The Polytope Formalism:

isomerism and associated

unimolecular isomerisation

A thesis submitted in partial fulfilment of the

requirements for admission to the degree of

Doctor of Philosophy

by

Peter J. Canfield

THE UNIVERSITY OF

SYDNEY

School of Chemistry

December, 2021

I. Authorship attribution

The work described in this thesis was carried out in the School of Chemistry at The University of Sydney between 3 November 2017 and 20 December 2021, under the supervision of Professor Maxwell J. Crossley and Professor Jeffrey R. Reimers (University of Technology Sydney and Shanghai University). This includes study periods spent at Shanghai University. All experiments, calculations and informatics reported have been carried out by me, except when specific reference to the contributions of others as described in the text.

The material in this Thesis was developed through extensive collaborative discussions with my supervisors Professor Maxwell J. Crossley and Professor Jeffrey R. Reimers. The drug docking calculations were carried out by Professor David Hibbs, Dr Jonathan Du, and Paul Groundwater.

Sections of this work have been submitted for publication and/or have been presented at scientific conferences:

Contents of Chapters 3 and 4:

- Canfield, P. J.; Blake, I. M.; Cai, Z.-L.; Luck, I. J.; Krausz, E.; Kobayashi, R.; Reimers, J. R.; Crossley, M. J., A New Fundamental Type of Conformational Isomerism. *Nature Chemistry* **2018**, *10*, 615-624.

- Canfield, P. J.; Reimers, J. R.; Crossley, M. J., Non-Canonical Atropisomers Versus Akamptisomers: *E-Letter* Re: "Total Synthesis Reveals Atypical Atropisomerism in a Small-Molecule Natural Product, Tryptorubin a" by Reisberg, Gao, Walker, Helfrich, Clardy, and Baran. *Science*, **2020**, *367*, 458-463.

Contents of Chapter 5:

- Canfield, P. J.; Govenlock, L. J.; Reimers, J.; Crossley, M. J., Recent Advances in Stereochemistry Reveal Classification Shortcomings. *ChemRxiv* **2020**, *12488525*.

Conference and other presentations:

- **Canfield, P. J.**; Akamptisomerism. **2017**, *Shanghai University ICQMS presentation.*

- **Canfield, P. J.**; Fundamentals of Akamptisomerism. **2018**, 1st International Akamptisomerisation Symposium, Shanghai University, China.

- **Canfield, P. J.**; Future directions. **2018**, 1st International Akamptisomerisation Symposium, Shanghai University, China.

- **Canfield, P. J.**; Reimers, J. R.; Crossley, M. J., The Full Potential Energy Surface for N-H Tautomerization in Free-base Porphyrin. **2019**, *RACI PhysChem2019, Perth.*

- **Canfield, P. J.**; Reimers, J. R.; Crossley, M. J., The Full Potential Energy Surface for N-H Tautomerization in Free-base Porphyrin. **2019**, *14th International Conference on Tetrapyrrole Photoreceptors in Photosynthetic Organisms, Sydney.*

- **Canfield, P. J.**; Reimers, J. R.; Crossley, M. J., The Full Potential Energy Surface for N-H Tautomerization in Free-base Porphyrin Reveals Deep Insights into General Structural Possibilities for Macrocyclic Complexes. **2019**, *APATCC 2019.*

- **Canfield, P. J.**; Tay, A.; Brothers, P.; Reimers, J. R.; Crossley, M. J., Macrocycle induced compression of coordinated B–O–B groups controls the barriers to akamptisomeric inversion. **2019** *USyd School of Chemistry Student Seminar Series*

- **Canfield, P. J.**; Blake, I. M.; Cai, Z-L.; Luck, I. J.; Krausz, E.; Kobayashi, R.; Reimers, J. R.; Crossley, M. J., Akamptisomerism, the final fundamental type of stereoisomerism: how it shows the way forward to the complete conceptualisation of Chemical Space. **2020,** *USyd School of Chemistry Final PhD talk.*

- **Canfield, P. J**.; Blake, I. M.; Cai, Z-L.; Luck, I. J.; Krausz, E.; Kobayashi, R.; Reimers, J. R.; Crossley, M. J., Akamptisomerism, the final fundamental type of stereoisomerism: how it shows the way forward to the complete conceptualisation of Chemical Space. **2020**, *RACI Physical Chemistry Division talks*.

Many presentations on the topics described in this Thesis have also been given by my supervisors Professor Maxwell J. Crossley and Professor Jeffrey R. Reimers.

Other work done during my PhD not described in this thesis:

- Peiris, C. R.; Ciampi, S.; Dief, E. M.; Zhang, J.; Canfield, P. J.; Le Brun, A. P.; Kosov, D. S.; Reimers, J. R.; Darwish, N., Spontaneous S–Si Bonding of Alkanethiols to Si(111)–H: Towards Si–Molecule–Si Circuits. *Chemical Science* **2020**, *11*, 5246-5256.

- Fares, M., Canfield, P., Alsherbiny, M. A., Lewis, W., Willis, A. C., Guang Li, C., Neyts, J., Jochmans, J., Gale, P. A., Keller, P. A., Synthesis, X-ray crystallographic analysis, DFT studies and biological evaluation of triazolopyrimidines and 2-anilinopyrimidines. *Journal of Molecular Structure* **2021**, doi:10.1016/j.molstruc.2021.132092

II. Acknowledgments

This thesis is dedicated to the memory of Earl Muetterties (1927 – 1984) who introduced the polytopal-rearrangements model of stereochemistry, stereochemistry pioneer Kurt Mislow (1923 – 2017), and porphyrin physical chemist Martin Gouterman (1931 – 2020).

I would like to thank the technical staff at The University of Sydney for all their support in facilitating these research efforts in the School of Chemistry, especially given the significant challenges that the COVID-19 pandemic presented.

Also, I would like to thank my collaborators at other universities. Elmars Krausz of the Australian National University, Research School of Chemistry for his guidance with the MCD measurements and his inexhaustable enthusiam, and Rika Kobayashi of the Australian National University and the National Computational Infrastructure for her tireless support and profound knowledge of complex quantum chemical calculations and how to implement them.

A special thanks to Graham Price for his help with the Latin and Greek terms and reigniting my love of philology.

Next, I would like to express my deep gratitude to my supervisors, Professor Maxwell Crossley and Professor Jeffrey Reimers. I have known and worked closely with each of them over the past 18 years, and they have made an indelible contribution to who I am, both personally and professionally. It has been a true privilege to work with such dedicated and knowledgeable mentors and I am proud to call each of them, *Doktorvater*.

III. Abstract

This thesis, at its broadest, concerns the so-called *ontology* of isomerism and isomerisation, this encompassing the conceptual frameworks and relationships that comprise the subject matter; the necessary formal definitions, nomenclature, and representations that have impacts reaching into unexpected areas such as drug registration and patent specifications; the requisite controlled and precise vocabulary that facilitates nuanced communication; and the digital/computational formalisms that underpin the chemistry software and database tools that empower chemists to perform much of their work.

Using conceptual tools taken from Combinatorics, Group Theory, and Graph Theory, means are presented for the provision of a unified description of isomerism and associated unimolecular isomerisation spanning both constitutional isomerism and stereoisomerism called the Polytope Formalism. This includes unification of the varying approaches historically taken to describe and understand stereoisomerism in organic and inorganic compounds.

The work in this Thesis began, historically, with the synthesis, isolation, and characterisation of compounds that could not be adequately described using existing IUPAC recommendations. Generalisation of the polytopal-rearrangements model of stereoisomerisation used for inorganic chemistry led to the prescriptions that could deal with the synthesised compounds, revealing an unrecognised fundamental form of isomerism called *akamptisomerism*.

Following on, this Thesis describes how in attempting to place akamptisomerism within the context of existing stereoisomerism reveals significant systematic deficiencies in the IUPAC recommendations. Further, these shortcomings have limited the conceptualisation of broad classes of compounds and hindered development of molecules for both medicinal and technological applications.

