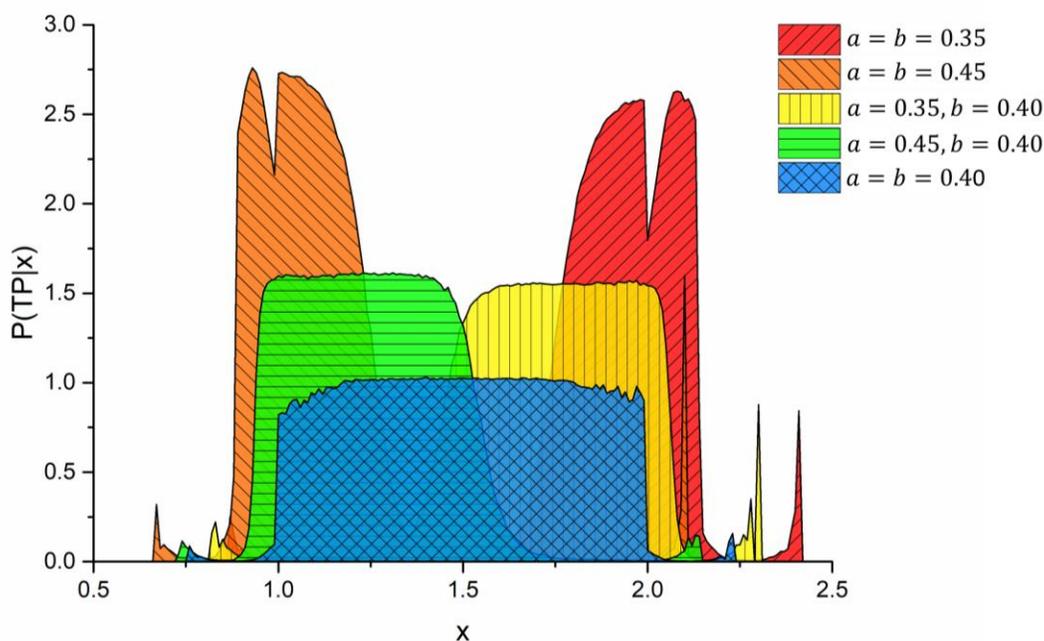
**Figure 5.21** Hypersurfaces of constant  $x$  for the piecewise reaction coordinate defined by eq. 5.12.

The overall reaction coordinate  $x$  can therefore be presented as the piecewise function of eq. 5.12.

$$x = \begin{cases} \max\left(\frac{r_1}{a}, \frac{r_2}{b}\right), & \text{if } \frac{r_1}{a} \text{ and } \frac{r_2}{b} \leq 1 \\ \min\left(\frac{r_1}{a}, \frac{r_2}{b}\right) + 1, & \text{if } \frac{r_1}{a} \text{ and } \frac{r_2}{b} \geq 1 \\ \frac{2}{\pi} \sin^{-1}\left(\frac{\max(r_1 - a, r_2 - b)}{\sqrt{(r_1 - a)^2 + (r_2 - b)^2}}\right) + 1, & \text{otherwise} \end{cases} \quad (5.12)$$

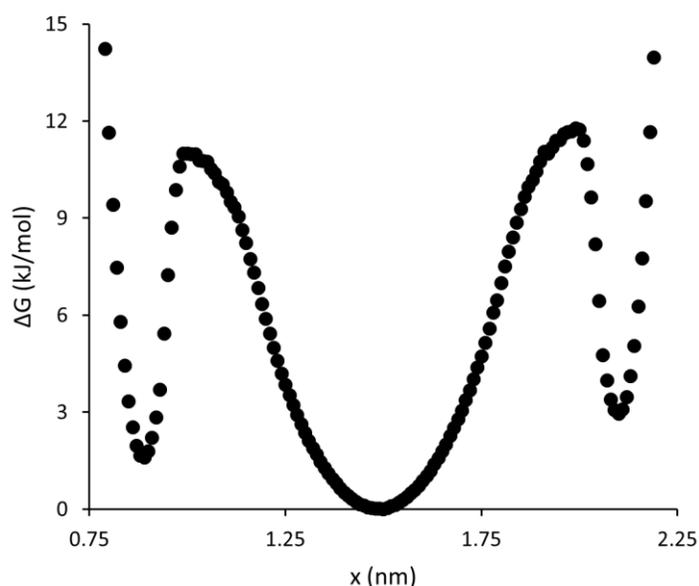
The constant  $x$  hypersurfaces for this coordinate are shown in Figure 5.21. The location of the “pivot point” at  $r_1 = a$ ,  $r_2 = b$  is considered to be the optimizable parameter for the Best-Hummer method.

Since the transition region is already defined as the fixed range of  $1 < x < 2$ , this width, and therefore the related height of the  $P(TP|x)$  distribution, can no longer be used as the target function for optimization. Instead, the location of the pivot point  $r_1 = a$ ,  $r_2 = b$  was optimized to achieve the maximum integrated area under  $P(TP|x)$  within  $1 < x < 2$ , or in other words, to maximize the total probability of finding transition paths within the pre-defined intermediate region. The sample  $P(TP|x)$  distributions of Figure 5.22 show that the optimal pivot point of  $a = b = 0.4$  nm produces the desired shape while the other pivot points produce distributions that are skewed towards the reactants or products.

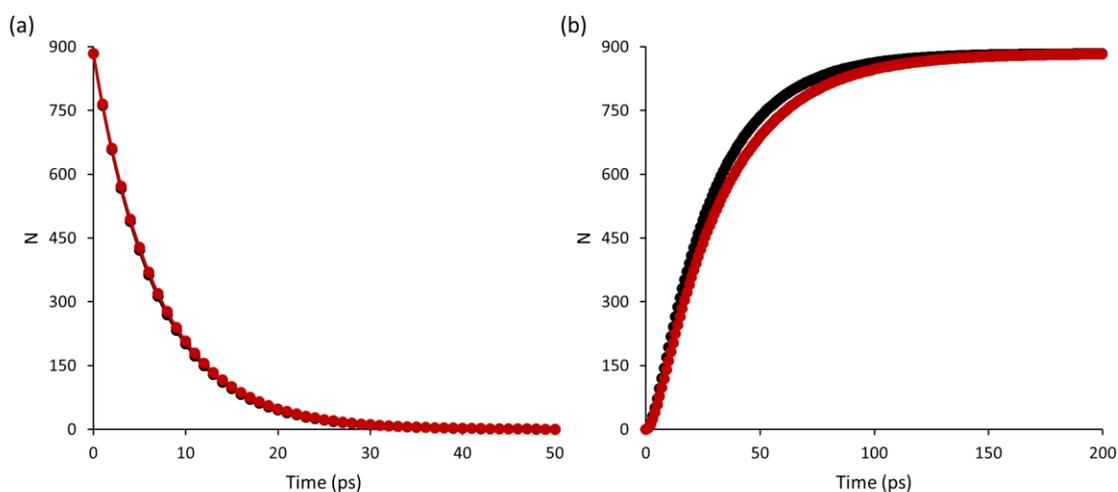


**Figure 5.22** Sample  $P(TP|x)$  distributions for the piecewise coordinate of eq. 5.12. Only  $a = b = 0.4$  nm produces a distribution contained entirely within the intermediate region of  $1 < x < 2$ .

The 1D  $\Delta G(x)$  profile for this optimal coordinate is shown in Figure 5.23. Unlike the two previous coordinates (linear combination and RMSD), this profile looks extremely symmetrical, as expected due to the symmetry of the GES. The kinetic curves for decomposition of the reactant and accumulation of the product in Figure 5.24 show good agreement with the kinetics obtained from the full 2D surface.



**Figure 5.23** One dimensional profile  $\Delta G(x)$  along the optimal two dimensional coordinate.



**Figure 5.24** Comparison of the kinetic curves derived from the 2D kinetic networks of Figure 5.11 (black) and from the 1D profile for the optimized piecewise coordinate of eq 5.12 (red) for (a) the decomposition of the reactant and (b) the accumulation of the product.

## 5.4. Reaction coordinates for multi-minima domains

While the 4-state linear triatomic provides a model for systems with multiple reaction pathways, it fails to capture the multi-minima nature of the reactant and product states on the free energy landscape for protein folding or other highly flexible systems. To assess the efficacy of the  $d_\infty$ -metric coordinate for multiple-minima reactant domains, we consider a 16-minima model surface described below.

### 5.4.1. Model surface

A Gibbs energy surface in generic structural coordinates  $q_1$  and  $q_2$  was constructed for  $0 \leq q_1, q_2 \leq 4$  using eq. 5.13, with a barrier height  $V$  of 10.8 kJ/mol, matching the average barrier height of the triatomic GES of Figure 5.10. The resulting surface contains 16 minima, labeled as  $S_{ij}$  ( $i, j = 1, 2, 3, 4$ ), and 16 maxima as shown in Figure 5.25.

$$\Delta G(q_1, q_2) = 0.5V(2 - \cos(2\pi q_1 - \pi) - \cos(2\pi q_2 - \pi)) \quad (5.13)$$

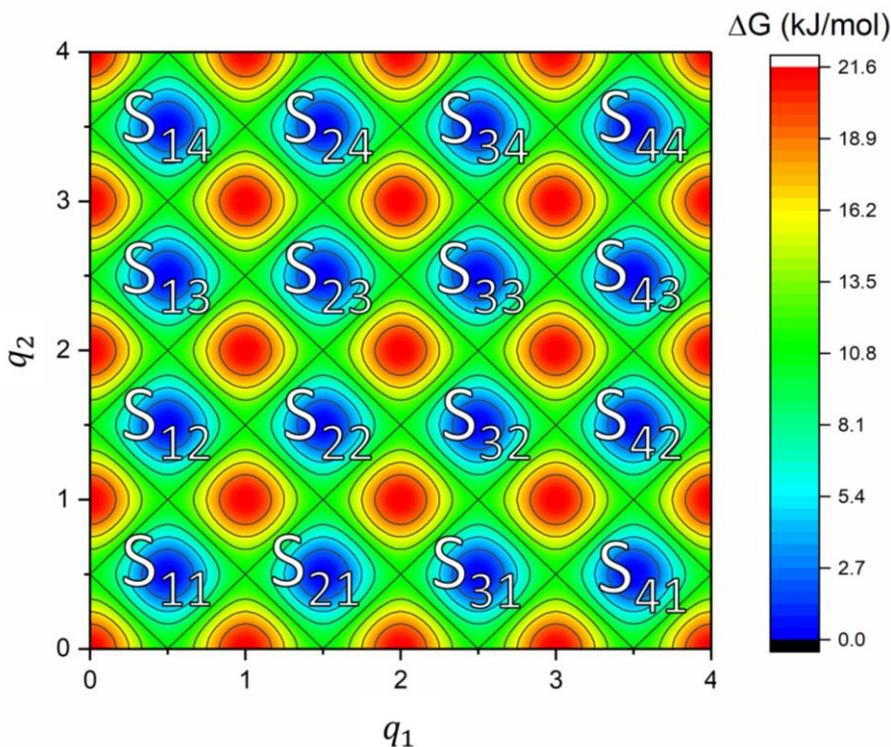


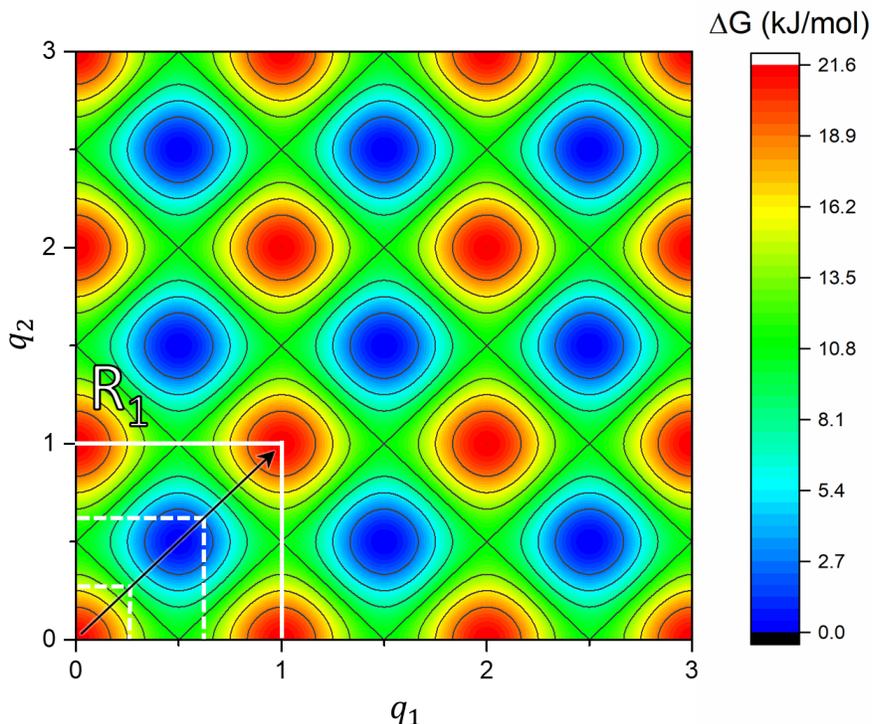
Figure 5.25 The model 16-state Gibbs energy surface of eq. 5.13. Blue configurations are minima, and red configurations are maxima.

Due to the symmetry of this surface, the activation barriers for transition between two adjacent minima are all equal to 10.8 kJ/mol. The value of the rate constant  $k$  for these transitions was obtained using the Eyring equation (eq. 5.5) and found to be  $8.23 \times 10^{10} \text{ s}^{-1}$ . The 1D and 2D kinetics for the depletion of the reactant was compared for four differently defined reactant domains as discussed below.

### 5.4.2. A single-minimum reactant domain

The first reactant domain ( $R_1$ , shown in Figure 5.26), encompasses only the  $S_{11}$  minimum and thus matches the previously discussed case of exit from the  $A$  state of the triatomic model system. As with the triatomic system, the reaction coordinate  $x_{R_1}$  is defined in eq. 5.14 and is scaled to be equal to 0 at the origin and equal to 1 at the  $R_1$  boundary.

$$x_{R_1} = \max(q_1, q_2) \quad (5.14)$$



**Figure 5.26** The boundary of the  $R_1$  domain (white solid line) with hypersurfaces of constant  $x_{R_1}$  (white dashed lines). Note that for clarity the surface has been zoomed in.

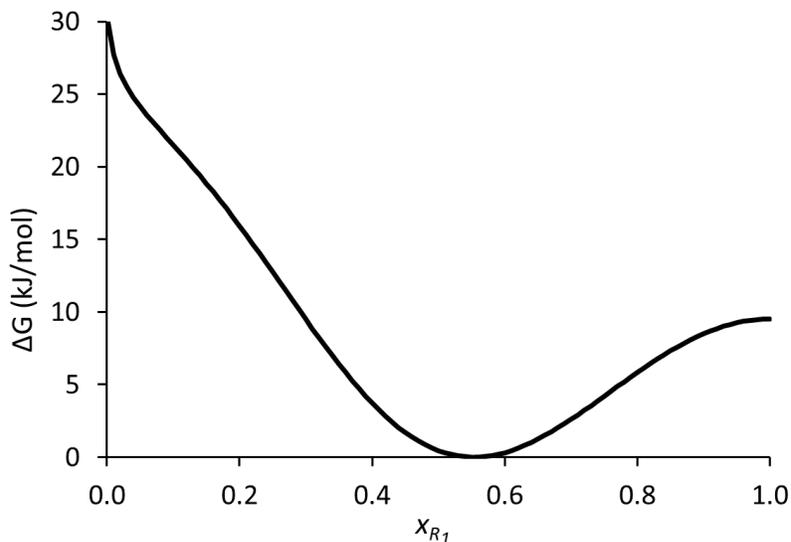
A one-dimensional probability density function (PDF)  $P(x_{R_1})$  for the  $x_{R_1}$  coordinate was generated by summing the probabilities along each hypersurface of constant  $x_{R_1}$  on the two-dimensional PDF  $P(q_1, q_2)$  obtained for  $T = 300\text{K}$  from the 2D GES  $\Delta G(q_1, q_2)$  using eq. 5.15.

$$P(q_1, q_2) = \exp\left(\frac{-\Delta G(q_1, q_2)}{RT}\right) \quad (5.15)$$

The  $P(x_{R_1})$  was then normalized and transformed according to eq. 5.16 to obtain the 1D Gibbs energy profile  $\Delta G(x_{R_1})$ . This method was used to obtain Gibbs energy profiles for all subsequent cases of multi-minima domains as well.

$$\Delta G(x_{R_1}) = -RT \ln\left(\frac{P(x_{R_1})}{\int P(x_{R_1}) dx_{R_1}}\right) \quad (5.16)$$

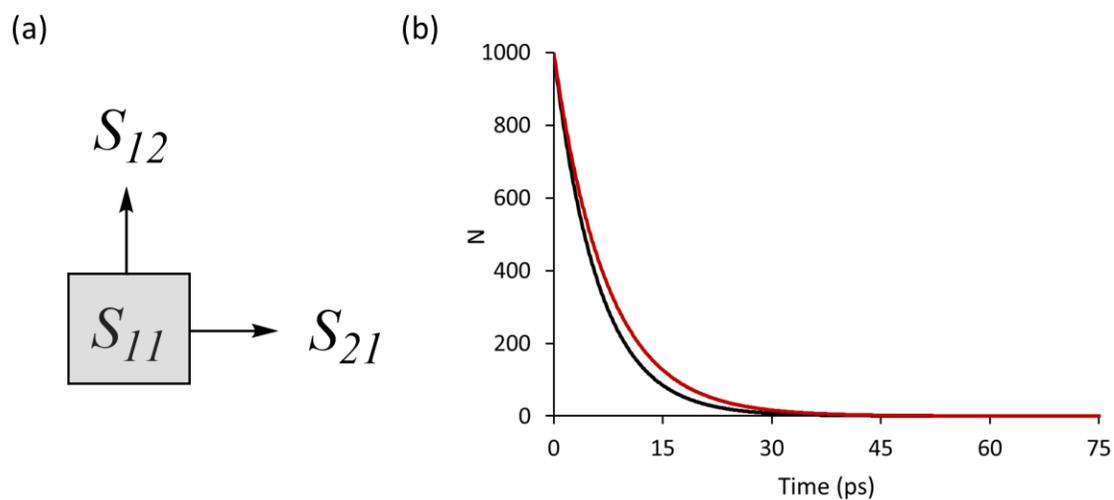
The resultant  $\Delta G(x_{R_1})$  is shown in Figure 5.27 and as expected, has a single minimum and a maximum at  $x_{R_1} = 1$ .



**Figure 5.27** Gibbs energy profile  $\Delta G(x_{R_1})$  along the  $x_{R_1}$  coordinate for the  $R_1$  domain.

The rate constant for exit from the  $R_1$  state, obtained from the activation barrier using the Eyring equation (eq. 5.5), was used to generate the first order kinetic curve for  $x_{R_1}$ . For comparison, the kinetic curve for exit from the  $R_1$  domain on the original 2D GES

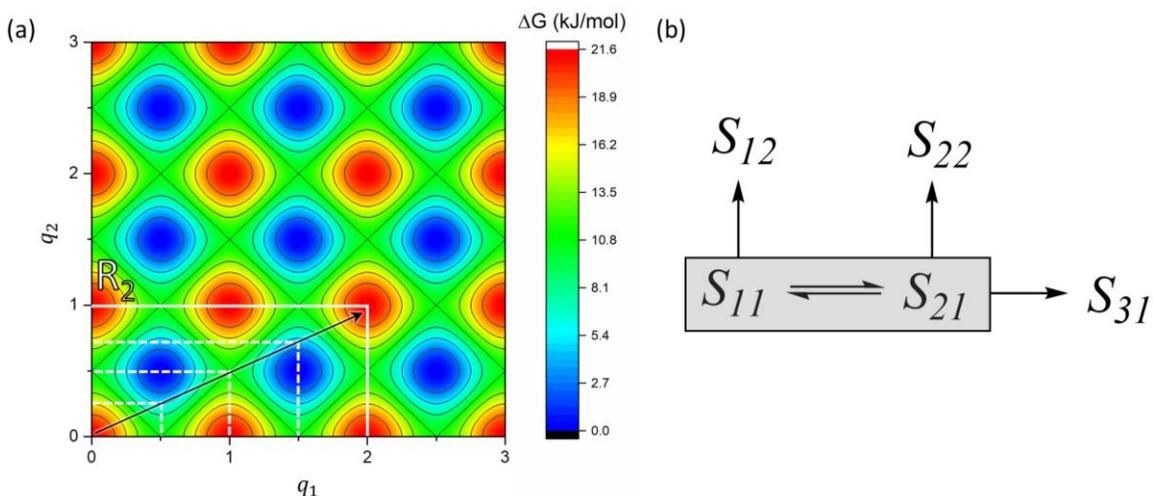
was calculated according to the kinetic scheme of Figure 5.28(a). The 1D and 2D results, shown in Figure 5.28(b), are in good agreement.



**Figure 5.28** (a) Kinetic scheme for exit from the  $R_1$  domain (represented by the grey square) on the 2D GES. (b) Comparison of the kinetic curves derived from the 2D kinetic network (black) and from the 1D  $\Delta G(x_{R_1})$  (red).

### 5.4.3. A two-minima reactant domain

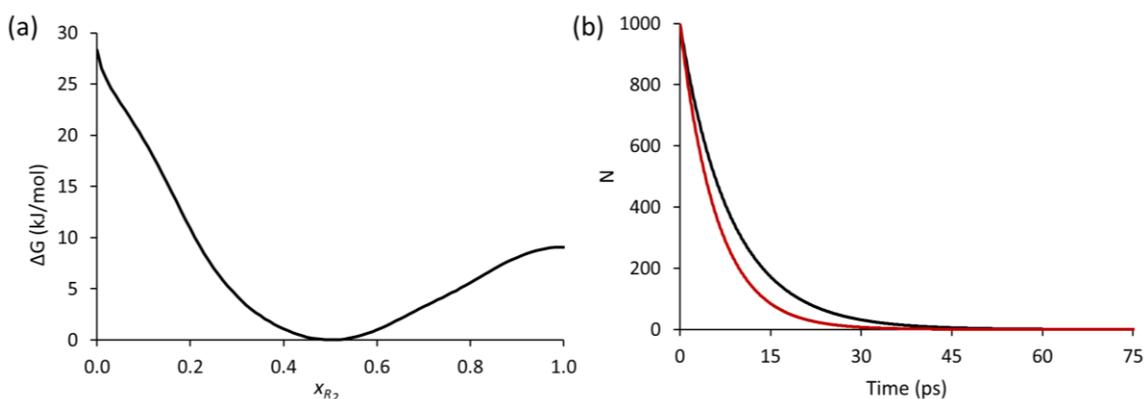
A two-minima reactant domain  $R_2$ , that encompasses both the  $S_{11}$  and the  $S_{21}$  minima is shown in Figure 5.29(a). The kinetic network on the 2D surface for the  $R_2$  boundary therefore follows the scheme in Figure 5.29(b).



**Figure 5.29** (a) The boundary of the  $R_2$  domain (solid white line) and hypersurfaces of constant  $x_{R_2}$  (dashed white lines). (b) The 2D reaction network for exit from the  $R_2$  domain, represented by the grey rectangle.

The reaction coordinate for the  $R_2$  domain was defined according to eq. 5.17.

$$x_{R_2} = \max\left(\frac{q_1}{2}, q_2\right) \quad (5.17)$$

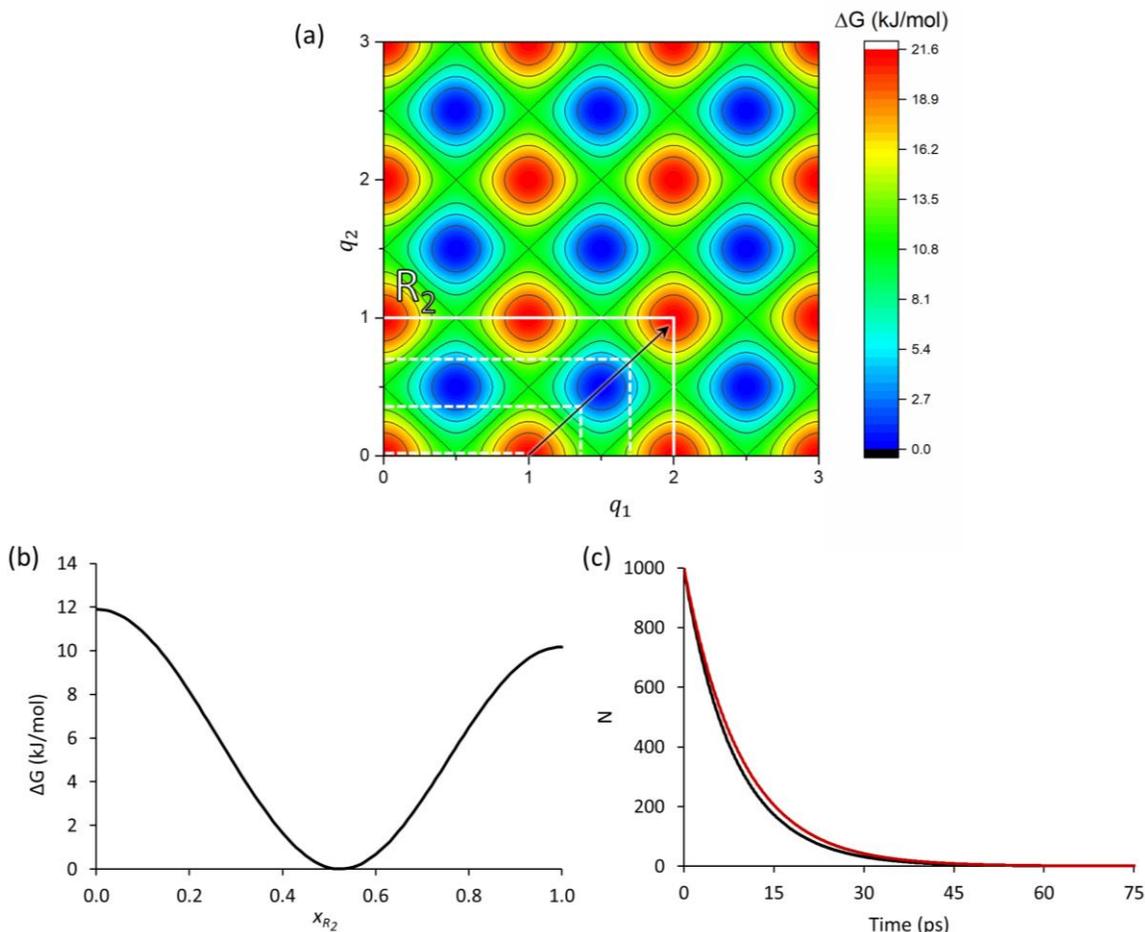


**Figure 5.30** (a) Gibbs energy profile  $\Delta G(x_{R_2})$  for the  $R_2$  domain along the  $x_{R_2}$  coordinate defined by eq. 5.17. (b) A comparison of the kinetic curve derived from the full reaction network of Figure 5.29(b) (black) with the two-state kinetics from  $\Delta G(x_{R_2})$  (red) shows a poorer quality match.

The Gibbs energy profile  $\Delta G(x_{R_2})$  was calculated along the  $x_{R_2}$  coordinate and is shown in Figure 5.30(a). It is very similar in shape to the profile obtained for the  $x_{R_1}$  coordinate. The 1D and 2D kinetics for exit from the  $R_2$  domain was calculated as in the

case of the  $R_1$  domain discussed in section 5.4.2. The resultant kinetic curves, shown in Figure 5.30(b), do not agree well with each other. The agreement improves substantially, as can be seen from Figure 5.31, when the  $x_{R_2}$  coordinate is redefined by shifting the reference point from (0,0) to (1,0) using eq. 5.18.

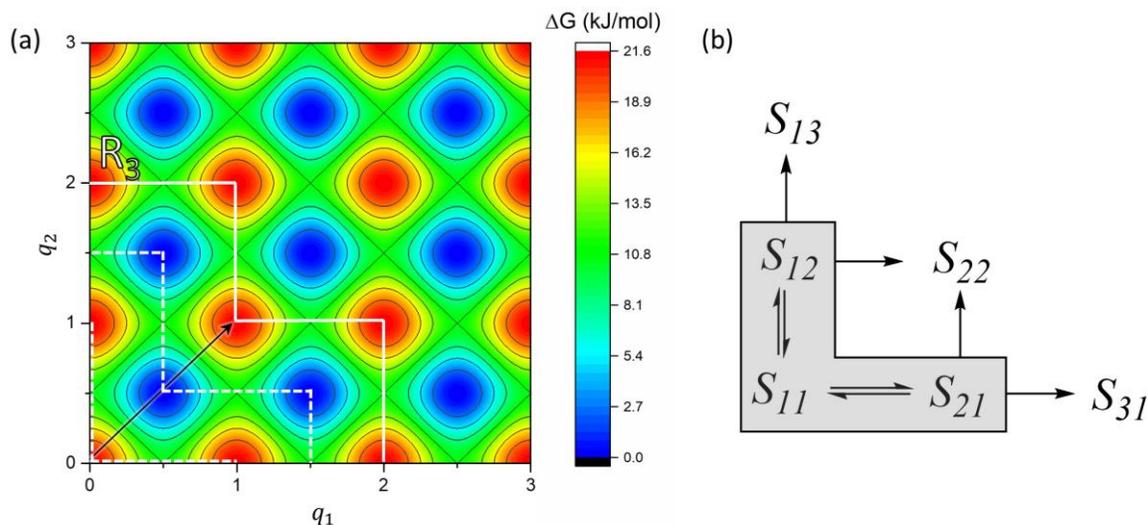
$$x_{R_2} = \max(q_1 - 1, q_2) \quad (5.18)$$



**Figure 5.31** (a) The hypersurfaces of constant  $x_{R_2}$  (white dashed lines) for the redefined coordinate of eq. 5.18. The boundary of the  $R_2$  domain (solid line) has not changed. (b) Gibbs energy profile  $\Delta G(x_{R_2})$  for the  $R_2$  domain along redefined  $x_{R_2}$  coordinate. (c) A comparison of the kinetic curves derived from the 2D reaction network of Figure 5.29(b) (black) and from the new 1D profile  $\Delta G(x_{R_2})$  (red).

#### 5.4.4. A three-minima reactant domain

An L-shaped three-minima reactant domain  $R_3$ , shown in Figure 5.32(a), contains the  $S_{11}$ ,  $S_{21}$  and  $S_{12}$  minima. There are four possible exits from this domain: from state  $S_{21}$  to state  $S_{31}$ , from state  $S_{21}$  to state  $S_{22}$ , from state  $S_{12}$  to state  $S_{13}$  and from state  $S_{12}$  to state  $S_{22}$ . The exit kinetics from the  $R_3$  domain thus follows the reaction network scheme of Figure 5.32(b).

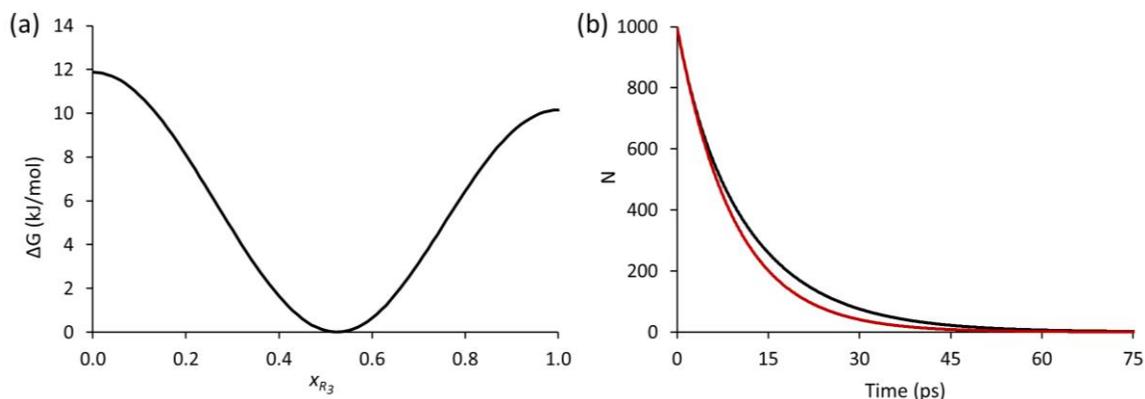


**Figure 5.32** (a) The boundary of the  $R_3$  domain (white solid line), encapsulating the  $S_{11}$ ,  $S_{21}$  and  $S_{12}$  minima. Hypersurfaces of constant  $x_{R_3}$  are shown as white dashed lines. (b) The 2D reaction network for exit from the  $R_3$  domain (grey).

A more complex shape of the boundary results in a more convoluted definition of the reaction coordinate  $x_{R_3}$  (eq. 5.19) that produces a system of constant  $x_{R_3}$  hypersurfaces parallel to the domain boundary.

$$x_{R_3} = \begin{cases} q_1, & q_1 > q_2 + 1 \\ q_2, & q_2 + 1 \geq q_1 > q_2 \\ q_1, & q_1 + 1 \geq q_2 > q_1 \\ q_2, & q_2 \geq q_1 + 1 \end{cases} \quad (5.19)$$

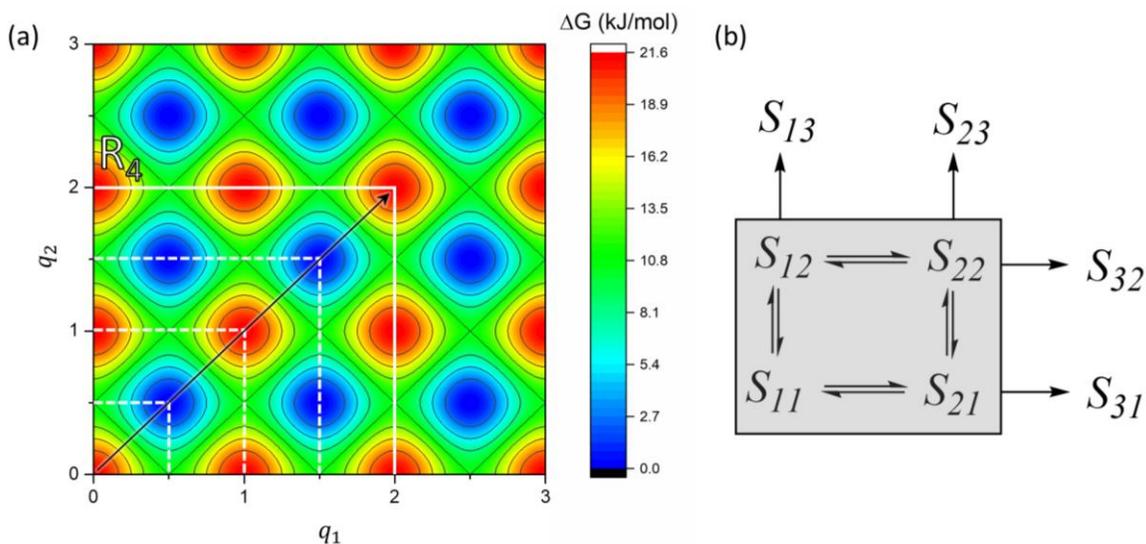
The one-dimensional Gibbs energy profile obtained for hypersurfaces of constant  $x_{R_3}$  is shown in Figure 5.33(a). This profile produces a kinetic curve for depletion of reactant in Figure 5.33(b) that is somewhat faster than that obtained for the 2D reaction network.



**Figure 5.33** (a) The Gibbs energy profile  $\Delta G(x_{R_3})$  for the  $R_3$  domain along the  $x_{R_3}$  coordinate defined by eq. 5.19. (b) A comparison of the kinetic curves derived from the 2D reaction network of Figure 5.29(b) (black) and from the 1D profile  $\Delta G(x_{R_3})$  (red).

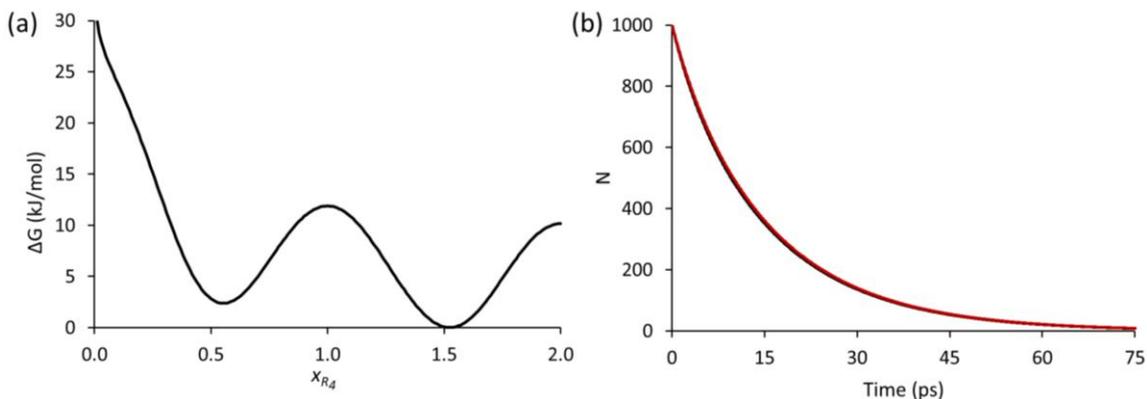
### 5.4.5. A four-minima reactant domain

A four-minima reactant domain  $R_4$ , which contains the  $S_{11}$ ,  $S_{21}$ ,  $S_{12}$  and  $S_{22}$  states is shown in Figure 5.34(a). Four exits from the  $R_4$  domain are possible: from  $S_{21}$  to  $S_{31}$ , from  $S_{12}$  to  $S_{13}$ , from  $S_{22}$  to  $S_{32}$  and from  $S_{22}$  to  $S_{23}$ . The kinetic reaction network for exit from this domain is shown in Figure 5.34(b).



**Figure 5.34** (a) The boundary of the  $R_4$  domain (white solid line), encapsulating the  $S_{11}$ ,  $S_{21}$ ,  $S_{12}$  and  $S_{22}$  minima. Hypersurfaces of constant  $x$  are shown as white dashed lines. (b) The 2D reaction network for exit from the  $R_4$  domain (grey square).

Since the shape of the boundary of the  $R_4$  domain is similar to that of the  $R_1$  domain, the same definition of the reaction coordinate can be used (eq. 5.14). The resultant Gibbs energy profile is shown in Figure 5.35(a) and contains two minima, interconversion between which should be considered reversible in accordance with the underlying 2D kinetic scheme of Figure 5.34(b).

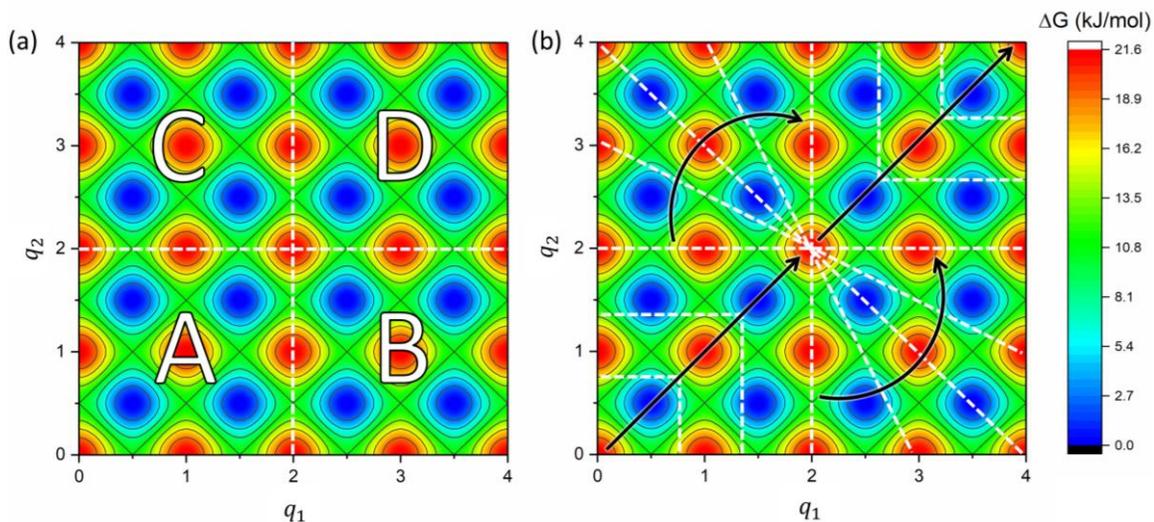


**Figure 5.35** (a) The Gibbs energy profile for the  $R_4$  domain. (b) A comparison of the kinetic curves derived from the 2D reaction network (black) and from the 1D profile (red).

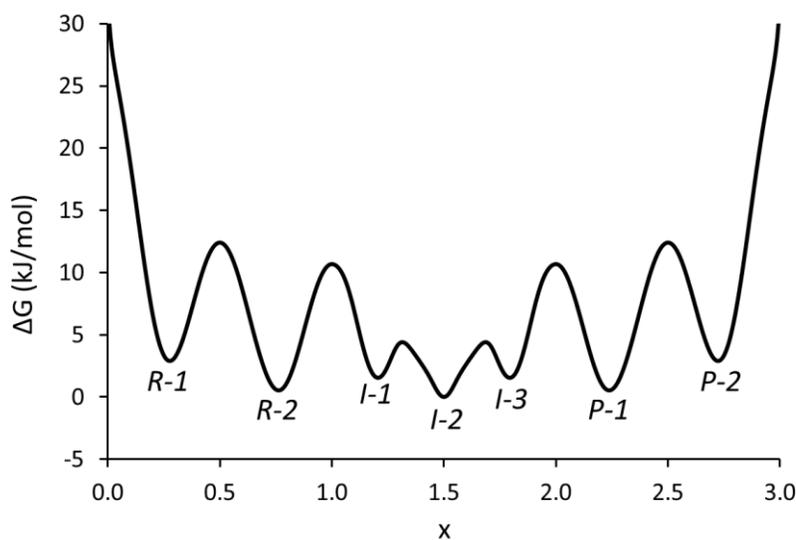
The kinetic curves obtained for the 1D three-state model and the 2D reaction network are compared in Figure 5.35(b), showing an excellent agreement.

#### 5.4.6. The piecewise coordinate

We now divide the model 16-minima surface into 4 domains consisting of 4 minima each, as shown in Figure 5.36, and apply the piecewise coordinate  $x$  of eq. 5.12 with  $a = b = 2$  to describe transition from the reactant domain  $A$  to the product  $D$  domain via the intermediate domains  $B$  and  $C$ .



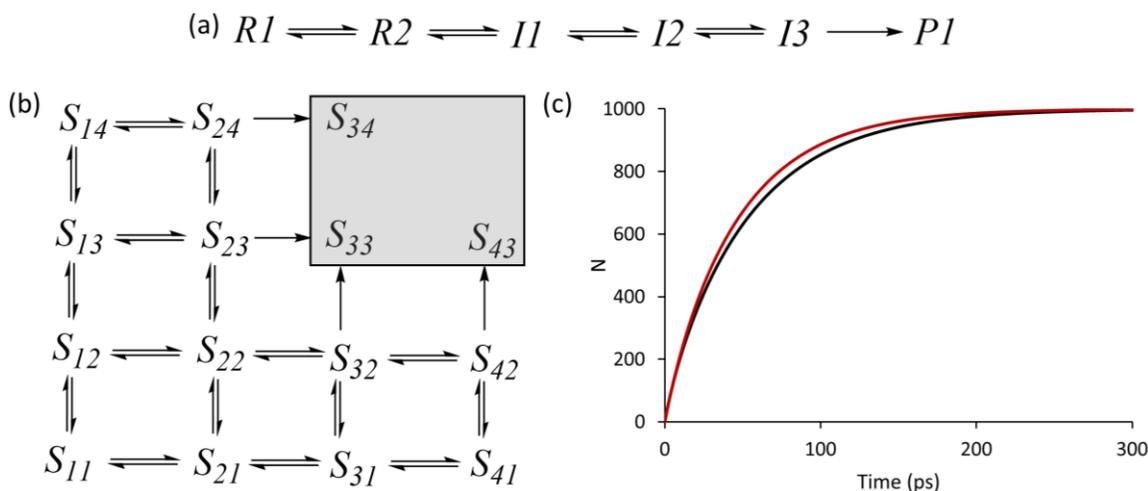
**Figure 5.36** (a) The reactant *A*, intermediate *B,C* and product *D* domains for the 16-state surface, each domain consisting of 4 minima. (b) The hypersurfaces of constant  $x$  for the piecewise coordinate.



**Figure 5.37** The Gibbs energy profile  $\Delta G(x)$  along the piecewise coordinate  $x$ , with minima labelled R-1 to P-2.

The resulting Gibbs energy profile  $\Delta G(x)$  is shown in Figure 5.37, and contains multiple minima. The first two of these minima (R-1 and R-2) are reactant minima and match those of the 1D profile of Figure 5.35. Similarly, the last two minima (P-1 and P-2) belong to the product, and the remaining minima I-1, I-2, and I-3 constitute the intermediate state.