It is shown how the Polytope Formalism can be applied to the description of constitutional isomerism in a practical manner. Finally, a radically different medicinal chemistry design strategy with broad application, based upon the principles, is described.

IV. Table of Contents

V. List of Tables

VI. List of Figures

VII. List of abbreviations

1D	One dimensional
2D	Two dimensional
3D	Three dimensional
ABS, Abs	Absorption (spectrum)
ak	Akamptisomerisation
AM1	Austin Method 1 (semi-empirical method)
API	Active pharmaceutical ingredient
B3LYP	Becke, 3-parameter, Lee–Yang–Parr functional
BAR	Bond angle reflection
CAM-B3LYP	Coulomb-attenuating method B3LYP functional
CAS	Chemical Abstracts Service
cc-pVDZ	Correlation-consistent polarised valence double zeta basis set
cc-pVTZ	Correlation-consistent polarised valence triple zeta basis set
CCSD(T)	Coupled-cluster method with singles and doubles and with perturbative correction for triples
CD	Circular dichroism (spectroscopy)
CIS	Configuration interaction, singles
cm	Centimetre
cm^{-1}	Wavenumbers in inverse cm
D3	Grimme's three-term empirical dispersion correction
D3(BJ)	Grimme's three-term empirical dispersion correction with Becke-Johnson damping
DFT	Density-functional theory
E-M	Angle between the electric transition dipole moment and the magnetic transition dipole moment
EMA	European Medicines Agency
Et	Ethyl
HPLC	High performance liquid chromatography
IKKβ	Inhibitor of nuclear factor kappa-B kinase subunit beta
InChI	IUPAC's International Chemical Identifier
IUPAC	International Union of Pure and Applied Chemistry
K	Kelvin
L	Litre
ln	Natural logarithm
M	Molar (mol L^{-1})
mL	Millilitre
mol	Moles
MCD	Magnetic circular dichroism
MeTHF	2-methyl-tetrahydrofuran

MP2	Møller–Plesset second-order perturbation theory
NMR	Nuclear magnetic resonance (spectroscopy)
Osc. Str.	Oscillator strength
p	Porphyrinate(2-)
Ph	Phenyl
PM3	Parametrised method 3 (semi-empirical method)
PM6	Parametrised method 6 (semi-empirical method)
PM6	Parametrised method 7 (semi-empirical method)
pqx	Quinoxalino[2,3−b']porphyrinate(2-)
(*R*)	*Rectus* (absolute configuration of tetrahedral centre)
R	Molar gas constant
Rot. Str.	Rotatory strength
(*S*)	*Sinistral* (absolute configuration of tetrahedral centre)
SMILES	Simplified molecular-input line-entry system
S_N2	Bimolecular nucleophilic substitution
THF	Tetrahydrofuran
TM	Transition moment
TMS	Tetramethylsilane
tpp	5,10,15,20-Tetraphenylporphyrin
TS	Transition structure, Transition state
ttp	5,10,15,20-Tetrakis(4-toluyl)porphyrin
USFDA	United States Food and Drug Administration
UV/Vis	Ultraviolet/visible (absorption spectrum)
λ_{max}	Wavelength of local maximum absorption

VIII. List of Electronic Files

In addition to the University of Sydney links, electronic files 2 – 5 and 8 are available for download as supplementary files to Canfield, P. J.; Blake, I. M.; Cai, Z.-L.; Luck, I. J.; Krausz, E.; Kobayashi, R.; Reimers, J. R.; Crossley, M. J., A New Fundamental Type of Conformational Isomerism. *Nature Chemistry* **2018**, *10*, 615-624.

E_File_1

https://cloudstor.aarnet.edu.au/plus/s/JDHhDYBYS9x0QDF

"E_File_1 - ML3_embedded_species.zip"
Cartesian coordinates (as ASCII ".dat" files) for the ML_3 polytope family and the Fortran code and compiled executable (for any 32-bit Windows platform) written by Professor Jeffrey R. Reimers, for generating them.

E_File_2

https://cloudstor.aarnet.edu.au/plus/s/D01oo1X7Zglvh7c

https://static-content.springer.com/esm/art%3A10.1038%2Fs41557-018-0043-6/MediaObjects/41557_2018_43_MOESM70_ESM.gif

This animates the four synthesised *transoid* quinoxalinoporphyrin stereoisomers such that their subtle differences are made apparent. Rings co-planar with in-plane (*parvo*) boron atoms are highlighted yellow. See Section 4.3.3.

(R_a,R_p)–B_2OF_2pqx (S_a,S_p)–B_2OF_2pqx
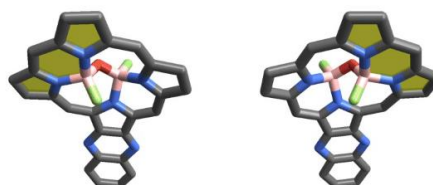
(R_p,R_a)–B_2OF_2pqx (S_p,S_a)–B_2OF_2pqx

E_File_3

https://cloudstor.aarnet.edu.au/plus/s/SyuswlevB3oQmst

https://static-content.springer.com/esm/art%3A10.1038%2Fs41557-018-0043-6/MediaObjects/41557_2018_43_MOESM71_ESM.gif

This animates the four conceptual *cisoid* quinoxalinoporphyrin stereoisomers such that their subtle differences are made apparent. Rings co-planar with in-plane (*parvo*) boron atoms are highlighted yellow. See Section 4.3.3.

(R_a,S_a)–B_2OF_2pqx (S_a,R_a)–B_2OF_2pqx
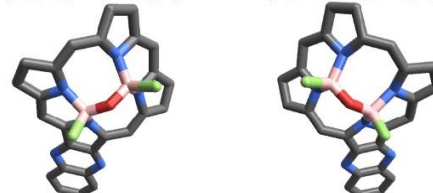
(R_p,S_p)–B_2OF_2pqx (S_p,R_p)–B_2OF_2pqx
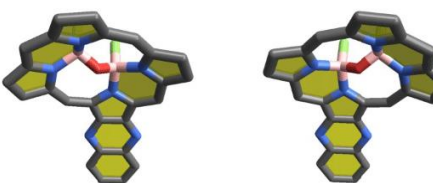
E_File_4

This animates the $R_{st}^c 1$ mechanism that must always be conceptually possible for the interconversion of *transoid* B_2OF_2-quinoxalinoporphyrin akamptisomers. For these isomers, the observed pattern of specific diastereomeric interconversion proves that this is the actual mechanism of thermal akamptisomerisation. See Section 4.5.
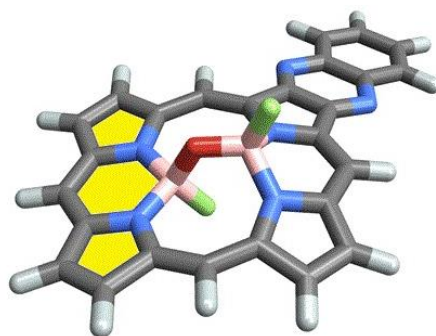
$(R_a,R_p)–B_2OF_2pqx \rightleftharpoons (R_p,R_a)–B_2OF_2pqx$

E_File_5

This animates normal vibrational mode #12 associated with the associated akamptisomerisation transition state, demonstrating that it closely resembles the $R_{st}^c 1$ akamptisomerisation mechanism shown in E_File_4. See Section 3.5 and Section 4.5.
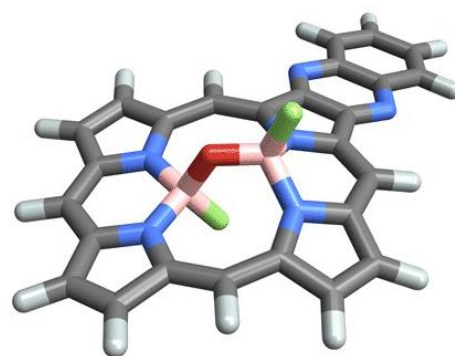
$(R_a,R_p)–B_2OF_2pqx$ normal vibrational mode 12

E_File_6

Animation of a forced hula-twist-like akamptisomerisation mechanism of *meso*-unsubstituted **2a** $\rightleftharpoons$ **3a** where the B–O–B angle is fixed at 130° (see Sections 2.4.11 and 4.5.2).