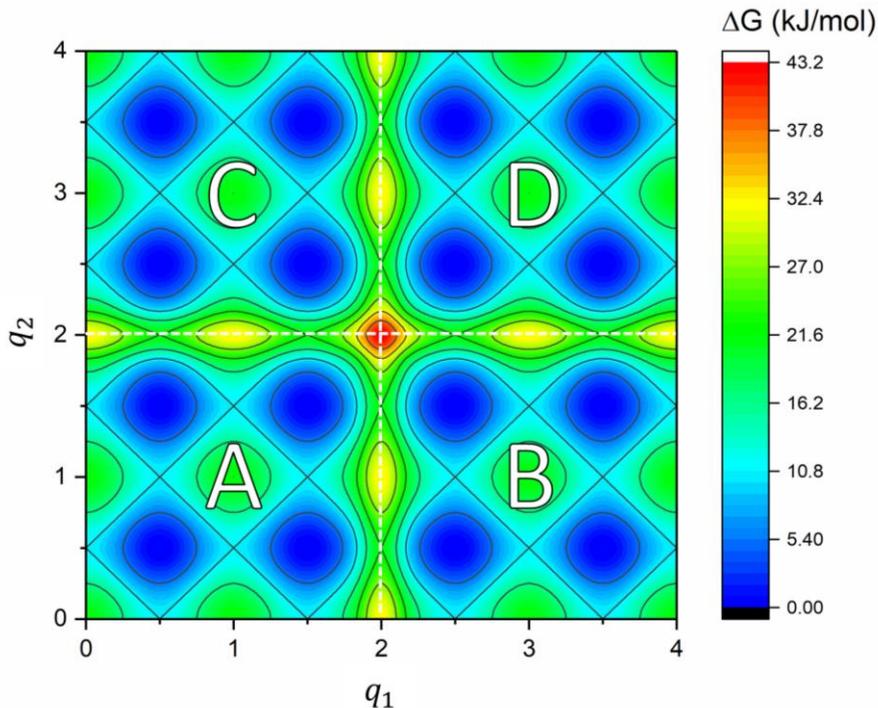
As the  $x$  coordinate covers the entire surface, we can use the  $\Delta G(x)$  profile to calculate kinetic curves both for the depletion of the reactant state  $A$ , and for accumulation of the product state  $D$ . The depletion kinetics from state  $A$  has already been discussed in the preceding section. To complement that result, the kinetics for accumulation of the product state  $D$  was calculated for the 1D and 2D reaction networks shown in Figure 5.38(a) and 5.38(b). The resultant kinetic curves, shown in Figure 5.38(c), display a good match.



**Figure 5.38** Kinetic schemes for accumulation of product for (a) the Gibbs energy profile  $\Delta G(x)$  and (b) the 2D surface. The states that comprise the D domain are highlighted with the grey square. (c) The resulting kinetic curves obtained from the 1D profile (red) and the 2D surface (black) agree well with each other.

In real systems, GESs are naturally split into multi-minima domains by higher barriers, and as a result the time required to cross between domains is longer than the time of interconversion between individual local states within a domain. To mimic this, we modified our 16-well surface by increasing the barriers between the four domains, as shown in Figure 5.39, by the addition of Gaussian functions along the border between domains. The resultant  $\Delta G(q_1, q_2)$  is described by eq. 5.20, with  $V = 10.8$  kJ/mol,  $\alpha_{q_1} = \alpha_{q_2} = 50$  and  $q_1^o = q_2^o = 2$ .

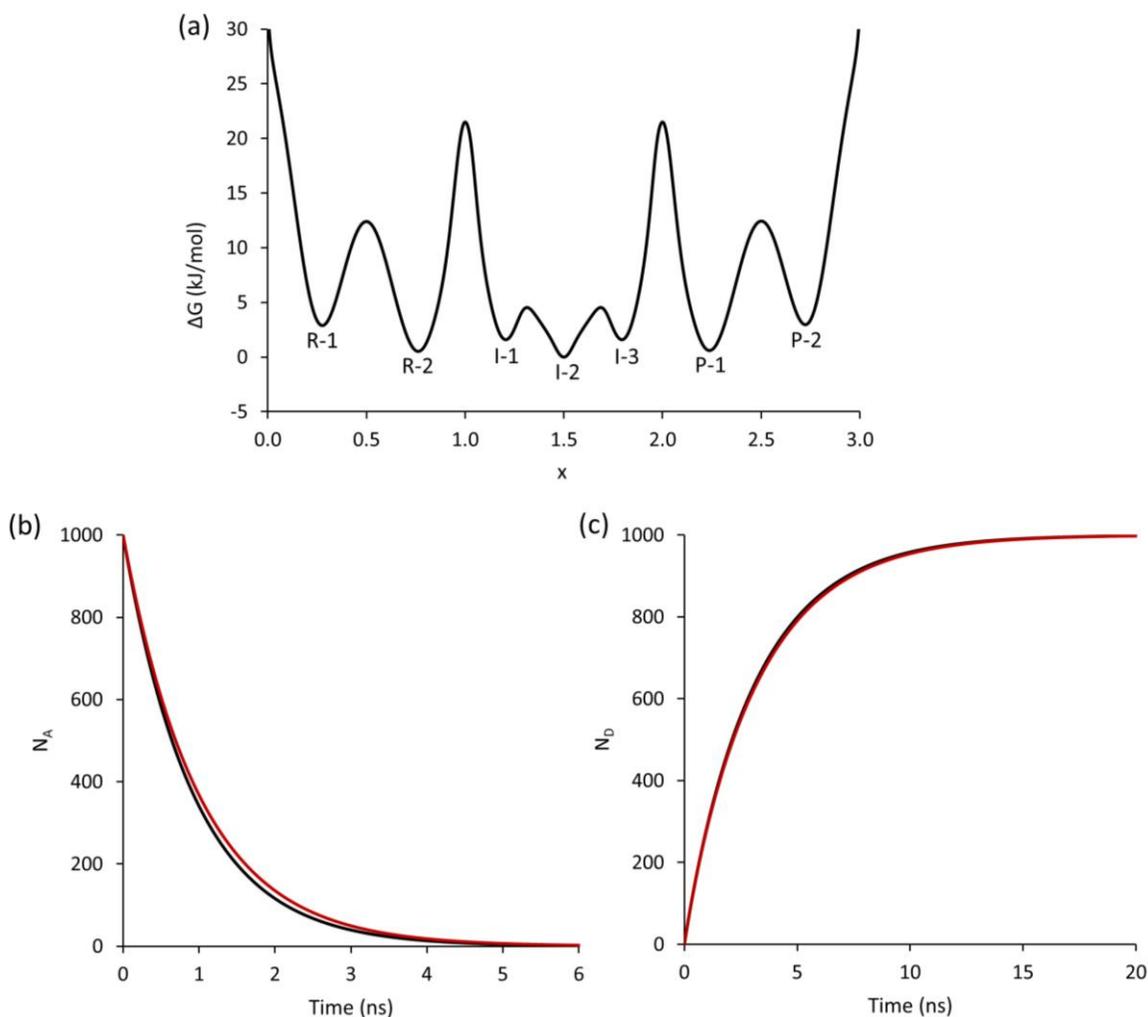
$$\Delta G(q_1, q_2) = 0.5V(2 - \cos(2\pi q_1 - \pi) - \cos(2\pi q_2 - \pi)) + V(e^{-\alpha_{q_1}(q_1 - q_1^o)^2} + e^{-\alpha_{q_2}(q_2 - q_2^o)^2}) \quad (5.20)$$



**Figure 5.39** The modified 16-well surface with increased activation barriers between the domains.

The rate constants for intra-domain transitions have not changed and are still equal to  $8.23 \times 10^{10} \text{ s}^{-1}$  but the rate constants for transition between domains are now equal to  $1.08 \times 10^9 \text{ s}^{-1}$ , nearly two orders of magnitude smaller.

The piecewise coordinate (eq. 5.12) was applied to this system in the same manner as above. The resultant Gibbs energy profile  $\Delta G(x)$  is shown in Figure 5.40(a), and looks similar to the original Gibbs energy profile of Figure 5.37(a) but with a notably higher barrier separating reactant, product, and intermediate states. A comparison of the kinetic curves obtained for the full reaction network and along the piecewise coordinate shows an excellent agreement for both the depletion of the reactant (Figure 5.40b) and accumulation of the product (Figure 5.40c). Thus, the piecewise coordinate is potentially a good reaction coordinate for systems with multiple minima and multiple reaction pathways.



**Figure 5.40** (a) The Gibbs energy profile  $\Delta G(x)$  along the piecewise coordinate  $x$ , with minima labelled R-1 to P-2. These minima correspond to the labels used in Figure 5.37(b). The kinetic curves obtained using 1D (red) and 2D (black) reaction networks show good agreement for both (b) depletion of reactant and (c) accumulation of product.

## 5.5. Conclusions

Reaction coordinates are difficult to define for flexible systems with multiple reaction pathways due to the multiplicity of reaction paths that must be taken into account. The Best-Hummer procedure provides a useful instrument for optimization of collective variables, but does not guarantee kinetic adequacy of the generated 1D profiles, which strongly depends on the chosen functional form of the optimized collective variable. We proposed a new reaction coordinate based on the  $d_\infty$  metric and demonstrated its efficiency in describing kinetics in selected model systems.

## Chapter 6.

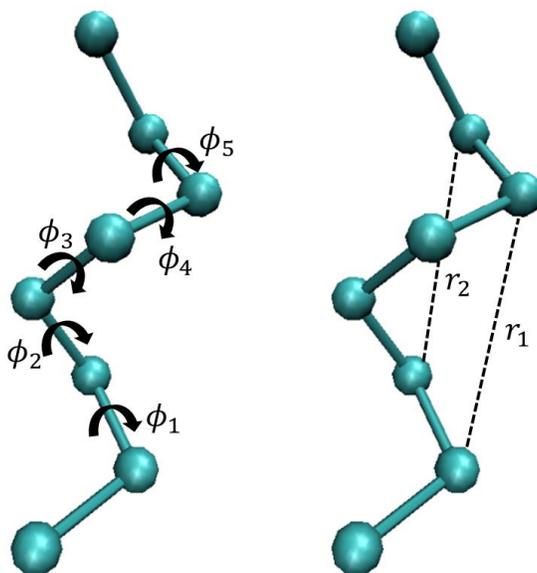
### Towards applications to biological systems

While biological systems are an attractive target for the volume profile method of Chapter 4, the high level of complexity in these systems, particularly in the definition of the reaction coordinate, makes the calculation of a volume profile quite challenging. In this concluding chapter, some preliminary results are presented for biological systems ranging in complexity from model peptides to full proteins.

#### 6.1. Reaction coordinate for the unfolding of a model helix

##### 6.1.1. System

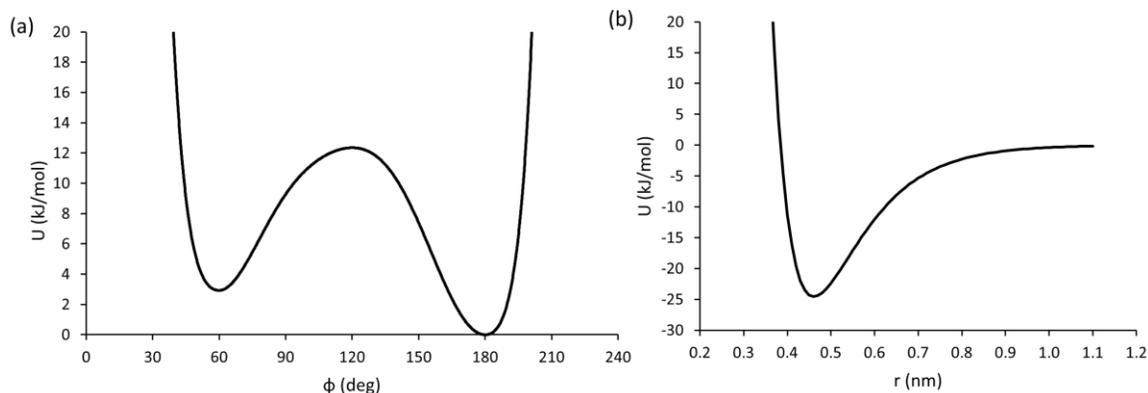
One of the most basic three-dimensional structures found in proteins is the alpha helix.<sup>28</sup> Therefore, as a first step we chose to investigate the form of the reaction coordinate for unfolding of a model helix, illustrated in Figure 6.1.



**Figure 6.1** The model helix system showing the five dihedral angles  $\phi_1 - \phi_5$  (left) and the two hydrogen bonds  $r_1$  and  $r_2$  (right).

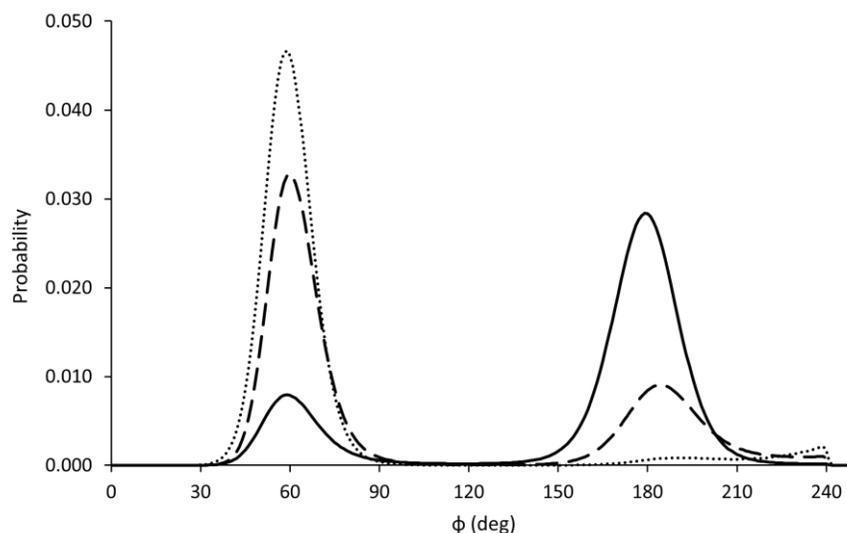
The bond and angle force field parameters are the same as those used for the model chain.<sup>209</sup> The 5 dihedral angles in the system ( $\phi_1$  to  $\phi_5$ ) are described by the double well

potential of Figure 6.2(a), with minima at  $60^\circ$  and  $180^\circ$  degrees. If all dihedral angles are in the  $60^\circ$  minimum, the 8-atom chain twists itself into a two-turn helical conformation. The favourability of this folded state is increased by the addition of two “hydrogen bonds” in the system ( $r_1$  and  $r_2$ ). These interactions are described by the Morse-like potential of Figure 6.2(b), with interaction energy of  $-50$  kJ/mol and equilibrium bond length  $r_e = 0.46$  nm. This potential was fitted to match the combined Lennard-Jones and Coulombic interaction energy of the NH-O hydrogen bond in the OPLS force field.



**Figure 6.2** (a) Double well potential for the dihedral angles in the model helix. The unfolded state at  $180^\circ$  is slightly lower in energy than the folded state at  $60^\circ$ . (b) The Morse bond potential for the two hydrogen bonds in this system.

This system was restrained to its folded state and was inserted in a box of 1000 Lennard-Jones particles at  $P = 1$  kbar,  $T = 300$  K. Once equilibrated, the restraint was released and a total of 2425 10.0 ns trajectories were initiated from the folded ensemble for use in the reaction coordinate analysis. Figure 6.3 shows probability distribution functions for each dihedral angle obtained over the course of these trajectories.



**Figure 6.3** Probability distribution functions for the five dihedral angles in the model helix grouped by symmetry:  $\phi_1$  and  $\phi_5$  (solid line),  $\phi_2$  and  $\phi_4$  (dashed line) and  $\phi_3$  (dotted line).

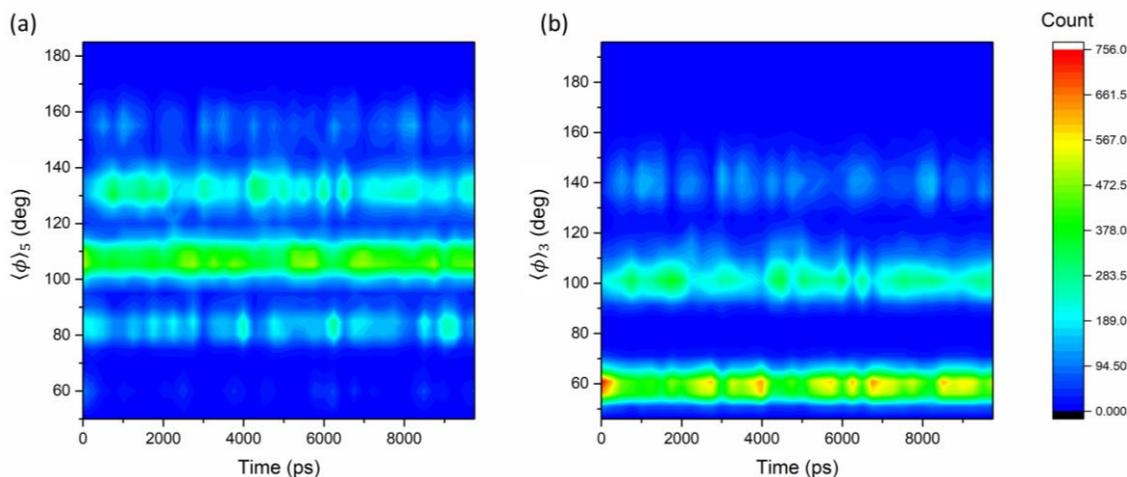
The outermost angles,  $\phi_1$  and  $\phi_5$ , favour the  $180^\circ$  conformation because they do not affect either of the “hydrogen bonded” distances, while the innermost angle  $\phi_3$  favours the  $60^\circ$  conformation as it affects both  $r_1$  and  $r_2$ . Angles  $\phi_2$  and  $\phi_4$  have population in both the  $60^\circ$  and  $180^\circ$  states since each of them affects only one of those distances.

### 6.1.2. Definition of reactant and product states

We chose to use the Best-Hummer optimization procedure,<sup>225</sup> described in section 5.2.2, to determine the optimal reaction coordinate for this model system. To do so, we needed to determine the important structural parameters for the system in order to define nonoverlapping reactant and product domains in terms of these parameters, and also to select a suitable functional form to combine these parameters into an optimizable collective variable  $x$ . Torsion angles, with their distinct preference for either  $60^\circ$  or  $180^\circ$  conformation, naturally lend themselves as the required structural parameters.

To define the reactant and product domains, we studied the time evolution of the average of all 5 dihedral angles  $\langle\phi\rangle_5$  and of only the middle 3 angles  $\langle\phi\rangle_3$ . A surface plot showing the population density of  $\langle\phi\rangle_5$  and  $\langle\phi\rangle_3$  vs. time for ten 10.0 ns trajectories is shown in Figure 6.4. While there is no clear transition region in the  $\langle\phi\rangle_5$  plot, as the density appears to be evenly spread across all values of  $\langle\phi\rangle_5$ , we do see a clear drop in density

in the region between  $70^\circ$  and  $90^\circ$  for  $\langle\phi\rangle_3$ . The reactant (folded) domain can thus be defined by  $\langle\phi\rangle_3 < 70^\circ$  and the product (unfolded) domain by  $\langle\phi\rangle_3 > 90^\circ$ .



**Figure 6.4** Population density for (a)  $\langle\phi\rangle_5$  and (b)  $\langle\phi\rangle_3$  as a function of time. For  $\langle\phi\rangle_3$ , there is a clear region of low population density between  $70^\circ < \langle\phi\rangle_3 < 90^\circ$ .

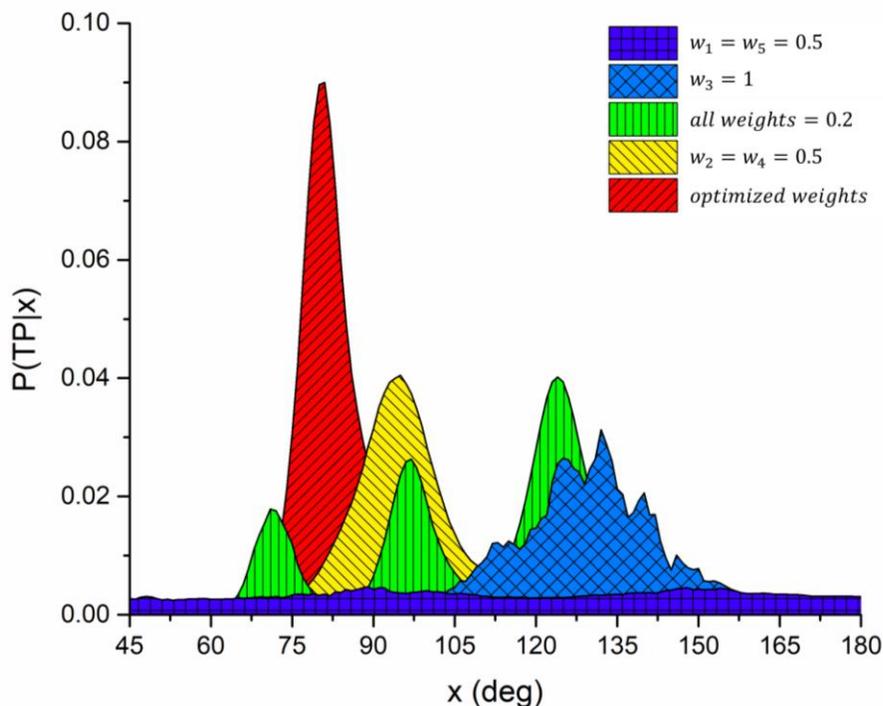
### 6.1.3. Trial reaction coordinate: Weighted average angle

For the optimizable reaction coordinate  $x$ , we chose to test the weighted average of the 5 dihedral angles (eq. 6.1). The initial weights were randomly selected and then optimized to satisfy the symmetry condition  $w_1 = w_5$  and  $w_2 = w_4$  reflecting the symmetry of the system. The optimal weights were determined as those that yielded the sharpest  $P(TP|x)$  distribution, calculated using eq. 6.2, via the same procedure as was used for the triatomic model system of Chapter 5.

$$x = \sum_{i=1}^5 w_i \phi_i \quad (6.1)$$

$$P(TP|x) = \frac{P(x|TP)P(TP)}{P(x)} \quad (6.2)$$

The optimal weights were found to be  $w_1 = w_5 = 0.02$ ,  $w_2 = w_4 = 0.25$  and  $w_3 = 0.46$ . The resulting  $P(TP|x)$  distribution for this weighting is shown in Figure 6.5, along with the distributions for some sample non-optimal weights. The optimized weights have the highest and most narrow distribution, with almost all of the probability density existing in the transition region between  $70^\circ < x < 90^\circ$ .