E_File_7
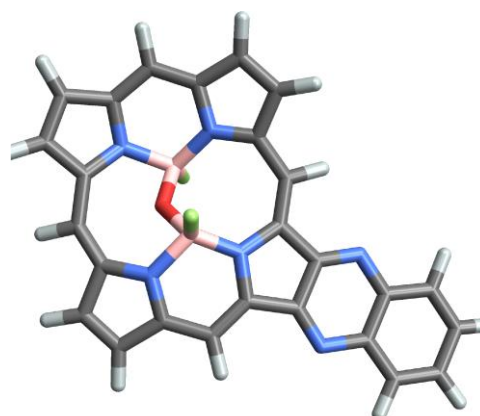
https://cloudstor.aarnet.edu.au/plus/s/ULL0Li0H08tRHdh

Animation of the asynchronous B(F)OB(F)-group strepsisomerisation "bond-walk" mechanism (see Section 2.4.26). See Section 4.5.3 for details.

E_File_8
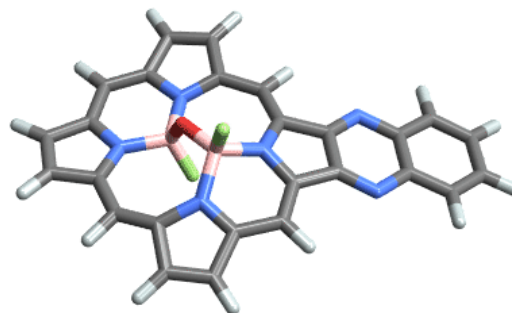
https://cloudstor.aarnet.edu.au/plus/s/l3StGq54j2bcuu8

https://static-content.springer.com/esm/art%3A10.1038%2Fs41557-018-0043-6/MediaObjects/41557_2018_43_MOESM74_ESM.zip

"E_File_8 - Chapter 4 DFT optimised coordinates.zip"
Zip archive folder of all computer-optimised molecular structures reported in Chapter 4.
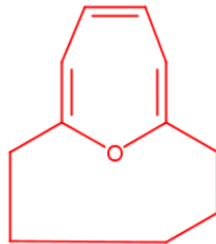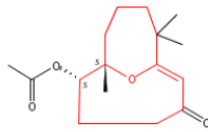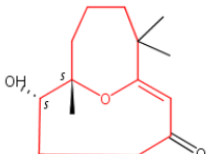
E_File_9

https://cloudstor.aarnet.edu.au/plus/s/tPgmj0R78pEO5ft

"E_File_9 - CAS SciFinder search results March 2019.zip"
A zip archive folder containing pdf files listing the 1409 CAS search results records of ring-locked oxabicyclo akamptisomers, as of March 2019, see Section 5.3.1.

SciFinder®   Page 2

1.
832111-05-4

$C_{11}H_{14}O$
12-Oxabicyclo[5.4.1]dodeca-7,9,11-triene
Key Physical Properties:
Molecular Weight
162.23
Boiling Point (Predicted)
Value: 318.6±11.0 °C | Condition: Press: 760 Torr
Density (Predicted)
Value: 1.02±0.1 g/cm3 | Condition: Temp: 20 °C Press: 760 Torr

Related Info:
~ 2 References
Reactions

2.
252231-40-6

Relative stereochemistry., Currently available stereo shown.
$C_{16}H_{24}O_4$
12-Oxabicyclo[5.4.1]dodec-1-en-3-one, 6-(acetyloxy)-7,11,11-trimethyl-, (6R,7R)-rel-
Key Physical Properties:
Molecular Weight
280.36
Boiling Point (Predicted)
Value: 375.8±31.0 °C | Condition: Press: 760 Torr
Density (Predicted)
Value: 1.09±0.1 g/cm3 | Condition: Temp: 20 °C Press: 760 Torr

Related Info:
~ 1 References

3.
252231-39-3

Rotation (-)., Absolute stereochemistry unknown., Currently available stereo shown.
$C_{14}H_{22}O_3$
12-Oxabicyclo[5.4.1]dodec-1-en-3-one, 6-hydroxy-7,11,11-trimethyl-, (6S,7S)-rel-(-)-
Key Physical Properties:
Molecular Weight
238.32
Boiling Point (Predicted)
Value: 365.1±31.0 °C | Condition: Press: 760 Torr
Density (Predicted)
Value: 1.09±0.1 g/cm3 | Condition: Temp: 20 °C Press: 760 Torr
pKa (Predicted)
Value: 14.04±0.60 | Condition: Most Acidic Temp: 25 °C

Related Info:
~ 1 References

E_File_10

https://cloudstor.aarnet.edu.au/plus/s/ny5DdNwVDmrRafb

"E_File_10 - Free-Base Porphyrin 'Distal' Structure Potential Energy Surface Explorer.cdf"
Mathematica computable document format as described in Chapter 6 and Appendix F. Can be executed in Mathematica[1] or using the free Wolfram Player available here (for Windows, Mac and Linux).

E_File_11

https://cloudstor.aarnet.edu.au/plus/s/tqkVSzJOoh21KjJ

"E_File_11 - Ch6 species.zip"
Zip archive folder containing the Fortran source code and PC executable for generating families, genera, and species for the D_{mh} site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioning approach to the Polytope Formalism of constitutional isomerism assuming non-dissociating and non-differentiable binders with m, $n \leq 8$ as described in Chapter 6 and Appendix G.

E_File_12

https://cloudstor.aarnet.edu.au/plus/s/o6FPKJtyewEqgrW

"E_File_12 - SmBn partitioned Polytope Formalism of constitutional isomerism graphing.nb"
The Mathematica script for interactively graphing D_{mh} site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioning approach to the Polytope Formalism of constitutional isomerism. This script uses the .dat output files from the Fortran software in **E_File_11** and outputs a wide variety of Graph Theory related files and properties. See details in described in Chapter 6 and Appendix H

E_File_13

https://cloudstor.aarnet.edu.au/plus/s/WD48Qt1sFBEiUo1

"E_File_13 - S4B2_genera{1-28}_all.zip"
This is a zip archive folder of the D_{4h} site symmetric $\mathcal{S}_4\mathcal{B}_2$ partitioned, full 1st order and 2nd-order motions graph output from the Mathematica[1] graphing software in **E_File_12**.

E_File_14

https://cloudstor.aarnet.edu.au/plus/s/sBYOPz560JbTXjA

"E_File_14 - Ch6 optimised coordinates.pdf"
DFT optimised coordinates for molecules from Chapter 6. Calculation details for all species are given along with systematic chemical names for most.

1   Introduction

1.1   Classifying isomerism and unimolecular isomerisation

Understanding isomerism – the different ways in which atoms can be rearranged into chemical structures – is one of the most important aspects to the understanding of chemistry. As such, it forms some of the earliest lessons for chemistry students. Isomerism is the subdiscipline of chemistry that rationalises and classifies/catalogues what chemical *structures* are possible for a given set of atoms, whilst isomerisation is concerned with how isomers may *interconvert*. Isomerisation of molecules can occur by many broadly described processes. Of these processes, *unimolecular* isomerisation is a special category in that a single molecular entity rearranges without the involvement of other molecular entities. For a given set of atoms, not all possible isomers are necessarily accessible to each other *via* a single-step rearrangement process. Understanding what processes are allowed, how many steps are required to connect any pair of isomers, and what the kinetics are for each step, thus become important to the understanding of both the dynamical behaviour of molecular entities and for how desired chemical reactions can be selected for or promoted, and undesired reactions suppressed.