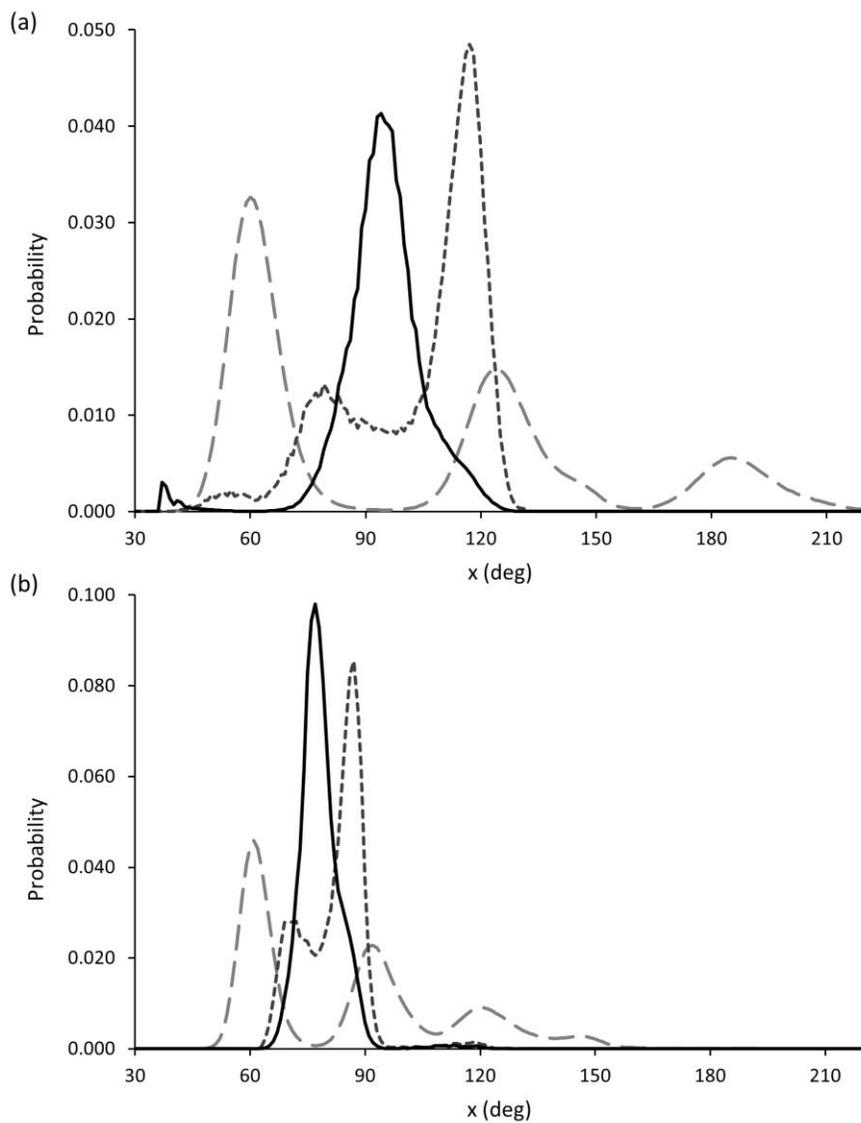


**Figure 6.5**  $P(TP|x)$  distributions for a selection of different weights: equal weighting (green), only  $\phi_1$  and  $\phi_5$  (dark blue), only  $\phi_2$  and  $\phi_4$  (yellow) only  $\phi_3$  (light blue) and the optimized weighting (red).

The low weights obtained for angles  $\phi_1$  and  $\phi_5$  are not unexpected, given the high population of the unfolded state relative to the folded state in their probability distributions of Figure 6.3. Therefore, the functional form of the optimizable reaction coordinate  $x$  was changed to exclude  $\phi_1$  and  $\phi_5$  (eq. 6.3), with symmetry condition  $w_2 = w_4$ .

$$x = \sum_{i=2}^4 w_i \phi_i \quad (6.3)$$

Optimization of this redefined coordinate produced weights  $w_2 = w_4 = 0.25$  and  $w_3 = 0.50$ . At first sight, the high weight placed on  $\phi_3$  appears inconsistent with the fact that it is largely biased towards the  $60^\circ$  conformation, as can be seen from the probability distributions of Figure 6.3, and is not therefore expected to be important for a description of the unfolding process. However, a comparison of  $P(TP|x)$  and its components  $P(x)$  and  $P(x|TP)$  for the optimized case of  $w_2 = w_4 = 0.25$ ,  $w_3 = 0.50$  and for the case of  $w_2 = w_4 = 0.5$ ,  $w_3 = 0$ , shown in Figure 6.6, explains this artifact as a consequence of the contraction of the scale of the  $x$  coordinate.



**Figure 6.6** The  $P(TP|x)$  distribution (black) and the component distributions  $P(x)$  (light grey dashed) and  $P(x|TP)$  (dark grey dashed) for (a) sample weights  $w_2 = w_4 = 0.5$ ,  $w_3 = 0$  and (b) the optimized weights.

Indeed, since  $\phi_3$  is practically constant, it does not contribute to the displayed probability distributions, which therefore show a remarkable similarity of shapes. At the same time, inclusion of  $\phi_3$  into  $x$  with a weight of  $w_3$  reduces the range of  $x$  over which  $P(TP|x)$  evolves by a factor of  $(1 - w_3)$ . Since Best-Hummer optimization aims at reducing the width of this range, it inevitably favors greater values of  $w_3$ .

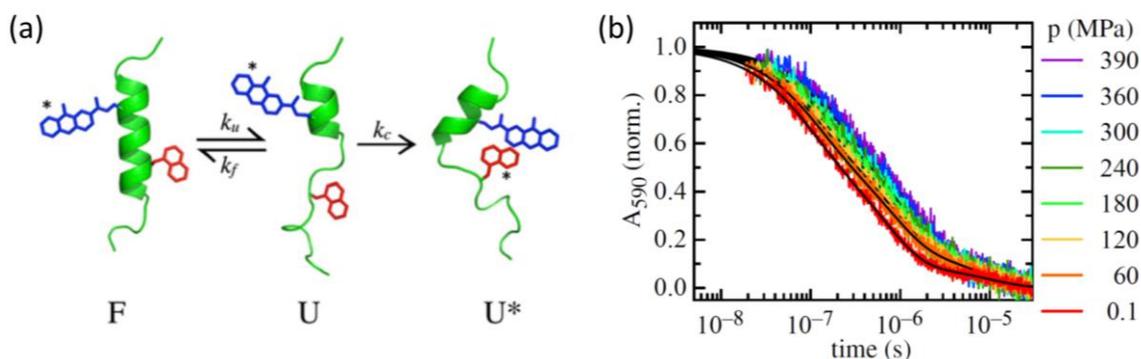
The above artifact is of a general nature - inclusion of any improper coordinate in the definition of  $x$  can lead to an erroneous composition of the reaction coordinate if the target function of the Best-Hummer optimization is the width of the transient path range.

For systems that are more complex than this model helix, it may not be possible to know which coordinates to include in  $x$  without prior knowledge of the full reactive dynamics. Therefore, this method should be used with caution when applied to biological systems.

## 6.2. Kinetic network analysis for an $\alpha$ -helical peptide

### 6.2.1. The Kiefhaber peptide

Thomas Kiefhaber and co-workers<sup>269–271</sup> undertook an extensive kinetic study of a 21-residue alanine-based peptide which undergoes a helix to coil transition upon the application of pressure. The primary sequence for this peptide (referred below as the Kiefhaber peptide) is  $^+\text{NH}_3\text{-(A)}_5\text{-(AAARA)}_3\text{-A-NH}_2$  with the labelling groups xanthone and naphthalene added to the alanine side chain in residues 7 and 13 respectively (Figure 6.7a).



**Figure 6.7** (a) The three-state model used to describe the unfolding of the Kiefhaber peptide. The helix must unfold in the region between the xanthone (blue) and naphthalene (red) groups in order for triplet-triplet energy transfer (TTET) to occur. (b) Bi-exponential experimental kinetic data for TTET at different pressures. Figure reprinted from ref. 271 with permission.

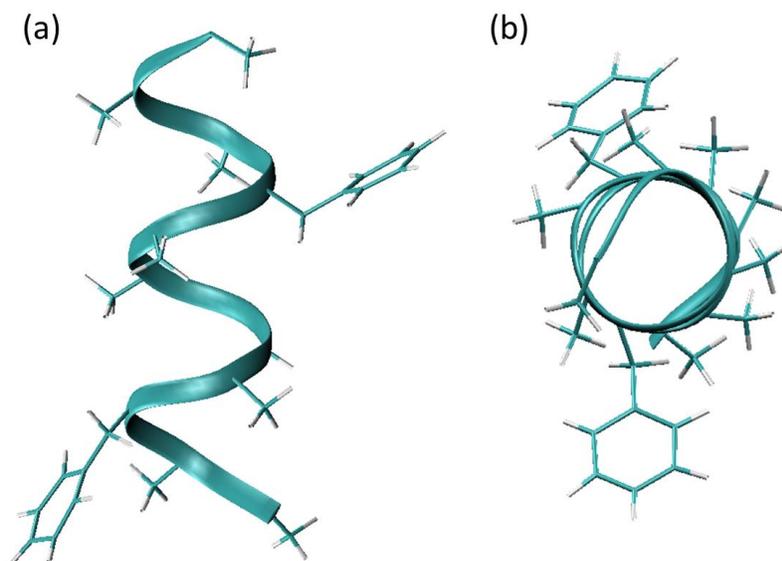
These large aromatic groups function as reporters of the folding process via triplet-triplet energy transfer (TTET),<sup>272</sup> where xanthone acts as the donor and naphthalene the acceptor. The xanthone triplet state absorbs very strongly at 590 nm. However, when it is in van der Waals contact with naphthalene, the two groups can undergo TTET, during which the triplet excited state is transferred from the xanthone to the naphthalene, resulting in a decrease in the absorbance at 590 nm. In the folded peptide, the reporter groups are oriented on opposite sides of the helix. For TTET to occur, the helix must unfold at least in the region between the two reporter groups.

The observed TTET kinetics for the Kiefhaber peptide is bi-exponential (Figure 6.7b) and was described by Kiefhaber *et al.* in terms of a three state model, with the unfolding occurring as the first (reversible) step. Local folding and unfolding of the helix occurs with rate constants  $k_u$  and  $k_f$ , and contact formation and subsequent TTET between the two groups in the unfolded state is described by  $k_c$ . The experimental activation volumes for this process have been determined<sup>271</sup> as  $\Delta V_u^\ddagger = 3.3 \pm 0.3$  cm<sup>3</sup>/mol,  $\Delta V_f^\ddagger = 1.7 \pm 0.5$  cm<sup>3</sup>/mol and  $\Delta V_c^\ddagger = 3.6 \pm 0.4$  cm<sup>3</sup>/mol and the reaction volume for the F  $\rightarrow$  U step is  $\Delta V = -1.6 \pm 0.6$  cm<sup>3</sup>/mol.

The relatively small size of this peptide, combined with the experimentally available activation volumes make this system a good target for the application of our volume profile method. While the choice of reaction coordinate for the contact formation step is straightforward (it can be identified as the distance between the reporter groups), the reaction coordinate for the unfolding step is less obvious. There are no direct measurements for this step in the experiment, and all information pertaining to it is inferred from the bi-exponential behavior of the TTET kinetics. MD simulations can fill this void and provide additional useful information on the mechanism of the folding/unfolding step.

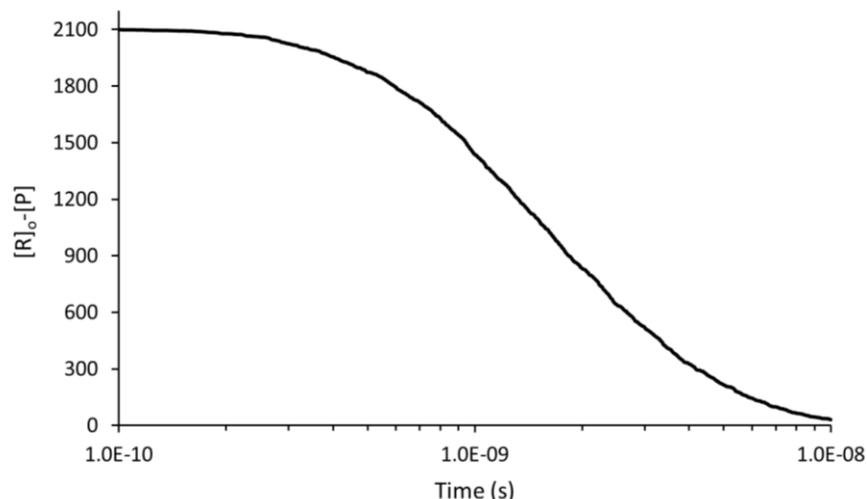
### 6.2.2. Model system: Polyalanine

To simplify the problem, we started with a smaller model closely mimicking the Kiefhaber peptide: an 11-residue polyalanine peptide with phenylalanines at positions 3 and 8, shown in Figure 6.8. The phenyl groups of the two phenylalanine residues provide proxies for the reporter groups used to measure the TTET kinetics. For the purpose of our simulations, when the phenyl groups were in van der Waals contact ( $< 0.35$  nm), it was taken to be a TTET “quenching” event and the unfolding reaction was considered to be completed.



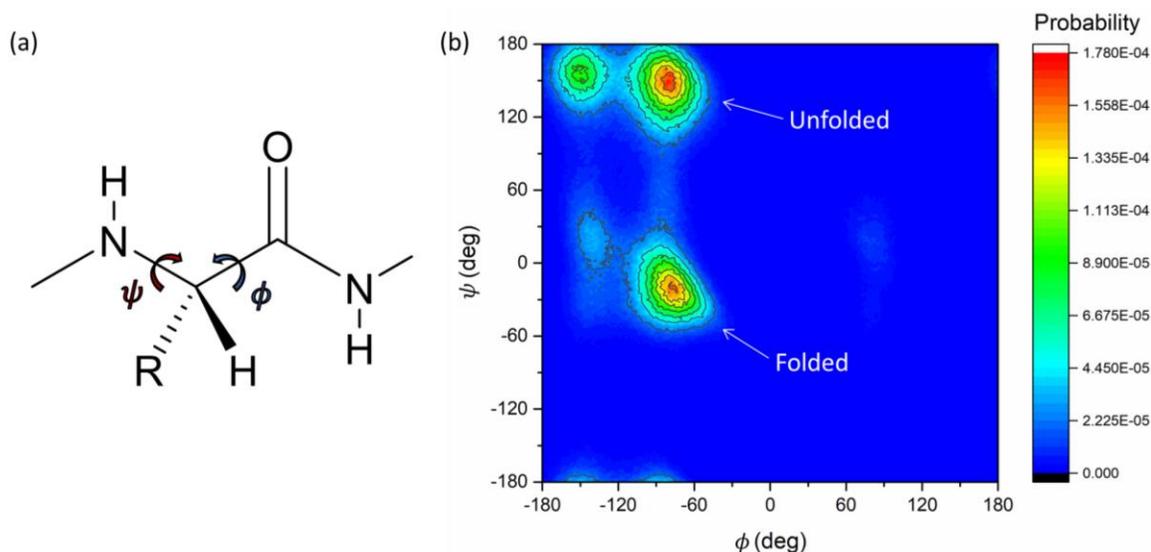
**Figure 6.8** The structure of the model polyaniline peptide showing the orientation of the phenylalanine R groups in the folded state: (a) a side view and (b) a view from above.

The quenching kinetics of this system was simulated using the first exit method.<sup>209</sup> For these simulations, the peptide was equilibrated with solvent in the folded state, with torsion angles along the protein backbone restrained to their energy-minimized values. The restraints were then released, and the system was allowed to evolve. The time it took for the phenyl groups to come within 0.35 nm of each other was recorded, and the simulation was repeated several thousand times to obtain good statistics. In the experimental setup, Kiefhaber *et al* were not observing the product state directly; rather they were using absorbance to monitor the disappearance of all species except for the product. To mimic their kinetic setup, we followed the evolution of  $[R]_o - [P]$ , where  $[R]_o$  is the sum of initial concentrations of all non-product species. The resulting plot of  $[R]_o - [P]$  vs. time is shown in Figure 6.9. The kinetic curve is polyexponential, as in the case of the experimentally-measured quenching kinetics of the Kiefhaber peptide.



**Figure 6.9 Polyexponential kinetic curve for the quenching of the polyaniline peptide obtained via first exit MD simulations.**

The MD trajectories were also used to generate a Ramachandran plot,<sup>28</sup> which is a 2D probability distribution function in  $\phi$  and  $\psi$  coordinates (protein backbone angles, defined in Figure 6.10a). Since only the 5 residues of the core region between the two reporter phenylalanine residues contribute to the TTET kinetics, only the backbone angles of these amino acids were used.



**Figure 6.10 (a) The  $\psi$  and  $\phi$  backbone torsion angles. (b) Ramachandran plot for the polyaniline helix. The folded and unfolded domains are cleanly separated from each other along the  $\psi$  axis.**

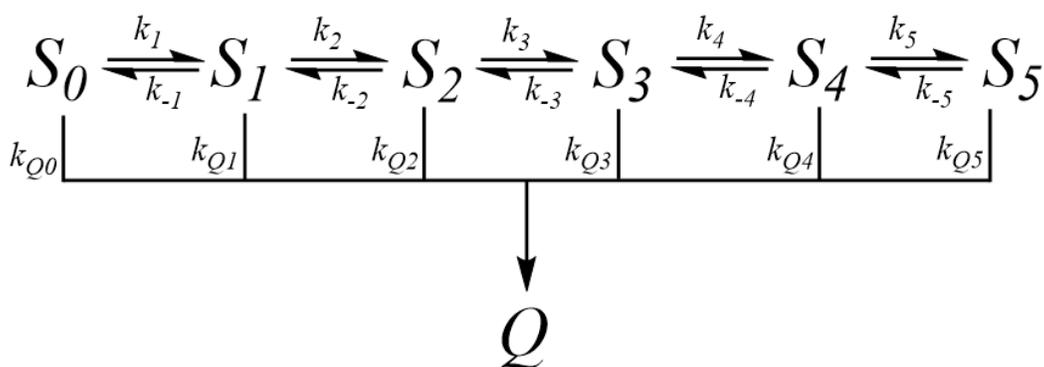
The resulting plot, shown in Figure 6.10(b), has four distinct sub-domains bundled into two clusters visibly separated along the  $\psi$  coordinate: the unfolded state at  $\psi \geq 100^\circ$

and the folded state at  $-70^\circ \leq \psi \leq 50^\circ$ . Although the  $\phi$  angles prefer *gauche* conformations in both states, the *anti* conformations of  $\phi$  do become more prevalent in the unfolded state.

### 6.2.3. Coarse-grained kinetic network for polyalanine unfolding

A detailed kinetic network to describe conformational changes in the polyalanine backbone would involve all of its  $\phi$  and  $\psi$  angles, but a network built on such a large number of states does not seem practical for the purpose of identifying the appropriate reaction coordinate. To reduce the size of the network, we made two assumptions: (1) that only the changes in  $\psi$  angles should be considered and (2) that only the total number rather than a detailed distribution of unfolded  $\psi$  angles in the core region is sufficient to describe the unfolding process in rough terms.

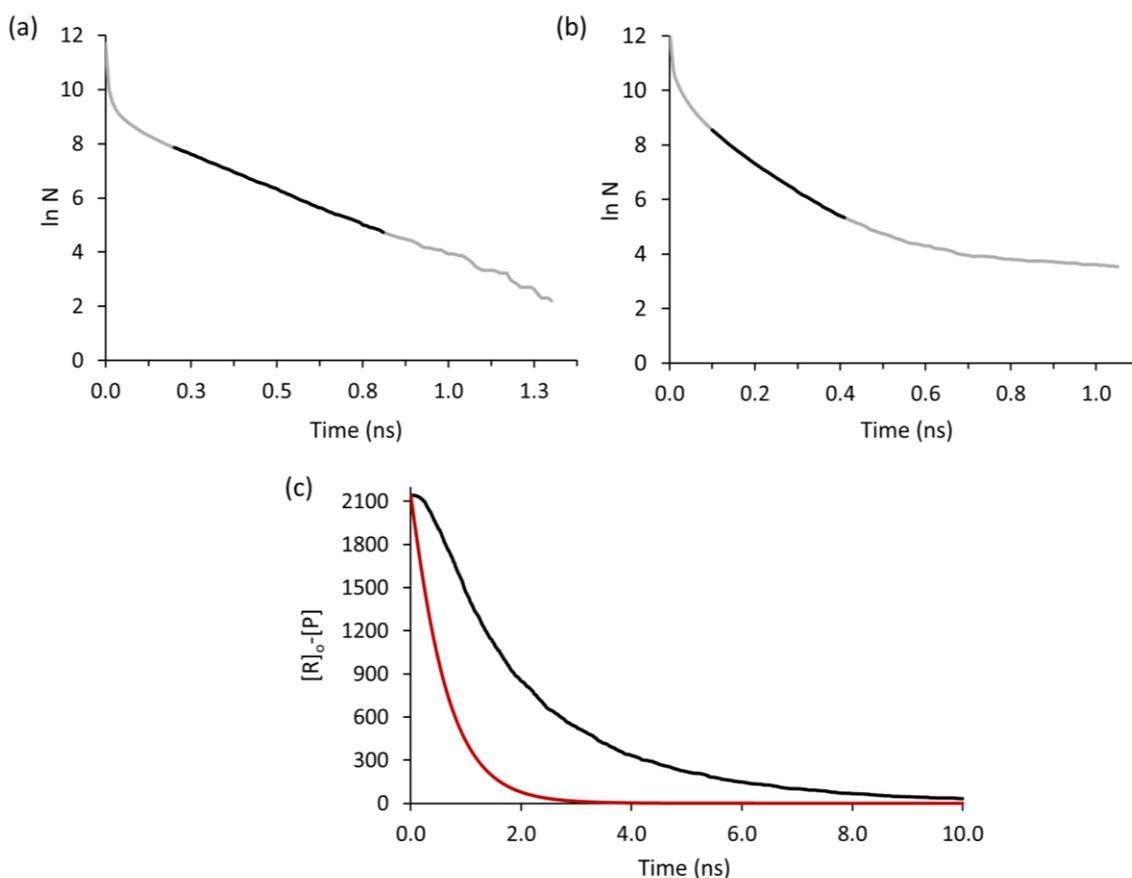
The first assumption is consistent with the fact evident from the Ramachandran plot of Figure 6.10(b) that folded and unfolded states differ by the value of  $\psi$  but not by the value of  $\phi$ . The second assumption is consistent with the design of the peptide, where van der Waals contacts of phenyl groups become possible as soon as at least one of  $\psi$  angles in the core region switches to the unfolded conformation. Accordingly, the coarse-grained kinetic network shown in Figure 6.11 involves 6 unquenched states, labeled  $S_0$  to  $S_5$  by the number of unfolded  $\psi$  angles in the core region, and one quenched state  $Q$ . We also assume that only transitions between states that differ by one unfolded  $\psi$  angle are allowed, and that the irreversible quenching reaction can occur from any of these states.



**Figure 6.11** Coarse-grained kinetic network for the unfolding and quenching of polyalanine.

To calculate the partial rate constants needed for this model, we utilized the trajectories previously used to calculate the kinetic curve of Figure 6.9 and determined the

distribution of times required for the system to exit out of each of the states  $S_0 - S_5$  into other states, and survival probability curves were generated for each possible transition. These curves, some samples of which are shown in Figure 6.12(a) and Figure 6.12(b), exhibit polyexponential behavior, likely due to nonequilibrium dynamic effects. Removal of the initial fast crossing period of the graphs yields a quasi-linear plot for some transitions (e.g. Figure 6.12a), the slope of which can be used to obtain the rate constant. Other transitions are highly non-linear (e.g. Figure 6.12b), and a linear fit is likely not appropriate. The kinetic curve for  $[R]_o - [P]$  based on the coarse-grained kinetic network of Figure 6.11 displays a significantly faster kinetics than that based on the MD simulations of section 6.2.2 (Figure 6.12c).