Most broadly, isomers are classified into two main types: *constitutional isomers*, in which the connections between the constituent atoms (bonding topologies) are different, and *stereoisomers* in which the bonding topologies are preserved but the spatial arrangements of the atoms vary. These are depicted graphically in Figure 1.1a. Whilst this classification is considered simple and intuitive, in its essence it relies on definitions concerning what constitutes a "bond" and so may be subject to ambiguities. If such ambiguities do arise, the traditional resolution to the problem has been simply use whichever options are most appropriate for the chemical system being considered.

Stereoisomers are further divided into subcategories. Some of these classifications rely on global geometric properties of molecules. For example, enantiomers are defined as non-superimposable mirror-image pairs of stereoisomers with those that do not fit this criterion called diastereomers as shown in in Figure 1.1a. Other classifications depart from this geometry-only approach and have arisen phenomenologically from the observation

(a)

Global properties

```
Isomerism ──→ Constitutional isomerism
              (Same formula, different connectivity)
```

```
              Stereoisomerism
(Same formula, same connectivity, different spatial arrangements)
```

```
Enantiomerism                    Diasteromerism
(non-superimposable              (all stereoisomers other
mirror images)                   than enantiomers)
```

(b)

Atom-centric stereoisomerism properties

Stereoisomers can interconvert by the following $R_{st}^c 1$ processes:

Conformational isomerism	Configurational isomerism
(*low energy* polytopal rearrangements including torsions)	(*high energy* polytopal rearrangements)

Fluxionality ≪ ≫ Practical isolability ≫ Configurationally stable

Torsional isomerism including rotamerism, atropisomerism, and torsion-mediated (*E,Z*)-diastereoisomerisation

Akamptisomerism including *e.g.*, (*E,Z*)-diastereoisomerisation in imines, oximes, hydrazones.

Trigonal pyramidal "inversion" including amines (low energy barriers) and phosphines (high energy barriers)

Higher order polytopal rearrangements falling under the general term of "*pseudorotation*" with specific variants *e.g.,* Berry pseudorotation in PF_5

eV: 0.   0.5   1.   1.5   2.   2.5   3.

kJ mol⁻¹: 0   20   40   60   80   100   120   140   160   180   200   220   240   260   280

kcal mol⁻¹: 0   5   10   15   20   25   30   35   40   45   50   55   60   65   70

$t_{1/2}$ at 298 K: 1 ns   1 µs   1 ms   1 s   1 h   1 yr   1 kyr   1 Myr   1 Gyr   1 Tyr   1 Pyr

$t_{1/2}$ at 423 K: 1 ns   1 µs   1 ms   1 s   1 h   1 yr   1 kyr   1 Myr   1 Gyr   1 Tyr   1 Pyr

Figure 1.1 Isomerism relationships. (a) Based upon global properties of a molecule, isomerism is divided into the two mutually exclusive categories of constitutional isomerism and stereoisomerism. Further, stereoisomerism can be divided into two mutually exclusive subcategories of enantiomerism and diastereomerism. (b) Alternatively, stereoisomerism can be described from an atom-centric level. Characteristic types of atom-centric stereoisomerism and associated $R_{st}^c 1$ mechanisms can be defined bused upon the coordination and geometry of the stereocentre. The traditional concepts of conformational isomerism and configurational isomerism represent a continuum of phenomena with each being applied to different thermal regimes where the associated $R_{st}^c 1$ processes are operative as indicated by the blue and yellow shaded regions. The positions of the dotted vertical lines indicate suggested property delineations, with "fluxionality" reflecting an interconversion between stereoisomers that is extremely rapid, "practical isolability" representing species stable for 1000 s at 298 K, and "configurationally stable" as stable for 24 hours at 423 K. IUPAC recognises stereoisomerism and stereoisomerisation types with respect to the common thermal regimes. Polytopal rearrangements, with torsions included can describe all possible classes of atom-centred stereoisomerism and associate stereoisomerisation.

and understanding of the dynamical and chemical-reaction behaviour of molecules. Whilst this approach has been eminently practical – and at its heart chemistry is foremost a practical science – it has nonetheless introduced a "common conditions" aspect into the classification scheme and mixes a reaction kinetics definition with a geometry-only definition.

A widely used subdivision of stereoisomerism where this mixing up of geometry with reaction kinetics occurs is in the concepts of *conformational* isomerism and *configurational* isomerism. Conformers are stereoisomers that interconvert (stereoisomerise) unimolecularly at ambient temperatures and pressure on the timescale of seconds or faster. Configurational isomers are defined as being configurationally "stable" at typically encountered temperature with some more informal "rule of thumb" definitions describing them as requiring bond breakage and reformation. The obvious issue with this subclassification is that the kinetics of stereoisomerisation represent a continuum of possibilities. As such, where does the divide lie? Any answer to this question inevitably introduces a degree of arbitrariness into the definition and reflects the "common conditions" bias and the desire for useful and practical definitions.

From a practical perspective, these subdivisions of stereoisomerism exist so that chemists can classify, name, and describe the "essential" aspects of some chemical. For example, if you have a bottle of a simple solvent like *n*-hexane, it is sufficient to know that each constituent molecule has six carbon atoms bonded sequentially with their valence satisfied by hydrogen atoms. Explicit knowledge that, at room temperature, these hexane molecules exist in a dynamic equilibrium of hundreds of different conformations interconverting *via* unimolecular stereoisomerisation reactions on the timescale of nanoseconds is typically far from the chemist's mind if all they are interested in is its utility as a solvent. Notwithstanding this perspective, technology is increasingly demanding the need to explicitly include such fine-grained knowledge into our endeavours. Technology allows us access to temperature and pressure conditions that depart far from ambient ones but, and perhaps more significantly, the rising prominence of machine-led chemistry has made chemists re-evaluate how the subject is approached. Frequently, it is no longer sufficient to rely on the vaguely described guiding principle of "chemical intuition": machines need all the details underpinning such intuition spelled out. As such, being able to generate a comprehensive catalogue of isomeric possibilities, no matter how large, empowers such research efforts and promises a way to transcend the limitations of the individual.

From another perspective, in a world of increasing regulation, the need for the complete cataloguing of isomers and allowed unimolecular reactions is becoming more important. There is a trend for regulatory authorities to demand more precise definitions and characterisation of chemical substances. This is especially the case for patent definitions where, for example, agencies such as the United States Food and Drug Administration (USFDA) and the European Medicines Agency (EMA) now require, where possible, single stereoisomer drugs.

A central issue for all these applications concerns the way that molecular structures and reactions are represented in databases. Some databases, such as crystallographic ones, store the coordinates (Cartesian or reciprocal space coordinates) of all, or a majority, of the atoms in the chemical or material. Such databases contain extensive data pertaining to isomerism, but typically contain no data pertaining to isomerisation. Yet despite intense effort, this level of structural characterisation is available for only a very small fraction of the tens of millions of compounds that have been synthesised.

The more common ways in which isomerism information is stored in databases is via 2D chemical structure representations – typically using "canonical" bond topology descriptions augmented with stereodescriptors – and in-line molecular descriptors such as SMILES or systematic chemical names. Each representation is equivalent, and software is available for converting between them. The Chemical Abstracts Service (CAS) database, for example, represents the bonding topology of a molecular structure as an "adjacency matrix" but converts this to the familiar 2D representation – technically, a "molecular graph" – for viewing by users. In addition, there is an increasing need for the generation of an approximate 3D geometry from each description, with such functionality now becoming standard for the popular chemical-drawing software packages. These approximate molecular geometries provide a starting point for numerous applications such as X-ray structural refinement, drug discovery platforms, and for high-level quantum chemistry calculations, *etc*.

All these approaches rely upon explicit understanding about what isomeric possibilities exist. Any deficiency in that understanding has a deep impact throughout all chemistry, and anything that depends upon it, by limiting both machine and human approaches alike. The authority that oversees this is the International Union of Pure and Applied Chemists (IUPAC). IUPAC provides the required sets of recommendations with these comprising their "Colour Books", each of which focusses on a sensibly delineated

chemistry subdiscipline or topic. For example, the "Blue Book"[2] is a compendium for organic chemistry nomenclature, the "Red Book"[3] for inorganic chemistry nomenclature, and the "Gold Book"[4] a compendium of chemical terminology. Whilst IUPAC is careful to call the contents of these compendia as "recommendations" – reflecting the fact that the field of chemistry is a constantly evolving discipline – these are, nonetheless, in large part treated as "rules". This is especially true the further one departs from "purely academic research". For example, governmental regulatory authorities treat the IUPAC recommendations as the basis for legal frameworks including those for patent specification.

The selection of current IUPAC recommendations pertinent to the material presented in this Thesis is reproduced in Appendices A and B, with some 2D stereochemistry drawing recommendations placed in context in Chapter 5. These reproduced recommendations include very many concepts that underpin chemistry teaching and practice and, whilst these could be considered as "common knowledge", in this work their precise definition and purpose often becomes the centrepiece of discussion.

Chemistry practice, and the IUPAC recommendations, has developed over two and a half centuries of discoveries of new chemical phenomena. These phenomena comprise the discovery and understanding of new compounds, new reactions, and a growing understanding of the underlying principles – what Kuhn first called "paradigm shifts"[5] – to rationalise what was observed. For the major subdisciplines of chemistry, there has been a degree of independent development with, for example, similar yet distinct systems of nomenclature for organic and inorganic compounds. In each instance, the system is fit-for-purpose, but it does highlight the fact that the understanding and approach to chemistry in its entirety is *not* a fully unified discipline.

This is particularly true of isomerism and associated unimolecular isomerisation. Organic chemistry, for all its rich diversity of structure and ubiquitous application, is nonetheless fixated on a single element – that of carbon. Isomerism and unimolecular isomerisation of organic compounds thus reflects what is typical of organic compounds. Tetrahedral coordination is typically the highest coordination geometry, with Walden inversion of configuration at a tetrahedral stereocentre requiring bimolecular processes where bonds are broken and made. Concerted unimolecular "inversion of configuration" at a tetrahedral carbon atom typically requires extreme energies and is thus outside the usual discussions of organic reactions. In contrast, bond torsion-related processes for organic

chemistry typically have very low barriers and present a broad landscape of conformational possibilities accessible at ambient temperatures – indeed, much biochemistry critically relies on this. This is well studied, and its teaching is part of the foundational concepts for chemistry students. An unfortunate side effect of this is that it can engender an unconscious misconception that these narrowly described isomerism and associated unimolecular isomerisation characteristics are representative of that for the rest of the periodic table. Indeed, for organic systems where there is a notable departure from the normal expectations, there is much fuss over it with special new terms being invented to account for it. As an example, rotamers and atropisomers are special cases of conformational torsional stereoisomers where the "rotation" (more correctly called a torsion) is described as "hindered" in contradistinction to typical torsional barriers being very low in energy. Even when armed with this more nuanced picture of torsional stereoisomerism and isomerisation, the recognition of the well-established examples of alkenes and imines as "hindered torsional isomers" tends to get glossed over or is simply ignored.

Chemists that start to specialise in inorganic chemistry soon learn that for inorganic systems, a much richer diversity of isomerism-related phenomena and wider range of energy regimes that isomerisation can operate at is seen as typical rather than rare or special cases. It is not surprising then that inorganic chemistry has developed a somewhat more general approach to isomerism and unimolecular isomerisation.

Whilst the current state of isomerism and unimolecular isomerisation makes sense from a historical perspective, from the perspective of physical chemistry, many of these subdiscipline-dependent idiosyncratic practices and definitions make little sense. Just as the periodic table brought order to the set of known chemical elements, a systematic, self-consistent, and general approach to isomerism and associated unimolecular isomerisation is thus very appealing.

A way forward, which is described in this Thesis, relies upon the mathematical fields of Combinatorics and Graph theory. Graph Theory, and its application to chemistry began in the later nineteenth century with the representation of molecules as bond topologies between atoms, *i.e.*, the 2D molecular structures that all chemists use *are* molecular graphs with few chemists today, apart from those practicing cheminformatics, recognising them for what they are. Anyone who has ever drawn the 2D representation of a molecule or performed a substructure search of a database has implicitly done so using Graph Theory techniques.

Graph Theory has been applied in numerous other ways to chemistry problems, one being the enumeration of constitutional isomers. For over half a century there has been an expectation that Combinatorics-Graph Theory applications could greatly advance chemistry, but this has not been realised. A critical issue limiting advancement was expressed by Vladimir Prelog (1975 Nobel Prize in Chemistry for "his research into the stereochemistry of organic molecules and reactions"), in the preface to the influential 1976 book edited by Balaban,[6] describing "the gap (or even abyss) between the chemical and mathematical literature".

Unifying the separate topics of constitutional isomerism and stereoisomerism through application of a Combinatorics-Graph Theoretic approach is the objective of this Thesis. It looks towards the development of a single formalism that unifies all aspects and requisite language for precise descriptions. Its central aim is to be able to count, name, represent, and catalogue all possible isomers, as well as define the unimolecular reaction mechanisms that interconvert them.

Concerning constitutional isomerism, IUPAC currently makes no recommendations pertaining to counting and cataloguing all possible constitutional isomers. There has been an extensive mathematical literature pertaining to this task, but this has not led to the development of practical schemes. A reason for this is that the number of constitutional isomers possible increases superexponentially up to the order $\mathcal{O}(2^{N^2})$ with the number of atoms N.

The starting impetus for this Thesis was the synthesis and basic characterisation, during 2006, by the author of this Thesis, Peter Canfield, in collaboration with his thesis advisers, other researchers at The University of Sydney, and researchers at The Australian National University, of compounds that could not be classified using existing IUPAC stereochemical recommendations. The initial work performed during this Thesis was the completion of the compound characterisation and the development of basic understanding as to why organic compounds could be made that could not be classified. This included a generalisation and mathematical completion of an IUPAC-recognised concept for isomerism and unimolecular isomerisation called the polytopal-rearrangements model of stereoisomerisation. Publication of the original work, plus the extensions done during this Thesis, then appeared in the publication: "A new fundamental type of conformational isomerism", Peter J. Canfield, Iain M. Blake, Zheng-Li Cai, Ian J. Luck, Elmars Krausz, Rika Kobayashi, Jeffrey R. Reimers & Maxwell J. Crossley, Nature Chemistry volume 10, pages

615–624 (2018).[7] Following on, this Thesis considers the consequences of this discovery for the understanding of known organic compounds and for the development of new stereoisomeric compounds that could not previously be anticipated. It also considers the underlying mathematical abstraction of isomerism and unimolecular isomerisation, and the centrality of Combinatoric and Graph Theory for its specification. This was further abstracted and generalised to provide a basis for application to constitutional isomerism. In doing so, the resulting formalism unifies both major branches of isomerism under a single conceptual framework.