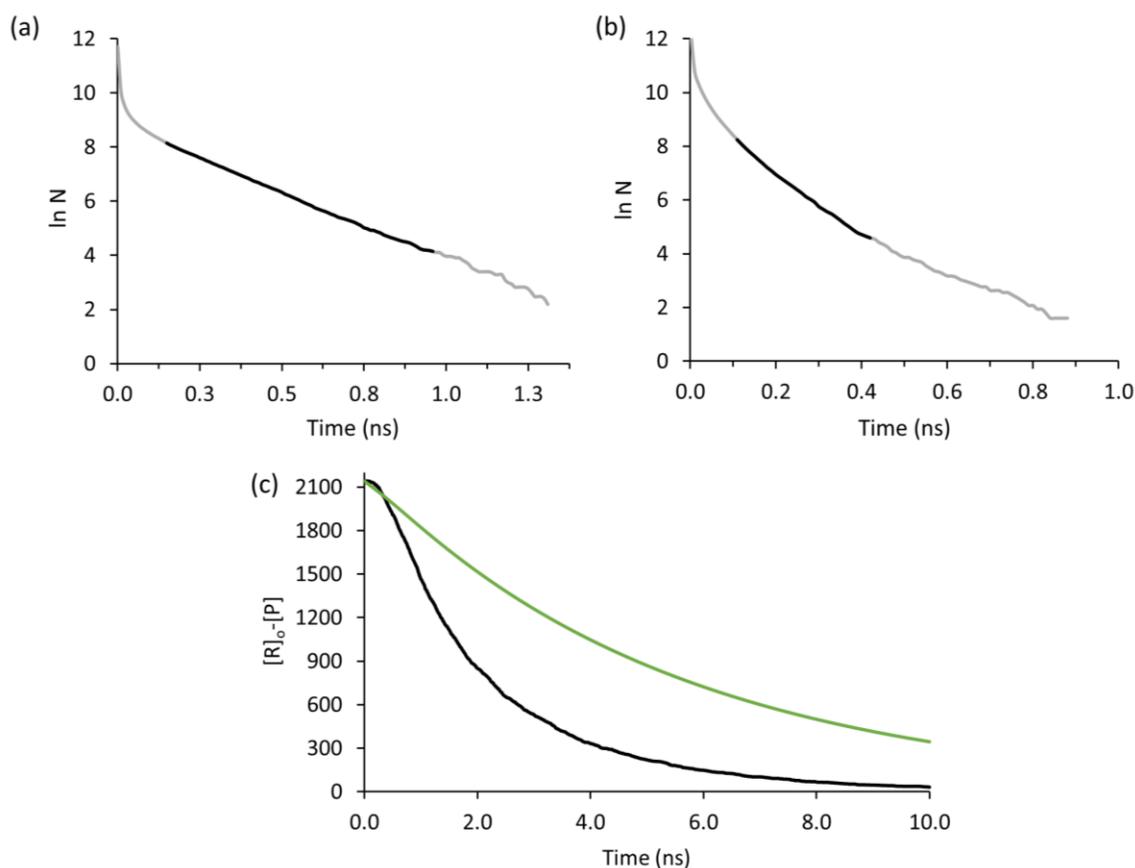


**Figure 6.12** (a) Survival probability curve for the  $S_0 \rightarrow S_1$  transition. The darker portion of the curve was used to determine the rate constant for the transition. (b) The survival probability curve for the  $S_4 \rightarrow S_3$ , showing highly nonlinear behaviour. (c) Comparison of the kinetic curves for the coarse-grained kinetic network (red) and MD simulations (black).

In an attempt to improve the accuracy of the determination of the partial rate constants, we decided to somewhat change our strategy. Survival probability plots were

now calculated for residence in each state  $S_0 - S_5$ , regardless of the direction of exit from the state. The net rate constants obtained from these kinetic curves were then partitioned into the individual rate constants for each transition,  $k_i$ , by eq. 6.4, where  $n_i$  is the number of transitions into state  $i$  from the state of interest, and  $N$  is the total number of transitions. While this method does help linearize the survival probability plots for the higher states (Figure 6.13a and b), there is still considerable curvature in the plots.

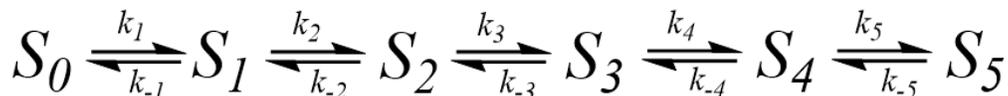
$$k_i = \frac{n_i}{N} k \quad (6.4)$$



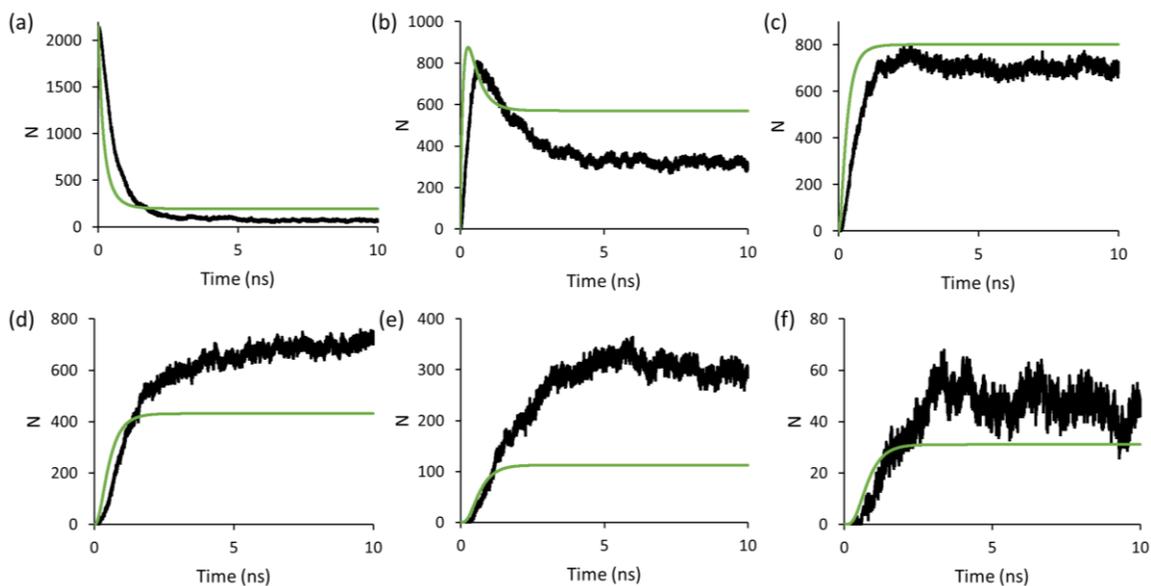
**Figure 6.13** Survival probability curves for residence in (a) the  $S_0$  state and (b) the  $S_4$  state. The darker portion of each curve was used to determine the rate constant. (c) Comparison of the kinetic curves for the new partitioning method (green) and MD simulations (black).

When the resulting rate constants are used to calculate  $[R]_0 - [P]$ , the network curve is now slower than the MD one (Figure 6.13c). A comparison of the values of the rate constants obtained by the survival probability method vs. the partitioning method showed that the rate constants for transitions between states were relatively unchanged,

while the rate constants for quenching changed by 2-3 orders of magnitude. This led us to believe that the quenching constants were the major source of error. To investigate this, we re-calculated the concentrations for all of our states using the revised coarse-grained kinetic network of Figure 6.14, where quenching transitions were ignored. As shown in Figure 6.15, some states are fairly well described by our model, while others predict completely wrong behavior.



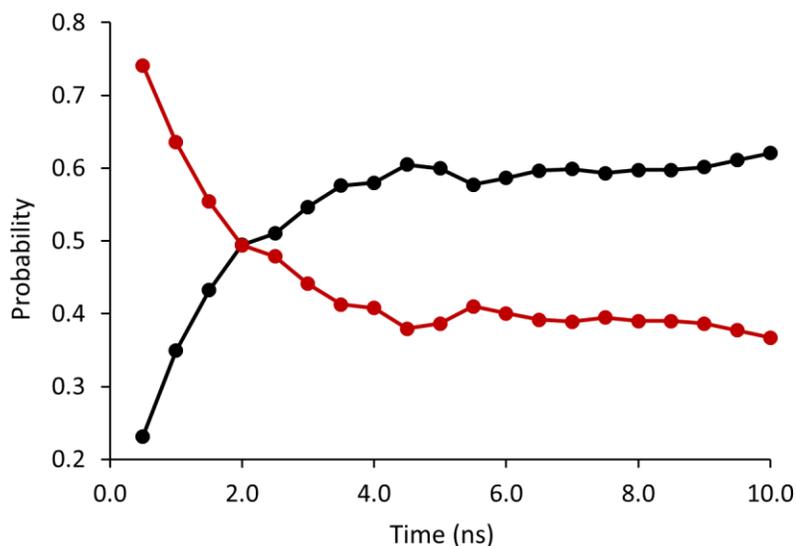
**Figure 6.14** Markov chain model where quenching transitions are not considered.



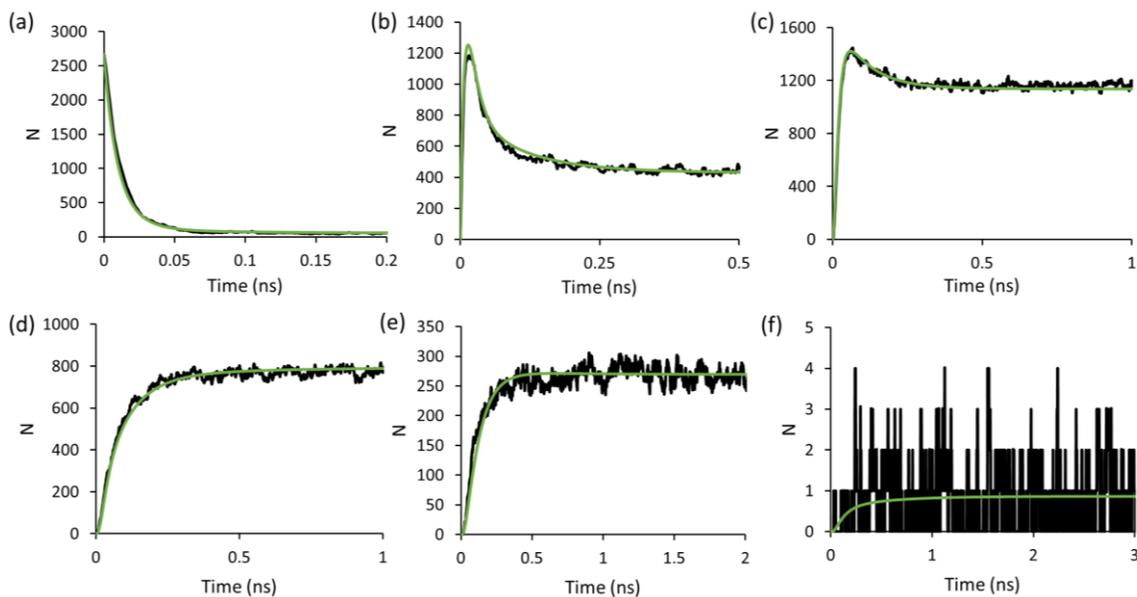
**Figure 6.15** Population vs. time curves for (a)  $S_0$ , (b)  $S_1$ , (c)  $S_2$ , (d)  $S_3$ , (e)  $S_4$  and (f)  $S_5$  based on the kinetic network of Figure 6.14 where quenching is no longer considered. For all plots, the black curve is from MD simulations and the green curve is from the kinetic network.

This may be due to relatively poor sampling of these states in our MD simulations, but it also may be due to a more fundamental flaw in the model. If the frequencies of specific transitions are plotted as a function of simulation time, we can see that these frequencies do not remain constant over the course of the simulation (Figure 6.16). Early in the trajectory the system is more likely to transition back into the state it just entered from, rather than transitioning forward into the next state. Our states are very broadly defined, so this time-dependent behavior can be understood by the system requiring time to diffuse within the current state from the border of the past state to the border of the next

state. This diffusion through multidimensional states likely takes place via the changing of  $\phi$  angles between the subdomains of each state.



**Figure 6.16** Probabilities of exit from state  $S_2$  forward into state  $S_3$  (black) and back into state  $S_1$  (red).



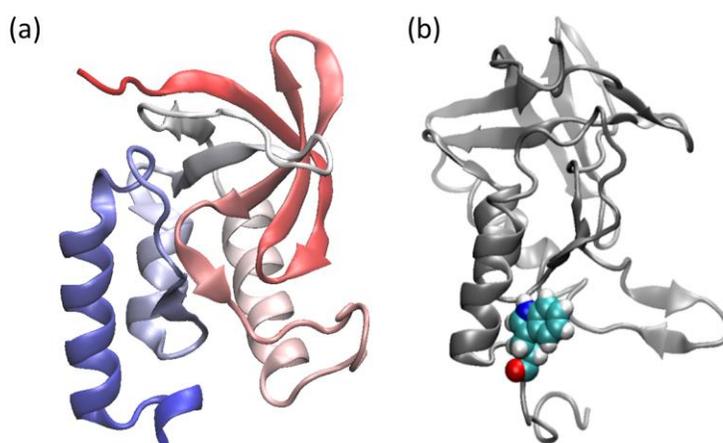
**Figure 6.17** Population vs. time curves for (a)  $S_0$ , (b)  $S_1$ , (c)  $S_2$ , (d)  $S_3$ , (e)  $S_4$  and (f)  $S_5$ . The black curves are obtained from MD simulations and the green curves are from the kinetic network of Figure 6.14 applied to the model helix of section 6.1.

This hypothesis was indirectly supported by application of the kinetic network of Figure 6.14 to the 8-atom model helix of section 6.1. As there are only 5 dihedral angles

in that system, internal diffusion within states is not expected to be important. The rate constants for transitions between states were determined using the partitioning method (eq. 6.4). As shown in Figure 6.17, the kinetic curves for the populations of the 6 states determined from kinetic network exactly match the curves determined directly from MD simulations. Therefore, it seems that a finer definition of the kinetic network states for polyalanine is necessary to better reproduce MD kinetics.

### 6.3. Staphylococcal nuclease

Our final and most complex system of interest is the staphylococcal nuclease (snase) protein, a 149 residue protein whose structure is shown in Figure 6.18. The pressure unfolding behaviour of snase has been well-studied by Catherine Royer and co-workers using both tryptophan fluorescence techniques<sup>40</sup> and small angle X-ray scattering (SAXS).<sup>44</sup>



**Figure 6.18** (a) The staphylococcal nuclease protein (PDB: 1SNO) shown in red at the N-terminus and blue at the C-terminus. (b) The tryptophan residue in snase is shown in the coloured space-filling representation.

The use of two experimental techniques allows for probing of both the local and global unfolding kinetics. The fluorescence emission maximum  $\lambda_{max}$  of tryptophan (TRP) is highly sensitive to the local environment.<sup>273,274</sup> Snase contains a single TRP residue, located at the base of one of the alpha helices (Figure 6.18b). In the folded state, the TRP residue is stacked against its neighbouring side chains, with one edge of the indole ring accessible to solvent. In this state, the fluorescence emission maximum is located at  $\lambda_{max} = 340$  nm. Upon unfolding, the TRP residue becomes accessible to solvent, causing

a red-shift in the  $\lambda_{max}$  which can be registered spectroscopically and used to monitor kinetics. Thus, this technique provides information about unfolding that occurs local to the TRP residue. SAXS, on the other hand, allows for calculation of the radius of gyration ( $R_g$ ) of the protein via the scattering vectors. As discussed in section 5.2.1,  $R_g$  is defined as the root mean square deviation of each atom from the center of mass, and it provides a measure of relative size of the protein. A change in  $R_g$ , therefore, can be viewed as a measure of the global folding/unfolding of the protein. For both experiments, a two-state model was used to determine the rate constants  $k_u$  and  $k_f$ , and the obtained activation volumes agree within experimental error (Table 6.1).

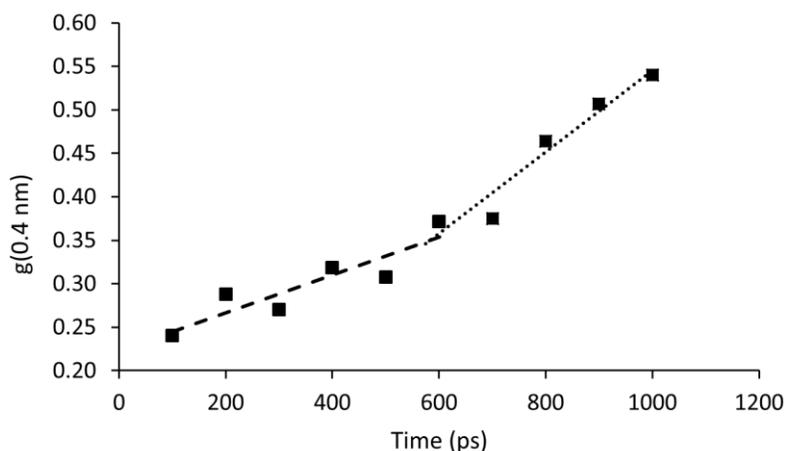
**Table 6.1 Experimental  $\Delta V^\ddagger$  and  $\Delta V$  data for snase.**

Technique	$\Delta V_f^\ddagger(\text{cm}^3\text{mol}^{-1})$	$\Delta V_u^\ddagger(\text{cm}^3\text{mol}^{-1})$	$\Delta V_u(\text{cm}^3\text{mol}^{-1})$	Ref.
Fluorescence	$92 \pm 4$	$20 \pm 3$	$-77 \pm 8$	40
	$99 \pm 12$	-	-81.6	44
SAXS	$89 \pm 10$	-	-	44

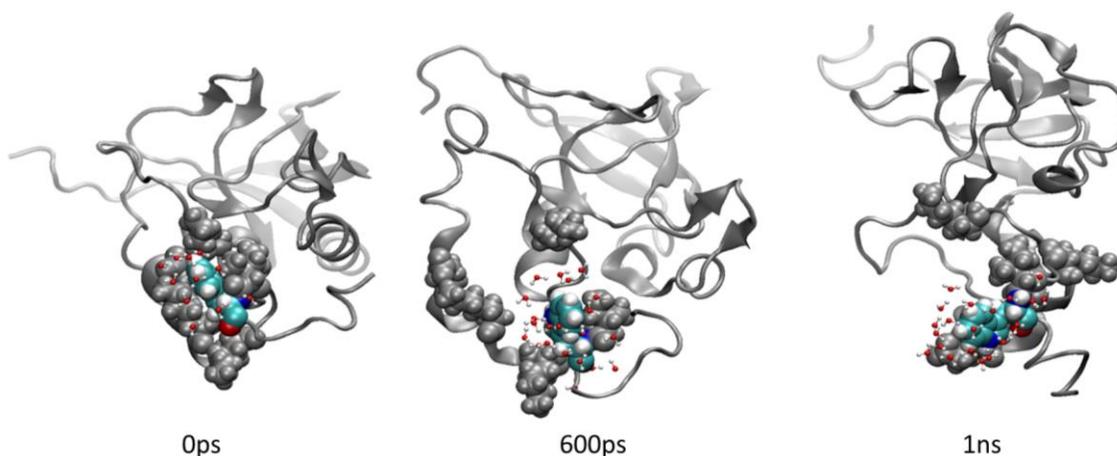
Of the two available experimental techniques, the SAXS data provides the most easily interpretable reaction coordinate, since  $R_g$  is readily calculated via MD simulations. To calculate a volume profile for the unfolding of snase, the value of the chosen reaction coordinate must be restrained for each volume calculation. It is possible to restrain  $R_g$  in GROMACS using the PLUMED<sup>275</sup> plugin, but this is achieved through the application of a substantial restraining force which, in view of a considerable size of the protein molecule, results a significant overestimation of pressure calculated using the virial. This, in turn, leads to a significant over-inflation of the volume of the simulated system aimed at reducing this perceived high pressure. As a test, we calculated the reaction volume for the unfolding of snase using  $R_g$  as a reaction coordinate. The value of  $R_g$  was restrained to the experimentally determined values for the folded and unfolded states: 1.8 nm and 2.9 nm respectively.<sup>44</sup> As a result, the  $\Delta V_u = 3047 \text{ cm}^3 \text{ mol}^{-1}$ , calculated by our method as the volume difference between the unfolded and folded states, was far larger than the experimental value. Therefore, in the context of our volume profile approach, the use of  $R_g$  as a reaction coordinate faces serious technical difficulties until a more appropriate method of restraining  $R_g$  is identified.

As an alternative, we turned to the TRP fluorescence experiments. This technique probes the solvent environment around TRP. Therefore, our reaction coordinate must be

somehow linked to the access of water to the face of the indole ring. To investigate this, we performed a high temperature unfolding simulation of snase. Over the course of 1.0 ns, the temperature was raised from 300 K to 500 K, resulting in the unfolding of the protein and exposure of the TRP group to water. The level of solvent exposure can be gauged by a radial distribution function (RDF)  $g(r)$  of water around the center of the indole ring. Plotting the value of  $g(r)$  at  $r = 0.4$  nm, the approximate value for the first solvation shell, demonstrates that the proportion of water in close proximity to the TRP increases more rapidly after approximately 600 ps of simulation time (Figure 6.19).



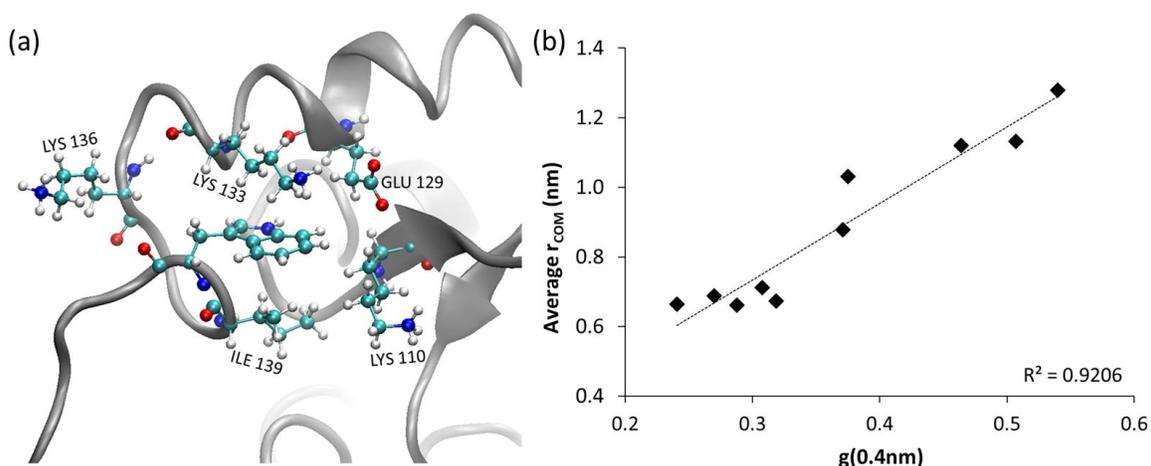
**Figure 6.19** The value of the TRP-water RDF taken at  $r = 0.4$  nm clearly shows an increasing access of water to the tryptophan as the trajectory progresses. The dashed and dotted lines show best fit lines to the data before and after 600 ps.