1.2   Stereogenic units and composite stereoisomerism

As mentioned earlier, enantiomerism (chirality) and diastereomerism are general types of stereoisomerism that are dependent upon *global properties* of the molecular system. A useful and alternative approach relies upon the definition of stereoisomers as having the same constitution, *i.e.*, the molecular connectivity (bond topology) remains the same for stereoisomers but their arrangements in space differ. From this definition, it is possible to focus attention on individual atoms or a small grouping of atoms and how these are bonded to the remainder of the molecule. This picture is *composite stereoisomerism* as the overall stereoisomerism of a molecule can be defined as being composed of its constituent parts and the global properties

Focussing on individual atoms, critical stereoisomeric properties are its coordination number and the coordination geometry. A further elaboration concerns whether the substituents to the atom are distinguishable. For the familiar case of a tetrahedral atom, if all four substituents are distinct or the interchange of any pair of substituents leads to a different stereoisomer, this is variously referred to as a single centred stereogenic unit, assymetric atom, or chirality centre. This principle applies more generally to atoms of arbitrary coordination number though increased scope of possibilities exist for coordination numbers of 5 or higher. For example, in a trigonal bipyramidal structure, ligand – ligand relationships are more complex with interchange of ligands occupying axial and equatorial positions leads to diastereomers and not enantiomers. Similarly, for octahedral complexes there exist *cis* and *trans* relationships between ligand positions providing scope for diastereomeric relationships in addition to enantiomeric ones.

The torsional relationships between contiguous atoms, where they exist, are also fundamental. To define a torsion, four atoms with no three contiguous collinear atoms is a requirement. Any four atoms that meet this requirement can be considered as a stereogenic unit, though IUPAC formally requires "stability" of the arrangement as an additional condition. Atropisomers, rotamers and *E/Z* configurations about double bonds all conform to this definition. The implicit requirement for stability in the single centred stereogenic unit, and explicit requirement in the case of torsional stereoisomers serves an additional example of where geometry and isomerisation mechanism become mixed.

For the purposes of this Thesis where general geometric principles of isomerism are examined, the geometric principles underlying the concept of the stereogenic unit can be generalised, regardless of whether substituent atoms are distinct or not. For convenience, throughout this Thesis such centres are called "stereocentres". The justification for this being that substituent atoms around a central atom are assigned distinct indices so that they can be tracked.

As the size of the stereocentre increases, so do the number of structural motifs that it can support and the number of unimolecular reactions available to interconvert between them. Some of these features for larger coordination numbers can often be recognised as aspects from smaller stereocentres that have been combined in some new and characteristic form. Fundamental stereoisomerism and stereoisomerisation processes are those that cannot be represented in terms of properties of smaller stereocentres. When the stereoisomerism of a molecule can adequately be described in terms of the properties of multiple stereocentres, then this is described as a composite stereoisomerism form.

1.3   Key historical discoveries pertaining to stereoisomerism

Some key discoveries concerning stereoisomerism include:

(1)   enantiomerism arising at a tetrahedral[8-9] or higher coordination number[10] chirality centre, as enunciated by Van't Hoff and Le Bel, to explain Pasteur's 1848 observations pertaining to optical activity;[11]

(2)   the understanding of, what would later be termed, *E/Z* torsional stereoisomerism involving double bonds in organic compounds, established around 1890[12-13] in *cis/trans* isomers;

(3) torsional stereoisomerism about single bonds (classical conformational stereoisomerism),[14-15] including atropisomerism ("hindered" torsional stereoisomerism producing isolable compounds, discovered in 1914[16-17]);

(4) pyramidal "inversion" – typically of amines[18] – and its "hindered" form as experimentally demonstrated in 1961;[19-20] and

(5) Fluxionality observed in phosphorus pentahalides[21] in 1953 and stereomutation in other 5-coordinate phosphorus compounds providing inspiration for the polytopal-rearrangements model of stereoisomerisation.

The consequence of this historical phenomenological development of basic concepts for stereochemistry is that the current IUPAC rules and recommendations constitute an *ad hoc*, though authoritative, collection, *i.e.*, they do not arise from the consistent and systematic application of a single conceptual framework describing all the possible underlying spatial arrangements and all possible concerted unimolecular stereoisomerisation rearrangement mechanisms (for which the symbol $R_{st}^c 1$ is introduced in this work).

Composite-isomerism forms, sometimes referred to as "non-canonical" forms, are describable as combinations of fundamental forms, whilst fundamental forms cannot have *all* their aspects described in simpler terms. Composite-isomerism forms are well known, including phenomena such as the familiar cyclohexane chair-boat as well as other complex forms that can result in stable, individually synthesisable compound atropisomers.[22-25]

1.4 Graph Theory and reaction graphs

Graph Theory is a mathematical subdiscipline that deals with the pair-wise relationships of abstract "objects". Graphs are composed of graph vertices (the "objects") with the relationships between them represented as graph edges, and the combination of vertices and edges is the "graph". Graphs may be "undirected" meaning the edges do not indicate directional relationships between vertices with such edges pictorially indicated as a simple conjoining line. Conversely, a "directed" graph features directed graph edges that indicate some directional or ordered relationship between the graph vertices with directed edges pictorially represented as arrows.

As any elementary chemical reaction or transformation represents one object (the reactant) changing into another (the product), representing chemical reactions as a graph is a natural fit. As an example, consider the reaction **a** → **b**, **c**, or **d** over independent transition states **TS^b**, **TS^c**, or **TS^d**. In physical chemistry, these are often depicted in the familiar and standard chemical form shown in Figure 1.2a. Figure 1.2b shows the same pathway information re-expressed as a reaction graph. If the energies of each of the species (now represented as graph vertices) is associated with each graph vertex, then the graph is equivalent to the traditional form in Figure 1.2a.
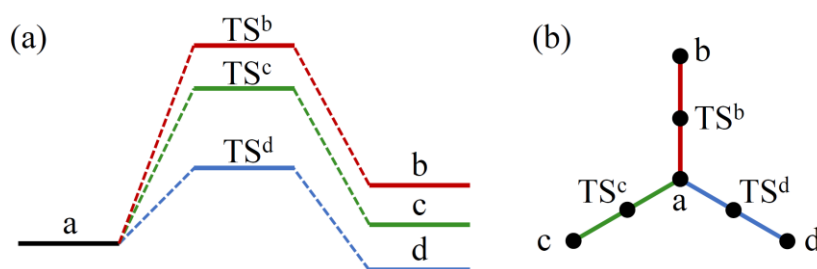
Figure 1.2 Traditional and Graph Theory depiction of chemical reactions. (a) The traditional depiction as typically used in physical chemistry of **a** → **b** or **c** or **d** over transition states **TS^b**, **TS^c**, or **TS^d** to yield products **b**, **c**, and **d**, respectively. (b) The Graph Theory "reaction graph" depiction of (a).

Figure 1.3 shows a related example detailed by the polytopal rearrangements model of stereoisomerism:[7, 26-29] Figure 1.3a shows representations of three of the polytopes of ML$_3$, with Figure 1.3b displaying them as a reaction graph that highlights their allowed interconversion pathways. Representing chemical reactions as graphs has its greatest utility when there are numerous interconnecting processes. Graph Theory provides a suite of powerful tools that can be brought to bear on problems of interest.
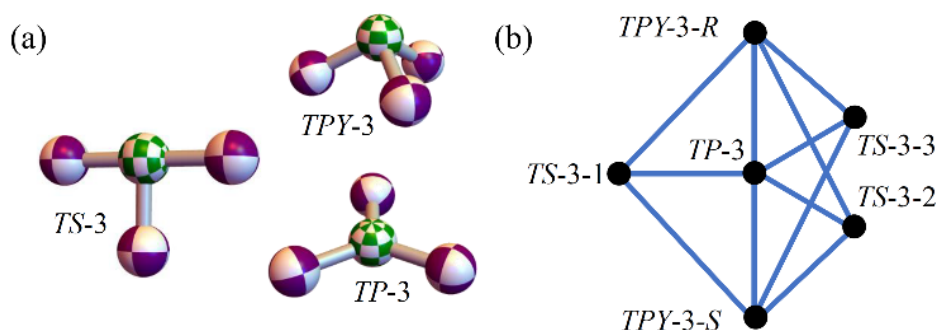
Figure 1.3 Three structural polytopes[7] of ML$_3$ are shown in (a), named by their IUPAC polyhedral symbols[3] (*TS*-3 is T-shaped, *TPY*-3 is trigonal pyramidal, *TP*-3 is trigonal planar), with (b) showing their representation as a reaction graph that highlights their concerted unimolecular reaction interconversion pathways (blue graph edges). Each graph vertex is labelled using IUPAC polyhedral configuration symbols.[3]

1.5   Overview of the thesis

As much of the work comprising this Thesis is pedagogical in nature, precise definitions and language are critical. Chapter 2 takes a more detailed look at some formal IUPAC stereoisomerism and stereoisomerisation-related terminology as listed in the "Gold Book",[4] and terms in common usage that either differ from the formal definitions or are not formally defined by IUPAC. It identifies some ambiguities and shortcomings and, where possible, addresses these. As this Thesis introduces an extensive reappraisal of isomerism within the context of a unifying mathematical formalism, a significant number of necessary new terms are introduced to facilitate a precise and concise description and discussion. These new terms are formally defined in detail and provide a reference for subsequent Chapters.

Chapter 3 seeks to describe a mathematically rigorous and complete framework for defining, cataloguing, and classifying stereoisomerism and unimolecular stereoisomerisation (R_{st}^c1) processes. The resulting framework is called the Polytope Formalism of stereoisomerism applicable to any general ML_n family. Many of the underlying principles for this formalism are an extension of the polytopal-rearrangements model of stereoisomerisation as introduced by Muetterties:[26-29] in the late 1960s onwards and recognised later by IUPAC. As there does not exist a sufficiently comprehensive or focussed description of the Muetterties model, one is provided. From the perspective of stereoisomerism *in general*, deficiencies of the Muetterties model are systematically identified and addressed to provide mathematical rigour and completeness. The Polytope Formalism of stereoisomerism is used to examine the ML_2 family with the concept of "embedding" introduced as a critical feature for manifesting the full scope of stereoisomeric possibilities for this special case. The Polytope Formalism of stereoisomerism analysis of the ML_3 family is also provided to indicate generality.

Chapter 4 presents the experimental demonstration of bond-angle reflexion stereoisomerism in four chirally resolved *transoid* B(F)–O–B(F)-quinoxalinoporphyrin stereoisomers where ML_2 is B–O–B. Focus is placed on characterisation of compounds synthesised prior to commencement of the work presented in this Thesis. This Chapter presents new nomenclature and terminology necessary to describe these synthesised stereoisomers and introduces the term "akamptisomers" to describe them. This Chapter also asks the question: *why were akamptisomers so difficult to unequivocally demonstrate experimentally?*

Chapter 5 asks the question: *what are the minimal molecular features required to guarantee akamptisomerism?* Armed with an answer, a restrictive search of the CAS database was performed to answer the follow-up question: *how common are akamptisomeric molecules and how relevant are they to chemistry and chemistry-related applications such as medicinal chemistr*y? A search is made seeking to identify significant deficiencies in the current treatment of such molecules demanding improvements. This Chapter investigates how these deficiencies can be overcome and describes important implications not just for stereochemistry but also anything that relies upon it such as patent definitions.

Chapter 6 asks the question: *what features do constitutional isomerism and stereoisomerism have in common?* In answering this, a mathematical framework is developed and called the Polytope Formalism of constitutional isomerism. Due to an identified size-scaling problem, an approach is developed to make the formalism tractable and provide practical solutions to problems of interest to chemists. The worked example of tautomerism of free-base porphyrin is presented with all possible bond-topology structures and interconversions listed and the accompanying full potential energy surface generated. A suite of software tools is provided to facilitate general application of this formalism.

In Chapter 7, findings from Chapter 5 are implemented as a drug-design strategy. The questions are asked: *is this strategy useful* and *how widely applicable is it?* A search of the CAS and RCSB.org[30] databases reveal significant opportunities and docking calculations on examples vindicate the premise. Density-functional theory modelling reveals why the strategy has merit. New opportunities for drug patents are identified.

2 Evaluation of existing IUPAC stereochemical terms and new ones introduced in this thesis

2.1 Summary

This Chapter reviews the formal IUPAC stereoisomerism and stereoisomerisation-related terminology as listed in the "Gold Book",[4] and terms in common usage that differ from these formal definitions. Ambiguities and shortcomings are noted. To overcome these deficiencies, herein new formally defined terms used throughout this thesis are introduced.

2.2 Introduction

In one of the first treatises of modern Chemistry, written in 1787, Antoine Lavoisier states: "We cannot improve the language of any science, without, at the same time improving the science itself; neither can we, on the other hand, improve a science, without improving the language or nomenclature which belongs to it".[31] The science of Chemistry encompasses a vast universe of possibilities and, in the two and a half centuries of discovery and understanding, there has been a clear and concerted effort made to formalise the language used to describe the ever expanding and evolving concepts.

To this end, the International Union of Pure and Applied Chemistry (IUPAC) has, since its founding in 1919, sought to act as a single authority on the subject. As stated on their website:[32]

> *"**What We Do**
> The International Union of Pure and Applied Chemistry (IUPAC) is the world authority on chemical nomenclature, terminology… It develops and maintains Recommendations that create a common language for the global chemistry community. The scientific work of IUPAC is conducted largely through a formal Project System in which proposals from chemists around the world are peer-reviewed and, if meritorious, are approved and supported. In addition, IUPAC has many other wide-ranging activities that ultimately impact both the chemical profession and the worldwide community as a whole. These include the publishing of technical reports, journals, books, databases and other information resources that facilitate the conduct of scientific research, conferences, and the provision of awards for the recognition of scientific excellence."*

The recommendations for terminology are collected online in the IUPAC "Gold Book"[4] and comprise an up-to-date collection of formal definitions for Chemists and all national and international authoritative bodies dealing with the description, handling, and processing of chemical concepts or substances.

Despite this and due to different teaching styles and source matériel (text-books, online sources including Wikipedia.org articles, some of which "Gold Book" entries link directly to), the concepts and terminology that comprise some of the *earliest lessons* for students of chemistry frequently differ in subtle and significant ways from the formal IUPAC recommendations. Professional chemists may go on to have successful and productive careers without ever having to systematically readdress such *foundational knowledge*. The result is that the chemistry community expresses a range of subtly different definitions and, given IUPAC represents this diversity of views, there is evidence of this in the IUPAC "Gold Book".[4] Some of these differences and ambiguities are discussed in Section 2.3.

Listed in Appendix A are some selected isomerism-related definitions copied *verbatim* from the IUPAC "Gold Book" and "Green Book". As this thesis deals with the systematisation of aspects of isomerism with a focus on precise language, it is suggested that the reader familiarise themself with these formal definitions.

This Chapter concludes with formal definitions of novel terminology introduced in this thesis and central to the work.

Within the IUPAC terminology, the definition for a stereogenic unit (see Section A.44) is a useful concept to use for this discussion. It explicitly invokes the same "grouping within a molecular entity that may be considered a focus of stereoisomerism" picture that is referred to here as fundamental stereoisomerism. Whilst the formal IUPAC definition is narrow and restricted to 4-coordinate or fewer centres, the term stereogenic unit can be broadened to encompass a centre of arbitrary coordination number ML_n.

2.3   Terms in common usage and their relationships to the IUPAC terms

The notion of isomerism, that is, the relationship between isomers (See Sections 0 and 0), from its earliest conception, has been about the geometric relationships between two or more molecular entities.

In the case of stereoisomerism – a subcategory of isomerism – the connectivity between the constituent atoms remains the same but the spatial arrangements and relationships are different. Stereoisomerism thus concerns itself with differences in geometries. As shown in Figure 1.1b, these differences can be dependent upon global properties or may be defined from an atom-centric perspective where spatial relationships are related by local geometric operators that change one stereoisomer into another. Here the term "local geometric operator" means a geometry altering operation that is applied to a *subset* of atoms and/or the connections between atoms but leaves the local geometric relationships of all other parts of the molecule unaltered. An internal coordinate description of a molecule provides the easiest implementation for a local geometric operator to act upon.