**Figure 6.20** Snapshots of the unfolding trajectory of snase at 0 ps (left), 600 ps (middle) and 1 ns (right). The amino acid side chains which are close to the tryptophan in the folded state are shown in the grey space-filling representation. At 600 ps, the surrounding residues have moved away from the TRP, allowing the access of water.

This was confirmed by a visual inspection of the trajectory (Figure 6.20). After 600 ps, the TRP residue is sufficiently mobile to rotate away from the surrounding side chains and become fully solvated.

Unfortunately, the TRP-water RDF value at a given distance cannot be constrained in MD simulations, so we cannot use it as our reaction coordinate. We must therefore find a geometrical parameter describing the snase protein that would directly correlate with the degree of water access to TRP. Inspection of the protein structure reveals that the TRP is surrounded by the side chains of five amino acids: lysine (LYS) 110, 133 and 136, glutamic acid (GLU) 129, and isoleucine (ILE) 139 (Figure 6.21a). As trial reaction coordinates, we tested a number of geometrical indexes describing the average distance between the aforementioned side chains and the center of mass of the tryptophan indole ring. A good correlation was achieved for the average distance between centers of mass, as shown in Figure 6.21(b).



**Figure 6.21** (a) The amino acid side chains that are in close proximity to tryptophan in the folded structure of snase. (b) Correlation of the average distance between the centers of mass of these amino acids and tryptophan with the value of the TRP-water RDF  $g$  at 0.4 nm.

However, while Royer's kinetic measurements probe the solvent exposure of TRP, they indirectly measure the restructuring of the entire protein since it is impossible to imagine that the TRP pocket is its only part that is unfolding in these experiments. It is equally unreasonable to expect that changing the average size of the TRP pocket will trigger global unfolding of the protein in MD simulations. Therefore, the ensembles that we would generate in our MD volume calculations using this reaction coordinate will not

match the ensembles that evolve under experimental conditions. Thus, the tryptophan environment may not be a useful reaction coordinate for the volume profile method.

## 6.4. Conclusions

Here, we presented some preliminary results towards understanding the reaction coordinate for the unfolding of model peptides and proteins. An analysis of the Best-Hummer optimization in its application to the unfolding of a model helix showed that this procedure can be vulnerable to a wrong choice of structural parameters, thus leading to an unwanted bias in the composition of the optimized reaction coordinate. Therefore, in application to the unfolding of a polyalanine alpha helix and the staphylococcal nuclease protein, we explored some alternative approaches of an *ad hoc* nature, which were more system-dependent and closely linked to the type of techniques used in the kinetic experiments. Encouraging results were obtained for the unfolding of helical systems, where selection of the essential structural parameters describing the process was based on the quality of kinetic data generated on reaction networks defined by a particular choice of these parameters. Although further analysis and computations are necessary to fully develop this approach, these preliminary results provide a useful framework upon which future studies can be built.

## Chapter 7.

### Concluding remarks and outlook

The goal of this work was to develop a computational methodology which can be used to provide a mechanistic interpretation of the effects of pressure on reaction systems in terms of their MD-calculated volume profiles. Two novel approaches were formulated, aimed at predicting the behaviour of chemical systems at high pressures and at elucidating TS structures for systems of varying complexity. The results presented in this work clearly demonstrate the efficacy and accuracy of volume profiles as computational mechanistic tools.

As an example of the utility of MD-calculated volume profiles for high pressure analysis, a comprehensive investigation of the effects of pressure on the methyl-methane hydrogen transfer system was presented in Chapter 2. The volume profile for the system was calculated and used in conjunction with the ambient pressure Gibbs energy profile to calculate high pressure Gibbs energy profiles. Consistent with the negative sign of the activation volume, the Gibbs activation energy for this reaction was found to decrease with pressure, which resulted in a complete elimination of the reaction barrier at a pressure of 125 kbar. High pressure enthalpy profiles were also calculated, and a reversal in the trend of the pressure-dependent activation enthalpies was found to occur after a threshold pressure of 500 bar. The pressure-induced viscosity effects for this system were found to be insignificant due to a relatively weak solvent-solute coupling, manifesting through the parallel alignment of reactant and product states on the 2D Gibbs energy surface in solvent-solute coordinates.

The demonstrated ability of pressure to push molecular systems towards the conformation with the smaller volume was harnessed in Chapter 3 for the control of molecular machines. This was exemplified by low-molecular-weight models of polythiophene molecular wires, displaying a shift in the Gibbs energy minimum from a twisted conformation at ambient pressure to a planar conformation at high pressure. The pressure needed to drive this transition was observed to be dependent on the level of substitution in the systems, indicating that a reduction of the required pressure could be achieved through a structural modification which changes the level of steric interactions.

We also proposed the first prototype of a pressure-driven unidirectional molecular rotor, which utilizes the fact that, whereas application of positive pressure drives a system towards its lowest volume conformation, the application of negative pressure will similarly drive the system towards its highest volume conformation. Analysis of a hypothetical prototype showed that a molecular system possessing a volume profile with extrema slightly shifted ahead of those of the Gibbs energy profile can be continuously propelled in a single direction via a sequence of positive and negative pressure pulses. A molecular rotor based on this model was proposed and tested, but further work is needed to suggest a potential synthetic prototype.

In Chapter 4, we developed and tested a novel methodology for the identification of TS ensembles using MD-calculated volume profiles in conjunction with experimentally-determined activation volumes. This method is particularly useful for conformationally flexible systems, where the TS structures are not easily obtainable using traditional computational methods due to the high dimensionality of the systems and long timescales of the required reaction simulations. We demonstrated the predictive power of this method through the accurate determination of the TS ensembles for two flexible systems in comparison with those determined by traditional, but time-consuming, computational techniques. In order to use this method to find the TS ensembles for larger, more complex systems, a properly defined reaction coordinate is needed. Some preliminary work towards understanding the mathematical form of a suitable reaction coordinate for systems with multiple reaction pathways was undertaken in Chapter 5. Using the Best-Hummer optimization procedure, we found that the kinetic accuracy of 1D Gibbs energy profiles was highly dependent on the functional form of the collective variables along which they were generated. We proposed a reaction coordinate with a new functional form, and showed that a 1D Gibbs energy profile along this coordinate can be used to accurately model the kinetics across the entire 2D surface. In Chapter 6, we presented some preliminary results towards the definition of a reaction coordinate for conformational changes in biological systems, ranging from a model helix to the staphylococcal nuclease protein.

In summary, the new methodology presented here constitutes an important step in developing theoretical and computational techniques aimed at describing high pressure systems and their transformations. Of particular interest are our results pertaining to the rational design of pressure-driven molecular machines and the determination of the

structure of unknown TS ensembles for flexible systems. The biggest challenge in extending the latter to larger systems of biological relevance is the lack of a conventionally accepted method for selecting a suitable reaction coordinate. Our analysis presented in Chapter 5 and the preliminary results of Chapter 6 will hopefully provide a good point of departure for further investigations.

## References

- (1) Hamann, S. D. *Physico-Chemical Effects of Pressure*; Academic Press: New York, 1957.
- (2) Whalley, E. Use of Volumes of Activation for Determining Reaction Mechanisms. *Adv. Phys. Org. Chem.* **1964**, 2, 93–162.
- (3) le Noble, W. J. Kinetics of Reactions in Solutions Under Pressure. In *Progress in Physical Organic Chemistry*; Streitwieser Jr., A., Taft, R. W., Eds.; John Wiley & Sons: Hoboken, N. J., 1967; Vol. 5, pp 207–330.
- (4) Klärner, F.-G.; Wurche, F. The Effect of Pressure on Organic Reactions. *J. Für Prakt. Chem.* **2000**, 342 (7), 609–636.
- (5) van't Hoff, J. H. *Lectures on Theoretical and Physical Chemistry*; Leffeldt, R. A., Translator; Edward Arnold: London, 1899.
- (6) Hamann, S. D. Chemical Kinetics. In *High Pressure Physics and Chemistry*; Bradley, R. S., Ed.; Academic Press: London and New York, 1963; Vol. 2, pp 163–208.
- (7) Planck, M. Ueber Das Princip Der Vermehrung Der Entropie. *Ann. Phys.* **1887**, 268 (11), 462–503.
- (8) Eyring, H. The Activated Complex in Chemical Reactions. *J. Chem. Phys.* **1935**, 3 (2), 107–115.
- (9) Glasstone, S.; Laidler, K. J.; Eyring, H. *The Theory of Rate Processes: The Kinetics of Chemical Reactions, Viscosity, Diffusion and Electrochemical Phenomena*; International chemical series; McGraw-Hill: New York, 1941.
- (10) Evans, M. G.; Polanyi, M. Some Applications of the Transition State Method to the Calculation of Reaction Velocities, Especially in Solution. *Trans. Faraday Soc.* **1935**, 31, 875–894.
- (11) Asano, T.; Le Noble, W. J. Activation and Reaction Volumes in Solution. *Chem. Rev.* **1978**, 78 (4), 407–489.
- (12) van Eldik, R.; Asano, T.; Le Noble, W. J. Activation and Reaction Volumes in Solution. 2. *Chem. Rev.* **1989**, 89 (3), 549–688.
- (13) Drljaca, A.; Hubbard, C. D.; van Eldik, R.; Asano, T.; Basilevsky, M. V.; le Noble, W. J. Activation and Reaction Volumes in Solution. 3. *Chem. Rev.* **1998**, 98 (6), 2167–2289.
- (14) Schettino, V.; Bini, R. Constraining Molecules at the Closest Approach: Chemistry at High Pressure. *Chem. Soc. Rev.* **2007**, 36 (6), 869–880.
- (15) Röntgen, W. C. Ueber Die Constitution Des Flüssigen Wassers. *Ann. Phys.* **1892**, 281 (1), 91–97.
- (16) Sander Jr., F. V. The Effects of High Pressure on the Inversion of Sucrose and the Mutarotation of Glucose. *J. Biol. Chem.* **1943**, 148 (2), 311–319.
- (17) Stern, O. Ueber Den Einfluss Des Druckes Auf Die Inversionsconstante Einiger Säuren. *Ann. Phys.* **1896**, 295 (11), 652–663.

- (18) Rothmund, V. Ueber Den Einfluss Des Druckes Auf Die Reaktionsgeschwindigkeit. *Z. Für Phys. Chem.* **1896**, *20U* (1), 168–179.
- (19) Koshkikallio, J.; Turpeinen, U. The Effect of Pressure and Competing A1 and A2 Mechanisms in the Acid-Catalysed Hydrolysis of Substituted Benzoic Anhydrides in Dioxan-Water Mixtures. *Acta Chem. Scand.* **1971**, *25* (9), 3360–3366.
- (20) Carroll, F. A. *Perspectives on Structure and Mechanism in Organic Chemistry*; John Wiley & Sons: Hoboken, N. J., 2011.
- (21) Jenner, G. Pericyclic Reactions. In *Organic High Pressure Chemistry*; le Noble, W. J., Ed.; Elsevier Science Publishers: New York, 1988; pp 143–203.
- (22) Houk, K. N.; Gonzalez, J.; Li, Y. Pericyclic Reaction Transition States: Passions and Punctilios, 1935-1995. *Acc. Chem. Res.* **1995**, *28* (2), 81–90.
- (23) Diels, O.; Alder, K. Synthesen in Der Hydroaromatischen Reihe. *Justus Liebigs Ann. Chem.* **1928**, *460* (1), 98–122.
- (24) Wassermann, A. The Mechanism of Additions to Double Bonds. Part I. Thermochemistry and Kinetics of a Diene Synthesis. *J. Chem. Soc.* **1935**, *0*, 828–839.
- (25) Woodward, R. B.; Katz, T. J. The Mechanism of the Diels-Alder Reaction. *Tetrahedron* **1959**, *5* (1), 70–89.
- (26) Stewart Jr., C. A. Competing Diradical and Electrocyclic Reactions. Difference in Activation Volumes. *J. Am. Chem. Soc.* **1972**, *94* (2), 635–637.
- (27) Klärner, F.-G.; Krawczyk, B.; Ruster, V.; Deiters, U. K. Evidence for Pericyclic and Stepwise Processes in the Cyclodimerization of Chloroprene and 1, 3-Butadiene from Pressure Dependence and Stereochemistry. Experimental and Theoretical Volumes of Activation and Reaction. *J. Am. Chem. Soc.* **1994**, *116* (17), 7646–7657.
- (28) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W.H. Freeman: New York, 1999.
- (29) Bridgman, P. W. The Coagulation of Albumen by Pressure. *J. Biol. Chem.* **1914**, *19* (4), 511–512.
- (30) Matthews, J. E.; Dow, R. B.; Anderson, A. K. The Effects of High Pressure on the Activity of Pepsin and Rennin. *J. Biol. Chem.* **1940**, *135* (2), 697–705.
- (31) Curl, A. L.; Jansen, E. F. Effect of High Pressures on Trypsin and Chymotrypsin. *J. Biol. Chem.* **1950**, *184* (1), 45–54.
- (32) Royer, C. A. Revisiting Volume Changes in Pressure-Induced Protein Unfolding. *Biochim. Biophys. Acta* **2002**, *1595* (1–2), 201–209.
- (33) Frye, K. J.; Royer, C. A. Probing the Contribution of Internal Cavities to the Volume Change of Protein Unfolding under Pressure. *Protein Sci.* **1998**, *7* (10), 2217–2222.
- (34) Rouget, J.-B.; Aksel, T.; Roche, J.; Saldana, J.-L.; Garcia, A. E.; Barrick, D.; Royer, C. A. Size and Sequence and the Volume Change of Protein Folding. *J. Am. Chem. Soc.* **2011**, *133* (15), 6020–6027.

- (35) Roche, J.; Caro, J. A.; Norberto, D. R.; Barthe, P.; Roumestand, C.; Schlessman, J. L.; Garcia, A. E.; Royer, C. A. Cavities Determine the Pressure Unfolding of Proteins. *Proc. Natl. Acad. Sci.* **2012**, *109* (18), 6945–6950.
- (36) Roche, J.; Dellarole, M.; Caro, J. A.; Guca, E.; Norberto, D. R.; Yang, Y.; Garcia, A. E.; Roumestand, C.; García-Moreno, B.; Royer, C. A. Remodeling of the Folding Free Energy Landscape of Staphylococcal Nuclease by Cavity-Creating Mutations. *Biochemistry* **2012**, *51* (47), 9535–9546.
- (37) Roche, J.; Dellarole, M.; Caro, J. A.; Norberto, D. R.; Garcia, A. E.; Garcia-Moreno, B.; Roumestand, C.; Royer, C. A. Effect of Internal Cavities on Folding Rates and Routes Revealed by Real-Time Pressure-Jump NMR Spectroscopy. *J. Am. Chem. Soc.* **2013**, *135* (39), 14610–14618.
- (38) Brower, K. R. A Method for Measuring the Activation Volumes of Fast Reversible Reactions. Ferric Thiocyanate Complex. *J. Am. Chem. Soc.* **1968**, *90* (20), 5401–5403.
- (39) Heremans, K.; Ceuterick, F.; Rijkenberg, J. High Pressure Pressure-Jump Apparatus. *Rev. Sci. Instrum.* **1980**, *51* (2), 252–253.
- (40) Vidugiris, G. J. A.; Markley, J. L.; Royer, C. A. Evidence for a Molten Globule-like Transition State in Protein Folding from Determination of Activation Volumes. *Biochemistry* **1995**, *34* (15), 4909–4912.
- (41) Mohana-Borges, R.; Silva, J. L.; Ruiz-Sanz, J.; de Prat-Gay, G. Folding of a Pressure-Denatured Model Protein. *Proc. Natl. Acad. Sci.* **1999**, *96* (14), 7888–7893.
- (42) Desai, G.; Panick, G.; Zein, M.; Winter, R.; Royer, C. A. Pressure-Jump Studies of the Folding/unfolding of Trp Repressor. *J. Mol. Biol.* **1999**, *288* (3), 461–475.
- (43) Pappenberger, G.; Saudan, C.; Becker, M.; Merbach, A. E.; Kiefhaber, T. Denaturant-Induced Movement of the Transition State of Protein Folding Revealed by High-Pressure Stopped-Flow Measurements. *Proc. Natl. Acad. Sci.* **2000**, *97* (1), 17–22.
- (44) Woenckhaus, J.; Köhling, R.; Thiyagarajan, P.; Littrell, K. C.; Seifert, S.; Royer, C. A.; Winter, R. Pressure-Jump Small-Angle X-Ray Scattering Detected Kinetics of Staphylococcal Nuclease Folding. *Biophys. J.* **2001**, *80* (3), 1518–1523.
- (45) Kitahara, R.; Royer, C.; Yamada, H.; Boyer, M.; Saldana, J.-L.; Akasaka, K.; Roumestand, C. Equilibrium and Pressure-Jump Relaxation Studies of the Conformational Transitions of P13MTCP1. *J. Mol. Biol.* **2002**, *320* (3), 609–628.
- (46) Tan, C.-Y.; Xu, C.-H.; Wong, J.; Shen, J.-R.; Sakuma, S.; Yamamoto, Y.; Lange, R.; Balny, C.; Ruan, K.-C. Pressure Equilibrium and Jump Study on Unfolding of 23-kDa Protein from Spinach Photosystem II. *Biophys. J.* **2005**, *88* (2), 1264–1275.
- (47) Kimura, T.; Sakamoto, K.; Morishima, I.; Ishimori, K. Dehydration in the Folding of Reduced Cytochrome C Revealed by the Electron-Transfer-Triggered Folding under High Pressure. *J. Am. Chem. Soc.* **2006**, *128* (3), 670–671.
- (48) Brun, L.; Isom, D. G.; Velu, P.; García-Moreno, B.; Royer, C. A. Hydration of the Folding Transition State Ensemble of a Protein. *Biochemistry* **2006**, *45* (11), 3473–3480.

- (49) Font, J.; Torrent, J.; Ribó, M.; Laurents, D. V.; Balny, C.; Vilanova, M.; Lange, R. Pressure-Jump-Induced Kinetics Reveals a Hydration Dependent Folding/Unfolding Mechanism of Ribonuclease A. *Biophys. J.* **2006**, *91* (6), 2264–2274.
- (50) Jenkins, D. C.; Pearson, D. S.; Harvey, A.; Sylvester, I. D.; Geeves, M. A.; Pinheiro, T. J. T. Rapid Folding of the Prion Protein Captured by Pressure-Jump. *Eur. Biophys. J.* **2009**, *38* (5), 625–635.
- (51) Osváth, S.; Quynh, L. M.; Smeller, L. Thermodynamics and Kinetics of the Pressure Unfolding of Phosphoglycerate Kinase. *Biochemistry* **2009**, *48* (42), 10146–10150.
- (52) Rouget, J.-B.; Schroer, M. A.; Jeworrek, C.; Pühse, M.; Saldana, J.-L.; Bessin, Y.; Tolan, M.; Barrick, D.; Winter, R.; Royer, C. A. Unique Features of the Folding Landscape of a Repeat Protein Revealed by Pressure Perturbation. *Biophys. J.* **2010**, *98* (11), 2712–2721.
- (53) Kramers, H. A. Brownian Motion in a Field of Force and the Diffusion Model of Chemical Reactions. *Physica* **1940**, *7* (4), 284–304.
- (54) Hanggi, P.; Talkner, P.; Borkovec, M. Reaction-Rate Theory: Fifty Years After Kramers. *Rev. Mod. Phys.* **1990**, *62* (2), 251–342.
- (55) Langer, J. S. Statistical Theory of the Decay of Metastable States. *Ann. Phys.* **1969**, *54* (2), 258–275.
- (56) Grote, R. F.; Hynes, J. T. The Stable States Picture of Chemical Reactions. II. Rate Constants for Condensed and Gas Phase Reaction Models. *J. Chem. Phys.* **1980**, *73* (6), 2715–2732.
- (57) Agmon, N.; Hopfield, J. J. Transient Kinetics of Chemical Reactions with Bounded Diffusion Perpendicular to the Reaction Coordinate: Intramolecular Processes with Slow Conformational Changes. *J. Chem. Phys.* **1983**, *78* (11), 6947–6959.
- (58) Agmon, N.; Hopfield, J. J. CO Binding to Heme Proteins: A Model for Barrier Height Distributions and Slow Conformational Changes. *J. Chem. Phys.* **1983**, *79* (4), 2042–2053.
- (59) Weidenmüller, H. A.; Jing-Shang, Z. Stationary Diffusion over a Multidimensional Potential Barrier: A Generalization of Kramers' Formula. *J. Stat. Phys.* **1984**, *34* (1), 191–201.
- (60) Berezhkovskii, A. M.; Zitserman, V. Y. Anomalous Regime for Decay of the Metastable State: An Extension of Multidimensional Kramers Theory. *Chem. Phys. Lett.* **1989**, *158* (5), 369–374.
- (61) Berezhkovskii, A. M.; D'yakov, Y. A.; Zitserman, V. Y. Smoluchowski Equation with a Sink Term: Analytical Solutions for the Rate Constant and Their Numerical Test. *J. Chem. Phys.* **1998**, *109* (11), 4182–4189.
- (62) Troe, J. Elementary Reactions in Compressed Gases and Liquids: From Collisional Energy Transfer to Diffusion Control. *J. Phys. Chem.* **1986**, *90* (3), 357–365.
- (63) Gehrke, C.; Schroeder, J.; Schwarzer, D.; Troe, J.; Voß, F. Photoisomerization of Diphenylbutadiene in Low-viscosity Nonpolar Solvents: Experimental

- Manifestations of Multidimensional Kramers Behavior and Cluster Effects. *J. Chem. Phys.* **1990**, *92* (8), 4805–4816.
- (64) Mohrschladt, R.; Schroeder, J.; Schwarzer, D.; Troe, J.; Vöhringer, P. Barrier Crossing and Solvation Dynamics in Polar Solvents: Photoisomerization of Trans-stilbene and E,E-diphenylbutadiene in Compressed Alkanols. *J. Chem. Phys.* **1994**, *101* (9), 7566–7579.
- (65) Asano, T.; Furuta, H.; Sumi, H. Two-Step Mechanism in Single-Step Isomerizations. Kinetics in a Highly Viscous Liquid Phase. *J. Am. Chem. Soc.* **1994**, *116* (13), 5545–5550.
- (66) Asano, T. Kinetics in Highly Viscous Solutions: Dynamic Solvent Effects in “Slow” Reactions. *Pure Appl. Chem.* **1999**, *71* (9), 1691–1704.
- (67) Asano, T. High-Pressure Kinetics and Highly Viscous Media. In *High Pressure Chemistry: Synthetic, Mechanistic, and Supercritical Applications*; van Eldik, R., Klärner, F.-G., Eds.; Wiley-VCH Verlag GmbH: Weinheim, Germany, 2002; pp 97–128.
- (68) Goto, Y.; Takahashi, T.; Ohga, Y.; Asano, T.; Hildebrand, M.; Weinberg, N. Dynamic Solvent Effects on the Thermal Cyclization of a Hexadienone Formed from a Diphenylnaphthopyran: An Example of a System with Distinctly Separate Medium and Chemical Contributions to the Overall Reaction Coordinate. *Phys. Chem. Chem. Phys.* **2003**, *5* (9), 1825–1830.
- (69) Kono, H.; Osako, H.; Sasaki, M.; Takahashi, T.; Ohga, Y.; Asano, T.; Hildebrand, M.; Weinberg, N. N. Dynamic Solvent Effects on the Fading of a Merocyanine Formed from a Spiroanthoxazine. A Case of a Strong Coupling between Chemical and Medium Coordinate. *Phys. Chem. Chem. Phys.* **2004**, *6* (9), 2260–2266.
- (70) Swaddle, T. W. Homogeneous versus Heterogeneous Self-Exchange Electron Transfer Reactions of Metal Complexes: Insights from Pressure Effects. *Chem. Rev.* **2005**, *105* (6), 2573–2608.
- (71) Gulam, R. M.; Takahashi, T.; Ohga, Y. Dynamic Solvent Effects on the Thermal Isomerization of Zinc Dithizonate. *Phys. Chem. Chem. Phys.* **2009**, *11* (25), 5170–5174.
- (72) Stearn, A. E.; Eyring, H. Pressure and Rate Processes. *Chem. Rev.* **1941**, *29* (3), 509–523.
- (73) Kitaigorodskii, A. I. *Organic Chemical Crystallography*; Consultants Bureau: New York, 1961.
- (74) Bondi, A. Van Der Waals Volumes and Radii. *J. Phys. Chem.* **1964**, *68* (3), 441–451.
- (75) Gavezzotti, A. The Calculation of Molecular Volumes and the Use of Volume Analysis in the Investigation of Structured Media and of Solid-State Organic Reactivity. *J. Am. Chem. Soc.* **1983**, *105* (16), 5220–5225.
- (76) Lustig, R. Surface and Volume of Three, Four, Six and Twelve Hard Fused Spheres. *Mol. Phys.* **1985**, *55* (2), 305–317.
- (77) Lee, B.; Richards, F. M. The Interpretation of Protein Structures: Estimation of Static Accessibility. *J. Mol. Biol.* **1971**, *55* (3), 379–400.

- (78) Connolly, M. L. Solvent-Accessible Surfaces of Proteins and Nucleic Acids. *Science* **1983**, *221* (4612), 709–713.
- (79) Richards, F. M. Areas, Volumes, Packing, and Protein Structure. *Annu. Rev. Biophys. Bioeng.* **1977**, *6* (1), 151–176.
- (80) Pavlov, M. Y.; Fedorov, B. A. Improved Technique for Calculating X-Ray Scattering Intensity of Biopolymers in Solution: Evaluation of the Form, Volume, and Surface of a Particle. *Biopolymers* **1983**, *22* (6), 1507–1522.
- (81) Richards, F. M. Calculation of Molecular Volumes and Areas for Structures of Known Geometry. *Methods Enzymol.* **1985**, *115*, 440–464.
- (82) Kundrot, C. E.; Ponder, J. W.; Richards, F. M. Algorithms for Calculating Excluded Volume and Its Derivatives as a Function of Molecular Conformation and Their Use in Energy Minimization. *J. Comput. Chem.* **1991**, *12* (3), 402–409.
- (83) Spooner, J.; Wiebe, H.; Boon, N.; Deglint, E.; Edwards, E.; Yanciw, B.; Patton, B.; Thiele, L.; Dance, P.; Weinberg, N. Molecular Dynamics Calculation of Molecular Volumes and Volumes of Activation. *Phys. Chem. Chem. Phys.* **2012**, *14* (7), 2264–2277.
- (84) Bader, R. F. W.; Henneker, W. H.; Cade, P. E. Molecular Charge Distributions and Chemical Binding. *J. Chem. Phys.* **1967**, *46* (9), 3341–3363.
- (85) Bader, R. F. W.; Carroll, M. T.; Cheeseman, J. R.; Chang, C. Properties of Atoms in Molecules: Atomic Volumes. *J. Am. Chem. Soc.* **1987**, *109* (26), 7968–7979.
- (86) Lim, C.; Chan, S. L.; Tole, P. Solvation Free Energies from a Combined Quantum Mechanical and Continuum Dielectric Approach. In *Structure and Reactivity in Aqueous Solution*; Cramer, C. J., Truhlar, D. G., Eds.; American Chemical Society, 1994; Vol. 568, pp 50–59.
- (87) Wiberg, K. B.; Rablen, P. R.; Rush, D. J.; Keith, T. A. Amides. 3. Experimental and Theoretical Studies of the Effect of the Medium on the Rotational Barriers for N,N-Dimethylformamide and N,N-Dimethylacetamide. *J. Am. Chem. Soc.* **1995**, *117* (15), 4261–4270.
- (88) Wiberg, K. B.; Keith, T. A.; Frisch, M. J.; Murcko, M. Solvent Effects on 1, 2-Dihaloethane Gauche/trans Ratios. *J. Phys. Chem.* **1995**, *99* (22), 9072–9079.
- (89) Wong, M. W.; Wiberg, K. B.; Frisch, M. J. Ab Initio Calculation of Molar Volumes: Comparison with Experiment and Use in Solvation Models. *J. Comput. Chem.* **1995**, *16* (3), 385–394.
- (90) Smith, B. J.; Hall, N. E. Atomic Radii: Incorporation of Solvation Effects. *J. Comput. Chem.* **1998**, *19* (13), 1482–1493.
- (91) Matubayasi, N.; Levy, R. M. Thermodynamics of the Hydration Shell. 2. Excess Volume and Compressibility of a Hydrophobic Solute. *J. Phys. Chem.* **1996**, *100* (7), 2681–2688.
- (92) Lin, C.-L.; Wood, R. H. Prediction of the Free Energy of Dilute Aqueous Methane, Ethane, and Propane at Temperatures from 600 to 1200 C and Densities from 0 to 1 G/Cm<sup>-3</sup> Using Molecular Dynamics Simulations. *J. Phys. Chem.* **1996**, *100* (40), 16399–16409.

- (93) Lockwood, D. M.; Rossky, P. J. Evaluation of Functional Group Contributions to Excess Volumetric Properties of Solvated Molecules. *J. Phys. Chem. B* **1999**, *103* (11), 1982–1990.
- (94) Chai, C. C.; Jhon, M. S. Molecular Dynamics Study on Protein and Its Water Structure at High Pressure. *Mol. Simul.* **2000**, *23* (4–5), 257–274.
- (95) Imai, T.; Kinoshita, M.; Hirata, F. Theoretical Study for Partial Molar Volume of Amino Acids in Aqueous Solution: Implication of Ideal Fluctuation Volume. *J. Chem. Phys.* **2000**, *112* (21), 9469–9478.
- (96) Dadarlat, V. M.; Post, C. B. Insights into Protein Compressibility from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2001**, *105* (3), 715–724.
- (97) DeVane, R.; Ridley, C.; Larsen, R. W.; Space, B.; Moore, P. B.; Chan, S. I. A Molecular Dynamics Method for Calculating Molecular Volume Changes Appropriate for Biomolecular Simulation. *Biophys. J.* **2003**, *85* (5), 2801–2807.
- (98) Floris, F. M. Nonideal Effects on the Excess Volume from Small to Large Cavities in TIP4P Water. *J. Phys. Chem. B* **2004**, *108* (41), 16244–16249.
- (99) Ridley, C.; Stern, A. C.; Green, T.; DeVane, R.; Space, B.; Miksovskva, J.; Larsen, R. W. A Combined Photothermal and Molecular Dynamics Method for Determining Molecular Volume Changes. *Chem. Phys. Lett.* **2006**, *418* (1–3), 137–141.
- (100) Moghaddam, M. S.; Chan, H. S. Pressure and Temperature Dependence of Hydrophobic Hydration: Volumetric, Compressibility, and Thermodynamic Signatures. *J. Chem. Phys.* **2007**, *126* (11), 114507.
- (101) Sangwai, A. V.; Ashbaugh, H. S. Aqueous Partial Molar Volumes from Simulation and Individual Group Contributions. *Ind. Eng. Chem. Res.* **2008**, *47* (15), 5169–5174.
- (102) Sarupria, S.; Ghosh, T.; García, A. E.; Garde, S. Studying Pressure Denaturation of a Protein by Molecular Dynamics Simulations. *Proteins Struct. Funct. Bioinforma.* **2010**, *78* (7), 1641–1651.
- (103) Yu, I.; Tasaki, T.; Nakada, K.; Nagaoka, M. Influence of Hydrostatic Pressure on Dynamics and Spatial Distribution of Protein Partial Molar Volume: Time-Resolved Surficial Kirkwood-Buff Approach. *J. Phys. Chem. B* **2010**, *114* (38), 12392–12397.
- (104) Patel, N.; Dubins, D. N.; Pomès, R.; Chalikian, T. V. Parsing Partial Molar Volumes of Small Molecules: A Molecular Dynamics Study. *J. Phys. Chem. B* **2011**, *115* (16), 4856–4862.
- (105) Medvedev, N. N.; Voloshin, V. P.; Kim, A. V.; Anikeenko, A. V.; Geiger, A. Calculation of Partial Molar Volume and Its Components for Molecular Dynamics Models of Dilute Solutions. *J. Struct. Chem.* **2013**, *54* (S2), 271–288.
- (106) Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. Determination of Partial Molar Volumes from Free Energy Perturbation Theory. *Phys. Chem. Chem. Phys.* **2015**, *17* (13), 8407–8415.
- (107) Voloshin, V. P.; Medvedev, N. N.; Smolin, N.; Geiger, A.; Winter, R. Exploring Volume, Compressibility and Hydration Changes of Folded Proteins upon Compression. *Phys. Chem. Chem. Phys.* **2015**, *17* (13), 8499–8508.

- (108) Del Galdo, S.; Amadei, A. The Unfolding Effects on the Protein Hydration Shell and Partial Molar Volume: A Computational Study. *Phys. Chem. Chem. Phys.* **2016**, *18* (40), 28175–28182.
- (109) Wiktor, J.; Bruneval, F.; Pasquarello, A. Partial Molar Volumes of Aqua Ions from First Principles. *J. Chem. Theory Comput.* **2017**, *13* (8), 3427–3431.
- (110) Deglint, E.; Martens, H.; Edwards, E.; Boon, N.; Dance, P.; Weinberg, N. Molecular Dynamics Calculation of Activation Volumes. *Phys. Chem. Chem. Phys.* **2011**, *13* (2), 438–440.
- (111) Wiebe, H.; Spooner, J.; Boon, N.; Deglint, E.; Edwards, E.; Dance, P.; Weinberg, N. Calculation of Molecular Volumes and Volumes of Activation Using Molecular Dynamics Simulations. *J. Phys. Chem. C* **2012**, *116* (3), 2240–2245.
- (112) le Noble, W. J. The Volume Profile of Chemical Reactions (Modern Aspects of Physical Chemistry at High Pressure: The 50th Commemorative Volume). *Rev. Phys. Chem. Jpn.* **1980**, *50*, 207–216.
- (113) Jenner, G. Volume Profiles of Pericyclic Reactions Involving Transition States. A Pressure Study. *New J. Chem.* **1991**, *15* (12), 897–899.
- (114) Stochel, G.; van Eldik, R. Elucidation of Inorganic Reaction Mechanisms through Volume Profile Analysis. *Coord. Chem. Rev.* **1999**, *187* (1), 329–374.
- (115) Hubbard, C. D.; van Eldik, R. Mechanistic Studies of Reactions of Coordination Compounds. Some Recent Highlights. *J. Coord. Chem.* **2007**, *60* (1), 1–51.
- (116) Hubbard, C. D.; van Eldik, R. Mechanistic Information on Some Inorganic and Bioinorganic Reactions from Volume Profile Analysis. *Inorganica Chim. Acta* **2010**, *363* (11), 2357–2374.
- (117) Jenner, G. High-Pressure Mechanistic Delineation Based on Activation Volumes. *J. Phys. Org. Chem.* **2002**, *15* (1), 1–13.
- (118) Spooner, J.; Yanciw, B.; Wiebe, B.; Weinberg, N. Reaction Profiles and Energy Surfaces of Compressed Species. *J. Phys. Chem. A* **2014**, *118* (4), 765–777.
- (119) Leach, A. R. *Molecular Modelling: Principles and Applications*, 2nd ed.; Pearson Education; Prentice Hall: Upper Saddle River, NJ, 2001.
- (120) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447.
- (121) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; et al. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29* (7), 845–854.
- (122) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (123) Jones, J. E. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proc. R. Soc. Lond. Ser. A* **1924**, *106* (738), 463–477.
- (124) Ryckaert, J.-P.; Bellemans, A. Molecular Dynamics of Liquid Alkanes. *Faraday Discuss. Chem. Soc.* **1978**, *66*, 95–106.

- (125) Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110* (6), 1657–1666.
- (126) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* **1996**, *118* (45), 11225–11236.
- (127) Wang, L.-P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theory Comput.* **2013**, *9* (1), 452–460.
- (128) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684.
- (129) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. *Gaussian 09, Revision E.01*; Gaussian, Inc.: Wallingford CT, 2016.
- (130) Burklé-Vitzthum, V.; Michels, R.; Scacchi, G.; Marquaire, P.-M. Mechanistic Modeling of the Thermal Cracking of Decylbenzene. Application to the Prediction of Its Thermal Stability at Geological Temperatures. *Ind. Eng. Chem. Res.* **2003**, *42* (23), 5791–5808.
- (131) Lannuzel, F.; Bounaceur, R.; Michels, R.; Scacchi, G.; Marquaire, P.-M. Reassessment of the Kinetic Influence of Toluene on N-Alkane Pyrolysis. *Energy Fuels* **2010**, *24* (7), 3817–3830.
- (132) Spooner, J.; Smith, B.; Weinberg, N. Effect of High Pressure on the Topography of Potential Energy Surfaces. *Can. J. Chem.* **2016**, *94* (12), 1057–1064.
- (133) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.
- (134) Flyvbjerg, H.; Petersen, H. G. Error Estimates on Averages of Correlated Data. *J. Chem. Phys.* **1989**, *91* (1), 461–466.
- (135) Hess, B. Determining the Shear Viscosity of Model Liquids from Molecular Dynamics Simulations. *J. Chem. Phys.* **2002**, *116* (1), 209.
- (136) Wurche, F.; Sicking, W.; Sustmann, R.; Klärner, F.-G.; Rüdhardt, C. The Effect of Pressure on Hydrogen Transfer Reactions with Quinones. *Chem. - Eur. J.* **2004**, *10* (11), 2707–2721.
- (137) Chen, B.; Hoffmann, R.; Cammi, R. The Effect of Pressure on Organic Reactions in Fluids—a New Theoretical Perspective. *Angew. Chem. Int. Ed.* **2017**, *56* (37), 11126–11142.
- (138) Walling, C.; Schugar, H. J. Organic Reactions under High Pressure. VII. Volumes of Activation for Some Diels-Alder Reactions. *J. Am. Chem. Soc.* **1963**, *85* (5), 607–612.
- (139) El'yanov, B. S.; Gonikberg, E. M. Effect of Pressure and Temperature on Some Kinetic and Thermodynamic Parameters of Non-Ionic Reactions. Volume Changes on Activation and Reaction. *J. Chem. Soc. Faraday Trans. 1* **1979**, *75*, 172–191.

- (140) Span, R. *Multiparameter Equations of State - An Accurate Source of Thermodynamic Property Data*; Springer-Verlag: Berlin, Heidelberg, 2000.
- (141) Bridgman, P. W. *The Physics of High Pressure.*; G. Bell: London, 1958.
- (142) Saluja, P. P. S.; Whalley, E. Chemical Kinetics at Very High Pressures. A Pressure-Induced Cage Effect in an Ion-Producing Reaction. *J. Chem. Soc. Chem. Commun.* **1983**, 0 (10), 552–553.
- (143) Whalley, E.; le Noble, W. J. Carbocationic Reactions. In *Organic High Pressure Chemistry*; le Noble, W. J., Ed.; Elsevier Science Publishers: New York, 1988; pp 238–253.
- (144) Wiebe, H.; Prachnau, M.; Weinberg, N. Hydrogen Transfer Reactions in Viscous Media — Potential and Free Energy Surfaces in Solvent–solute Coordinates and Their Kinetic Implications. *Can. J. Chem.* **2013**, 91 (9), 787–794.
- (145) Dhaliwal, M.; Basilevsky, M. V.; Weinberg, N. Dynamic Effects of Nonequilibrium Solvation: Potential and Free Energy Surfaces for Z/E Isomerization in Solvent-Solute Coordinates. *J. Chem. Phys.* **2007**, 126 (23), 234505.
- (146) Huber, L.; Edwards, E.; Basilevsky, M. V.; Weinberg, N. Reactions in Viscous Media: Potential and Free Energy Surfaces in Solvent-Solute Coordinates. *Mol. Phys.* **2009**, 107 (21), 2283–2291.
- (147) Burrows, H. D.; Hartshorn, R. M. The 2016 Nobel Prize in Chemistry. *Pure Appl. Chem.* **2016**, 88 (10–11), 917–918.
- (148) *Molecular Machines and Motors*; Credi, A., Silvi, S., Venturi, M., Eds.; Topics in Current Chemistry; Springer International Publishing: Switzerland, 2014; Vol. 354.
- (149) Abendroth, J. M.; Bushuyev, O. S.; Weiss, P. S.; Barrett, C. J. Controlling Motion at the Nanoscale: Rise of the Molecular Machines. *ACS Nano* **2015**, 9 (8), 7746–7768.
- (150) Kay, E. R.; Leigh, D. A. Rise of the Molecular Machines. *Angew. Chem. Int. Ed.* **2015**, 54 (35), 10080–10088.
- (151) Kassem, S.; van Leeuwen, T.; Lubbe, A. S.; Wilson, M. R.; Feringa, B. L.; Leigh, D. A. Artificial Molecular Motors. *Chem Soc Rev* **2017**, 46 (9), 2592–2621.
- (152) Bustamante, C.; Chemla, Y. R.; Forde, N. R.; Izhaky, D. Mechanical Processes in Biochemistry. *Annu. Rev. Biochem.* **2004**, 73 (1), 705–748.
- (153) Feringa, B. L.; van Delden, R. A.; Koumura, N.; Geertsema, E. M. Chiroptical Molecular Switches. *Chem. Rev.* **2000**, 100 (5), 1789–1816.
- (154) Credi, A.; Silvi, S.; Venturi, M. Light-Operated Machines Based on Threaded Molecular Structures. In *Molecular Machines and Motors: Recent Advances and Perspectives*; Credi, A., Silvi, S., Venturi, M., Eds.; Topics in Current Chemistry; Springer International Publishing: Switzerland, 2014; pp 1–34.
- (155) Guldi, D. M.; Nishihara, H.; Venkataraman, L. Molecular Wires. *Chem Soc Rev* **2015**, 44 (4), 842–844.
- (156) Lo, W.-Y.; Zhang, N.; Cai, Z.; Li, L.; Yu, L. Beyond Molecular Wires: Design Molecular Electronic Functions Based on Dipolar Effect. *Acc. Chem. Res.* **2016**, 49 (9), 1852–1863.

- (157) Kottas, G. S.; Clarke, L. I.; Horinek, D.; Michl, J. Artificial Molecular Rotors. *Chem. Rev.* **2005**, *105* (4), 1281–1376.
- (158) Hirokawa, N. Kinesin and Dynein Superfamily Proteins and the Mechanism of Organelle Transport. *Science* **1998**, *279* (5350), 519–526.
- (159) Schliwa, M.; Woehlke, G. Molecular Motors. *Nature* **2003**, *422* (6933), 759–765.
- (160) Spooner, J.; Wiebe, H.; Louwse, M.; Reader, B.; Weinberg, N. Theoretical Analysis of High-Pressure Effects on Conformational Equilibria. *Can. J. Chem.* **2017**, *Ahead of print*, DOI: 10.1139/cjc-2017-0411.
- (161) Diaz-Quijada, G. A.; Weinberg, N.; Holdcroft, S.; Pinto, B. M. Investigation of Barriers To Conformational Interchange in Oligothiophenes and Oligo(Thienyl)furans. *J. Phys. Chem. A* **2002**, *106* (7), 1266–1276.
- (162) Diaz-Quijada, G. A.; Weinberg, N.; Holdcroft, S.; Pinto, B. M. Conformational Analysis of Oligothiophenes and Oligo(thienyl)furans by Use of a Combined Molecular Dynamics/NMR Spectroscopic Protocol. *J. Phys. Chem. A* **2002**, *106* (7), 1277–1285.
- (163) Meng, S.; Ma, J.; Jiang, Y. Solvent Effects on Electronic Structures and Chain Conformations of  $\alpha$ -Oligothiophenes in Polar and Apolar Solutions. *J. Phys. Chem. B* **2007**, *111* (16), 4128–4136.
- (164) Sato, M.; Tanaka, S.; Kaeriyama, K. Soluble Conducting Polythiophenes. *J. Chem. Soc. Chem. Commun.* **1986**, *0* (11), 873–874.
- (165) Elsenbaumer, R. L.; Jen, K. Y.; Oboodi, R. Processible and Environmentally Stable Conducting Polymers. *Synth. Met.* **1986**, *15* (2–3), 169–174.
- (166) Hotta, S.; Rughooputh, S. D. D. V.; Heeger, A. J.; Wudl, F. Spectroscopic Studies of Soluble Poly (3-Alkylthienylenes). *Macromolecules* **1987**, *20* (1), 212–215.
- (167) Patil, A. O.; Ikenoue, Y.; Wudl, F.; Heeger, A. J. Water-Soluble Conducting Polymers. *J. Am. Chem. Soc.* **1987**, *109* (6), 1858–1859.
- (168) Kelly, T. R.; De Silva, H.; Silva, R. A. Unidirectional Rotary Motion in a Molecular System. *Nature* **1999**, *401* (6749), 150–152.
- (169) Koumura, N.; Zijlstra, R. W. J.; van Delden, R. A.; Harada, N.; Feringa, B. L. Light-Driven Monodirectional Molecular Rotor. *Nature* **1999**, *401* (6749), 152–155.
- (170) Noji, H. F1-Motor of ATP Synthase. In *Molecular Motors*; Schliwa, M., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2002; pp 141–151.
- (171) Berry, R. The Bacterial Flagellar Motor. In *Molecular Motors*; Schliwa, M., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2002; pp 111–140.
- (172) Kelly, T. R.; Cai, X.; Damkaci, F.; Panicker, S. B.; Tu, B.; Bushell, S. M.; Cornella, I.; Piggott, M. J.; Salives, R.; Cavero, M.; et al. Progress toward a Rationally Designed, Chemically Powered Rotary Molecular Motor. *J. Am. Chem. Soc.* **2007**, *129* (2), 376–386.
- (173) Fletcher, S. P.; Dumur, F.; Pollard, M. M.; Feringa, B. L. A Reversible, Unidirectional Molecular Rotary Motor Driven by Chemical Energy. *Science* **2005**, *310* (5745), 80–82.

- (174) Lin, Y.; Dahl, B. J.; Branchaud, B. P. Net Directed 180° Aryl–aryl Bond Rotation in a Prototypical Achiral Biaryl Lactone Synthetic Molecular Motor. *Tetrahedron Lett.* **2005**, *46* (48), 8359–8362.
- (175) Wilson, M. R.; Solà, J.; Carlone, A.; Goldup, S. M.; Lebrasseur, N.; Leigh, D. A. An Autonomous Chemically Fuelled Small-Molecule Motor. *Nature* **2016**, *534* (7606), 235–240.
- (176) van Delden, R. A.; ter Wiel, M. K. J.; Pollard, M. M.; Vicario, J.; Koumura, N.; Feringa, B. L. Unidirectional Molecular Motor on a Gold Surface. *Nature* **2005**, *437* (7063), 1337–1340.
- (177) Štacko, P.; Kistemaker, J. C. M.; van Leeuwen, T.; Chang, M.-C.; Otten, E.; Feringa, B. L. Locked Synchronous Rotor Motion in a Molecular Motor. *Science* **2017**, *356* (6341), 964–968.
- (178) Truhlar, D. G.; Garrett, B. C. Variational Transition-State Theory. *Acc. Chem. Res.* **1980**, *13* (12), 440–448.
- (179) Truhlar, D. G.; Garrett, B. C. Variational Transition State Theory. *Annu. Rev. Phys. Chem.* **1984**, *35* (1), 159–189.
- (180) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. *Proteins Struct. Funct. Bioinforma.* **1995**, *21* (3), 167–195.
- (181) Onuchic, J. N.; Socci, N. D.; Luthey-Schulten, Z.; Wolynes, P. G. Protein Folding Funnels: The Nature of the Transition State Ensemble. *Fold. Des.* **1996**, *1* (6), 441–450.
- (182) Royer, C. A. The Nature of the Transition State Ensemble and the Mechanisms of Protein Folding: A Review. *Arch. Biochem. Biophys.* **2008**, *469* (1), 34–45.
- (183) Hehre, W. J.; Radom, L.; von Schleyer, P. R.; Pople, J. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
- (184) Shaik, S. S.; Schlegel, H. B.; Wolfe, S. *Theoretical Aspects of Physical Organic Chemistry: The SN2 Mechanism*; Wiley: New York, 1992.
- (185) Fueno, T. *Transition State: A Theoretical Approach*; Kodansha, Ltd.: Tokyo, 1999.
- (186) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; John Wiley & Sons: Chichester, England, 2007.
- (187) Lewars, E. G. *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, 2nd ed.; Springer: Dordrecht, 2011.
- (188) Wolfe, S.; Buckley, A. V.; Weinberg, N. A Comprehensive Computational Examination of Transannular Diels Alder Reactions of Unsubstituted C14 Trienes Barriers, Template Effects, and the Curtin Hammett Principle. *Can. J. Chem.* **2001**, *79* (8), 1284–1292.
- (189) Hsieh, Y.-H.; Weinberg, N.; Yang, K.; Kim, C.-K.; Shi, Z.; Wolfe, S. Hydration of the Carbonyl Group Acetic Acid Catalysis in the Co-Operative Mechanism. *Can. J. Chem.* **2005**, *83* (6–7), 769–785.

- (190) Meng, Q.; Zhang, C.; Huang, M.-B. A Theoretical Model Study on the Cyclic Reaction of 4-Hydroxybutanal Catalyzed by Bronsted Acid. *Can. J. Chem.* **2009**, *87* (11), 1610–1619.
- (191) Shi, Z.; Hsieh, Y.-H.; Weinberg, N.; Wolfe, S. The Neutral Hydrolysis of Methyl Acetate - Part 2. Is There a Tetrahedral Intermediate? *Can. J. Chem.* **2009**, *87* (4), 544–555.
- (192) Gómez-Bombarelli, R.; Calle, E.; Casado, J. Mechanisms of Lactone Hydrolysis in Neutral and Alkaline Conditions. *J. Org. Chem.* **2013**, *78* (14), 6868–6879.
- (193) Tílviz, E.; Cárdenas-Jirón, G. I.; Menéndez, M. I.; López, R. Understanding the Hydrolysis Mechanism of Ethyl Acetate Catalyzed by an Aqueous Molybdocene: A Computational Chemistry Investigation. *Inorg. Chem.* **2015**, *54* (4), 1223–1231.
- (194) Yuan, Y.; Mills, M. J. L.; Popelier, P. L. A.; Jensen, F. Comprehensive Analysis of Energy Minima of the 20 Natural Amino Acids. *J. Phys. Chem. A* **2014**, *118* (36), 7876–7891.
- (195) Chun, C. P.; Connor, A. A.; Chass, G. A. Ab Initio Conformational Analysis of N- and C-Terminally-Protected Valyl-Alanine Dipeptide Model. *J. Mol. Struct. THEOCHEM* **2005**, *729* (3), 177–184.
- (196) Wierzejewska, M.; Olbert-Majkut, A. Conformational Behavior of the Simplest Dipeptide: Formylglycine. Quantum Chemical and Matrix Isolation FTIR Study. *Chem. Phys. Lett.* **2009**, *476* (4–6), 287–292.
- (197) Atwood, R. E.; Urban, J. J. Conformations of the Glycine Tripeptide Analog Ac-Gly-Gly-NHMe: A Computational Study Including Aqueous Solvation Effects. *J. Phys. Chem. A* **2012**, *116* (5), 1396–1408.
- (198) Wang, C.; Lin, Z.; Zhang, R. Zwitterions Are the Most Stable Form for Neutral Arginylglycine in Gas Phase: Clear Theoretical Evidence. *Comput. Theor. Chem.* **2013**, *1008*, 96–102.
- (199) Xu, S.; Zhao, D.-X.; Gong, L.-D.; Liu, C.; Yang, Z.-Z. Search of the Conformations of Val-Dipeptide and Val-Tripeptide by Ab Initio Method and ABEEMσπ Polarizable Force Field. *Chem. Phys. Lett.* **2015**, *618*, 147–152.
- (200) Kaminský, J.; Jensen, F. Conformational Interconversions of Amino Acid Derivatives. *J. Chem. Theory Comput.* **2016**, *12* (2), 694–705.
- (201) Wiebe, H.; Weinberg, N. Theoretical Volume Profiles as a Tool for Probing Transition States: Folding Kinetics. *J. Chem. Phys.* **2014**, *140* (12), 124105.
- (202) Wiebe, H.; Louwarse, M.; Weinberg, N. Theoretical Volume Profiles for Conformational Changes: Application to Internal Rotation of Benzene Ring in 1,12-Dimethoxy-[12]-Paracyclophane. *J. Chem. Phys.* **2017**, *146* (10), 104107.
- (203) Toan, N. M.; Morrison, G.; Hyeon, C.; Thirumalai, D. Kinetics of Loop Formation in Polymer Chains. *J. Phys. Chem. B* **2008**, *112* (19), 6094–6106.
- (204) Fugmann, S.; Sokolov, I. M. Internal Friction and Mode Relaxation in a Simple Chain Model. *J. Chem. Phys.* **2009**, *131* (23), 235104.
- (205) Taylor, M. P.; Paul, W.; Binder, K. Two-State Protein-like Folding of a Homopolymer Chain. *Phys. Procedia* **2010**, *4*, 151–160.

- (206) Cheng, R. R.; Makarov, D. E. Failure of One-Dimensional Smoluchowski Diffusion Models to Describe the Duration of Conformational Rearrangements in Floppy, Diffusive Molecular Systems: A Case Study of Polymer Cyclization. *J. Chem. Phys.* **2011**, *134* (8), 85104.
- (207) Movahed, H. B.; van Zon, R.; Schofield, J. Free Energy Landscape of Protein-like Chains with Discontinuous Potentials. *J. Chem. Phys.* **2012**, *136* (24), 245103.
- (208) Brown, D.; Clarke, J. H. R. A Direct Method of Studying Reaction Rates by Equilibrium Molecular Dynamics: Application to the Kinetics of Isomerization in Liquid N-butane. *J. Chem. Phys.* **1990**, *92* (5), 3062–3073.
- (209) Perkins, J.; Edwards, E.; Kleiv, R.; Weinberg, N. Molecular Dynamics Study of Reaction Kinetics in Viscous Media. *Mol. Phys.* **2011**, *109* (15), 1901–1909.
- (210) Gonikberg, M. G. *Chemical Equilibria and Reaction Rates at High Pressures*; National Science Foundation, Washington, 1963.
- (211) Isaacs, N. S. *Liquid Phase High Pressure Chemistry*; John Wiley & Sons, 1981.
- (212) Kunugi, S.; Kitayaki, M.; Yanagi, Y.; Tanaka, N.; Lange, R.; Balny, C. The Effect of High Pressure on Thermolysin. *Eur. J. Biochem.* **1997**, *248* (2), 567–574.
- (213) Herberhold, H.; Marchal, S.; Lange, R.; Scheyhing, C. H.; Vogel, R. F.; Winter, R. Characterization of the Pressure-Induced Intermediate and Unfolded State of Red-Shifted Green Fluorescent Protein—A Static and Kinetic FTIR, UV/VIS and Fluorescence Spectroscopy Study. *J. Mol. Biol.* **2003**, *330* (5), 1153–1164.
- (214) Zheng, J.; Zhao, Y.; Truhlar, D. G. Thermochemical Kinetics of Hydrogen-Atom Transfers between Methyl, Methane, Ethynyl, Ethyne, and Hydrogen. *J. Phys. Chem. A* **2007**, *111* (21), 4632–4642.
- (215) Xu, X.; Alecu, I. M.; Truhlar, D. G. How Well Can Modern Density Functionals Predict Internuclear Distances at Transition States? *J. Chem. Theory Comput.* **2011**, *7* (6), 1667–1676.
- (216) Yamada, H.; Mukuno, K.; Umeda, M.; Maeda, T.; Sera, A. NMR Study of Pressure Effect on the Hindered Rotation in Cis-1,12-Disubstituted[12]paracyclophanes. *Chem. Lett.* **1996**, 437–438.
- (217) Yamada, H.; Yonehara, S.; Tanaka, S.; Muro, F.; Watanabe, A.; Nishikawa, K. High-Pressure NMR Study of Cis-1,n-Disubstituted[n]paracyclophanes. Effect of Increased Pressure on the Hindered Internal Rotation. *J. Am. Chem. Soc.* **2001**, *123* (2), 279–284.
- (218) Viana, C. A. N.; Reis, J. C. R. Pressure Dependence of Rate and Equilibrium Constants in Solution. A Guide to Analytical Equations. *Pure Appl. Chem.* **1996**, *68* (8), 1541–1554.
- (219) Spooner, J.; Weinberg, N. A Comparative Analysis of Empirical Equations Describing Pressure Dependence of Equilibrium and Reaction Rate Constants. *Can. J. Chem.* **2017**, *95* (2), 149–161.
- (220) El'yanov, B. S.; Vasylytskaya, E. M. Some Recent Approaches to the Quantitative Description of the Effect of High Pressure on the Thermodynamic and Kinetic Parameters of Chemical Reactions (Modern Aspects of Physical

- Chemistry at High Pressure: The 50th Commemorative Volume). *Rev. Phys. Chem. Jpn.* **1980**, *50*, 169–184.
- (221) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011–1021.
- (222) Shoemaker, B. A.; Wang, J.; Wolynes, P. G. Structural Correlations in Protein Folding Funnels. *Proc. Natl. Acad. Sci.* **1997**, *94* (3), 777–782.
- (223) Shoemaker, B. A.; Wang, J.; Wolynes, P. G. Exploring Structures in Protein Folding Funnels with Free Energy Functionals: The Transition State Ensemble. *J. Mol. Biol.* **1999**, *287* (3), 675–694.
- (224) Ma, B.; Kumar, S.; Tsai, C.-J.; Hu, Z.; Nussinov, R. Transition-State Ensemble in Enzyme Catalysis: Possibility, Reality, or Necessity? *J. Theor. Biol.* **2000**, *203* (4), 383–397.
- (225) Best, R. B.; Hummer, G. Reaction Coordinates and Rates from Transition Paths. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6732–6737.
- (226) Krivov, S. V. On Reaction Coordinate Optimality. *J. Chem. Theory Comput.* **2013**, *9* (1), 135–146.
- (227) Peters, B. Common Features of Extraordinary Rate Theories. *J. Phys. Chem. B* **2015**, *119* (21), 6349–6356.
- (228) McGibbon, R. T.; Husic, B. E.; Pande, V. S. Identification of Simple Reaction Coordinates from Complex Dynamics. *J. Chem. Phys.* **2017**, *146* (4), 44109.
- (229) Li, A.; Daggett, V. Characterization of the Transition State of Protein Unfolding by Use of Molecular Dynamics: Chymotrypsin Inhibitor 2. *Proc. Natl. Acad. Sci.* **1994**, *91* (22), 10430–10434.
- (230) Hummer, G.; García, A. E.; Garde, S. Conformational Diffusion and Helix Formation Kinetics. *Phys. Rev. Lett.* **2000**, *85* (12), 2637–2640.
- (231) García, A. E.; Onuchic, J. N. Folding a Protein in a Computer: An Atomic Description of the Folding/unfolding of Protein A. *Proc. Natl. Acad. Sci.* **2003**, *100* (24), 13898–13903.
- (232) Clementi, C.; García, A. E.; Onuchic, J. N. Interplay Among Tertiary Contacts, Secondary Structure Formation and Side-Chain Packing in the Protein Folding Mechanism: All-Atom Representation Study of Protein L. *J. Mol. Biol.* **2003**, *326* (3), 933–954.
- (233) Juraszek, J.; Bolhuis, P. G. Sampling the Multiple Folding Mechanisms of Trp-Cage in Explicit Solvent. *Proc. Natl. Acad. Sci.* **2006**, *103* (43), 15859–15864.
- (234) Wu, S.; Zhuravlev, P. I.; Papoian, G. A. High Resolution Approach to the Native State Ensemble Kinetics and Thermodynamics. *Biophys. J.* **2008**, *95* (12), 5524–5532.
- (235) Shao, Q.; Wei, H.; Gao, Y. Q. Effects of Turn Stability and Side-Chain Hydrophobicity on the Folding of  $\beta$ -Structures. *J. Mol. Biol.* **2010**, *402* (3), 595–609.
- (236) Pan, A. C.; Weinreich, T. M.; Shan, Y.; Scarpazza, D. P.; Shaw, D. E. Assessing the Accuracy of Two Enhanced Sampling Methods Using EGFR Kinase

- Transition Pathways: The Influence of Collective Variable Choice. *J. Chem. Theory Comput.* **2014**, *10* (7), 2860–2865.
- (237) Bursulaya, B. D.; Brooks, C. L., III. Folding Free Energy Surface of a Three-Stranded  $\beta$ -Sheet Protein. *J. Am. Chem. Soc.* **1999**, *121* (43), 9947–9951.
- (238) Prigozhin, M. B.; Gruebele, M. The Fast and the Slow: Folding and Trapping of  $\lambda 6$ –85. *J. Am. Chem. Soc.* **2011**, *133* (48), 19338–19341.
- (239) Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. Diffusive Dynamics of the Reaction Coordinate for Protein Folding Funnels. *J. Chem. Phys.* **1996**, *104* (15), 5860–5868.
- (240) Sheinerman, F. B.; Brooks, C. L., III. Calculations on Folding of Segment B1 of Streptococcal Protein G. *J. Mol. Biol.* **1998**, *278* (2), 439–456.
- (241) Best, R. B.; Hummer, G. Diffusive Model of Protein Folding Dynamics with Kramers Turnover in Rate. *Phys. Rev. Lett.* **2006**, *96* (22), 228104.
- (242) Best, R. B.; Hummer, G.; Eaton, W. A. Native Contacts Determine Protein Folding Mechanisms in Atomistic Simulations. *Proc. Natl. Acad. Sci.* **2013**, *110* (44), 17874–17879.
- (243) Yang, S.; Onuchic, J. N.; García, A. E.; Levine, H. Folding Time Predictions from All-Atom Replica Exchange Simulations. *J. Mol. Biol.* **2007**, *372* (3), 756–763.
- (244) Whitford, P. C.; Noel, J. K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K. Y.; Onuchic, J. N. An All-Atom Structure-Based Potential for Proteins: Bridging Minimal Models with All-Atom Empirical Forcefields. *Proteins Struct. Funct. Bioinforma.* **2009**, *75* (2), 430–441.
- (245) Hirst, J. D.; Brooks, C. L., III. Helicity, Circular Dichroism and Molecular Dynamics of Proteins. *J. Mol. Biol.* **1994**, *243* (2), 173–178.
- (246) Bulheller, B. M.; Rodger, A.; Hirst, J. D. Circular and Linear Dichroism of Proteins. *Phys. Chem. Chem. Phys.* **2007**, *9* (17), 2020–2035.
- (247) Asciotto, E. K.; Mikhonin, A. V.; Asher, S. A.; Madura, J. D. Computational and Experimental Determination of the  $\alpha$ -Helix Unfolding Reaction Coordinate. *Biochemistry* **2008**, *47* (7), 2046–2050.
- (248) Fedorov, M. V.; Schumm, S. The Effect of Sodium Chloride on Poly-L-Glutamate Conformation. *Chem Commun* **2009**, *0* (8), 896–898.
- (249) Piaggi, P. M.; Valsson, O.; Parrinello, M. Enhancing Entropy and Enthalpy Fluctuations to Drive Crystallization in Atomistic Simulations. *Phys. Rev. Lett.* **2017**, *119* (1), 15701.
- (250) Palazzesi, F.; Valsson, O.; Parrinello, M. Conformational Entropy as Collective Variable for Proteins. *J. Phys. Chem. Lett.* **2017**, *8* (19), 4752–4756.
- (251) Peters, B. Recent Advances in Transition Path Sampling: Accurate Reaction Coordinates, Likelihood Maximisation and Diffusive Barrier-Crossing Dynamics. *Mol. Simul.* **2010**, *36* (15), 1265–1281.
- (252) Chodera, J. D.; Pande, V. S. Splitting Probabilities as a Test of Reaction Coordinate Choice in Single-Molecule Experiments. *Phys. Rev. Lett.* **2011**, *107* (9), 98102.

- (253) Geissler, P. L.; Dellago, C.; Chandler, D. Kinetic Pathways of Ion Pair Dissociation in Water. *J. Phys. Chem. B* **1999**, *103* (18), 3706–3710.
- (254) Bolhuis, P. G.; Dellago, C.; Chandler, D. Reaction Coordinates of Biomolecular Isomerization. *Proc. Natl. Acad. Sci.* **2000**, *97* (11), 5877–5882.
- (255) Berezhkovskii, A.; Szabo, A. One-Dimensional Reaction Coordinates for Diffusive Activated Rate Processes in Many Dimensions. *J. Chem. Phys.* **2005**, *122* (1), 14503.
- (256) Ma, A.; Dinner, A. R. Automatic Method for Identifying Reaction Coordinates in Complex Systems. *J. Phys. Chem. B* **2005**, *109* (14), 6769–6779.
- (257) Rhee, Y. M.; Pande, V. S. One-Dimensional Reaction Coordinate and the Corresponding Potential of Mean Force from Commitment Probability Distribution. *J. Phys. Chem. B* **2005**, *109* (14), 6780–6786.
- (258) Peters, B.; Bolhuis, P. G.; Mullen, R. G.; Shea, J.-E. Reaction Coordinates, One-Dimensional Smoluchowski Equations, and a Test for Dynamical Self-Consistency. *J. Chem. Phys.* **2013**, *138* (5), 54106.
- (259) Li, W.; Ma, A. Recent Developments in Methods for Identifying Reaction Coordinates. *Mol. Simul.* **2014**, *40* (10–11), 784–793.
- (260) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How Fast-Folding Proteins Fold. *Science* **2011**, *334* (6055), 517–520.
- (261) Piana, S.; Sarkar, K.; Lindorff-Larsen, K.; Guo, M.; Gruebele, M.; Shaw, D. E. Computational Design and Experimental Testing of the Fastest-Folding  $\beta$ -Sheet Protein. *J. Mol. Biol.* **2011**, *405* (1), 43–48.
- (262) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. Protein Folding Kinetics and Thermodynamics from Atomistic Simulation. *Proc. Natl. Acad. Sci.* **2012**, *109* (44), 17845–17850.
- (263) Hoops, S.; Sahle, S.; Gauges, R.; Lee, C.; Pahle, J.; Simus, N.; Singhal, M.; Xu, L.; Mendes, P.; Kummer, U. COPASI—a COmplex PATHway Simulator. *Bioinformatics* **2006**, *22* (24), 3067–3074.
- (264) Carpenter, B. K. Nonstatistical Dynamics in Thermal Reactions of Polyatomic Molecules. *Annu. Rev. Phys. Chem.* **2005**, *56* (1), 57–89.
- (265) Carpenter, B. K. Energy Disposition in Reactive Intermediates. *Chem. Rev.* **2013**, *113* (9), 7265–7286.
- (266) Ma, X.; Hase, W. L. Perspective: Chemical Dynamics Simulations of Non-Statistical Reaction Dynamics. *Philos. Trans. R. Soc. A* **2017**, *375* (2092), 20160204.
- (267) Pitts, C. G. C. *Introduction to Metric Spaces*; Oliver and Boyd: Edinburgh, 1972.
- (268) Shreider, Y. A. *What Is Distance?*; Cohn, L., Edelberg, H., Translators; University of Chicago Press: Chicago, 1974.
- (269) Fierz, B.; Reiner, A.; Kiefhaber, T. Local Conformational Dynamics in  $\alpha$ -Helices Measured by Fast Triplet Transfer. *Proc. Natl. Acad. Sci.* **2009**, *106* (4), 1057–1062.

- (270) Neumaier, S.; Reiner, A.; Buttner, M.; Fierz, B.; Kiefhaber, T. Testing the Diffusing Boundary Model for the Helix-Coil Transition in Peptides. *Proc. Natl. Acad. Sci.* **2013**, *110* (32), 12905–12910.
- (271) Neumaier, S.; Buttner, M.; Bachmann, A.; Kiefhaber, T. Transition State and Ground State Properties of the Helix-Coil Transition in Peptides Deduced from High-Pressure Studies. *Proc. Natl. Acad. Sci.* **2013**, *110* (52), 20988–20993.
- (272) Fierz, B.; Joder, K.; Krieger, F.; Kiefhaber, T. Using Triplet-Triplet Energy Transfer to Measure Conformational Dynamics in Polypeptide Chains. In *Protein Folding Protocols*; Bai, Y., Nussinov, R., Eds.; Humana Press: Totowa, NJ, 2006; pp 169–187.
- (273) Royer, C. A. Fluorescence Spectroscopy. In *Protein Stability and Folding: Theory and Practice*; Shirley, B. A., Ed.; Humana Press: Totowa, NJ, 1995; pp 65–89.
- (274) Royer, C. A. Probing Protein Folding and Conformational Transitions with Fluorescence. *Chem. Rev.* **2006**, *106* (5), 1769–1784.
- (275) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput. Phys. Commun.* **2014**, *185* (2), 604–613.