Figure 2.1a shows a *local reflection operator* (indicated by the red disk) acting upon the nitrogen chiral stereogenic unit of (R)-[(R)-2-chloro-2-fluoroethyl](methyl)amine. with all other internal coordinates remaining unchanged including the halogenated carbon stereocentre and backbone conformation.

(a)

(R)-((R)-2-chloro-2-
fluoroethyl)(methyl)amine

(S)-((R)-2-chloro-2-
fluoroethyl)(methyl)amine

(b)
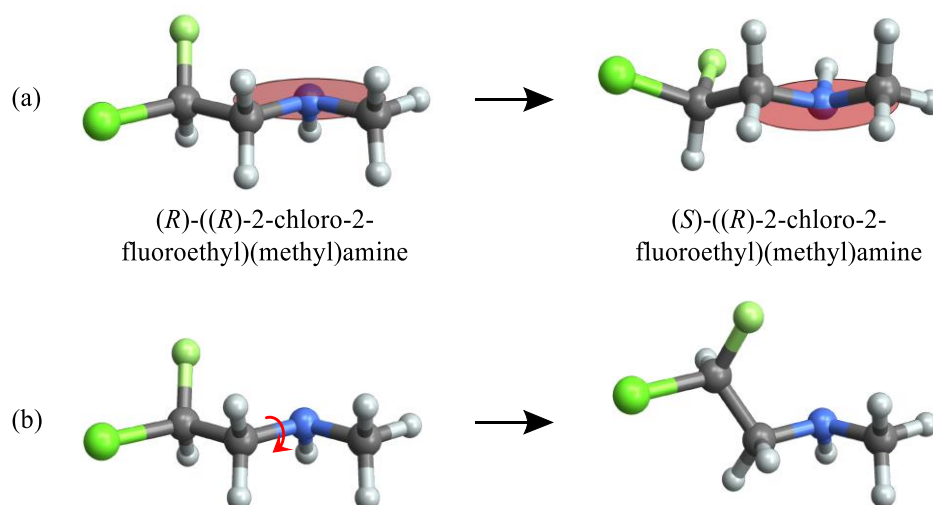
Figure 2.1   The effect of local geometrical operators on the example compound (R)-((R)-2-chloro-2-fluoroethyl)(methyl)amine. (a) A *local reflection operator*, indicated by the red disk, acting upon the nitrogen and immediate substituent atoms with all other internal coordinates remaining unchanged. This results in a change in the absolute configuration of the nitrogen but leaves the carbon stereocentre and overall staggered conformation unchanged. (b) The bond torsion operator indicated by the red arrow changing the backbone conformation.

Similarly, in Figure 2.1b a *local torsion operator* (indicated by the red arrow) is applied to a C–N bond changing the backbone conformation. In each of these examples in Figure 2.1, the stereoisomerism is described in *purely geometrical terms* with no reference to physical mechanism or process. It is important to note that the black arrow indicating the transformations are not referring to actual continuous physical motions.

In contrast, the concept of isomerisation *does* concern itself with real or conceptual continuous physical processes. Isomerisation is about *mechanism.* In general terms, isomerisation can mean *any* mechanism that brings about a change of isomers whether it be concerted, unimolecular, bimolecular or some complex multistep sequence of elementary processes. For constitutional isomerism, *any* isomerisation process necessarily involves a breaking and making of bonds as, by definition, there is a change in connectivity (molecular topology – see Section 2.4.27). In the case of stereoisomerisation where, by definition, there is no *net* change in connectivity, there are two distinct classes of processes; (i) stereoisomerisation mechanisms where the topology undergoes changes and (ii) mechanisms where the topology remains unchanged throughout.

An example of the case (i) is the hypothetical, multistep reaction where an (*S*) stereoisomer is changed to the (*R*) configuration *via* a trigonal planar prochiral carbanion intermediate as shown:

An example reaction scheme showing the (*S*) stereoisomer with COOEt and CN groups, converted via **-H⁺** to a trigonal planar carbanion intermediate, then via **+H⁺** to the (*R*) configuration.

(*S*)	trigonal planar carbanion intermediate
	(*R*)
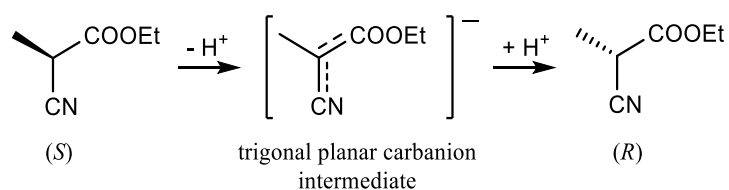
A well-known example of case (ii) is the "inversion" (actually, a reflexion) of NH_3 *via* a trigonal planar transition structure like the example in Figure 2.1:

trigonal planar
transition structure
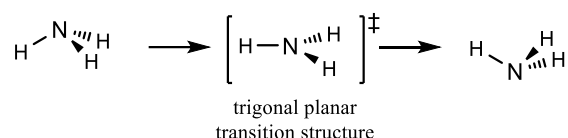
This represents an elementary (see Section A.14) concerted unimolecular process where the motion of the atoms is smooth and continuous and the formal covalent character of the N–H bonds remains, essentially, unchanged. Throughout this thesis, the concept of concerted unimolecular stereoisomerisation (rearrangement) mechanisms characterised by unchanging connectivity ("stereotropic" – see Section 2.4.21) are a central topic. The symbol $R_{st}^{c}1$ is introduced in this Thesis (see Section 2.4.20) to concisely describe this.

The distinction that *isomerism* be only concerned with geometric differences between isomers and *stereoisomerisation* be only concerned with real or conceptual physical processes is not a uniformly adopted view. The mixing up of the two concepts is particularly

prevalent throughout the terminology regarding stereoisomerism. Numerous definitions of stereoisomers and the relationships between them (the stereoisomerism) explicitly include reference to mechanism (stereoisomerisation) – see Appendix A Sections 0, 0, 0, and 0.

The central and long standing example of "conformational isomerism" which is traditionally described as a form of stereoisomerism where the stereoisomers *can be interconverted* by torsions about single bonds, explicitly intermingling the concepts of geometry and mechanism. IUPAC, however, does not formally define "conformational isomerism", instead defining "conformation" (see Section A.6) as "the spatial arrangement of the atoms affording distinction between stereoisomers *which can be* interconverted by rotations about formally single bonds…". By avoiding calling it a type of isomerism, and instead using the phrasing "arrangement of atoms", this definition encompasses all intermediate geometries and is consistent with the IUPAC definition that isomers represent potential energy minima deep enough to accommodate at least one vibrational mode. Despite the careful definitions by IUPAC (see Section A.7), the concept of "conformational isomerism" is commonplace.

Atropisomers and atropisomerism – presumably the relationship between atropisomers – presents an interesting example. Atropisomers are defined by IUPAC (see Section A.1) and, being a subcategory of conformer which in turn is a subcategory of isomer, must represent a local potential energy minimum deep enough to accommodate at least one vibrational mode. The notion that these isomers are defined as arising from *restricted rotation* (see Section A.18) provides an example of where the formal definition of a geometric relationship is now intermingled with mechanism. As a side note, "rotation" in this sense is more precisely called "torsion".

The existence of stereoisomers related by "restricted rotation" had the effect of placing the concept of potential energy barrier to rotation at the centre for all "conformational isomerism" discussions in textbooks[33] and teaching settings. It is not surprising, then, that the concepts of the geometric differences between the stereoisomers and the associated $R_{st}^c 1$ mechanism are often viewed as indistinguishable.

Concerning stereoisomerism, it is divided into two mutually exclusive subcategories conformational isomerism and configurational isomerism. The latter is defined as anything which is not conformational isomerism (see Section A.5). Whilst defining something by what it is not, may be convenient, it suggests an underlying flaw in the conceptualisation.

The obvious question is: *what physical properties or behaviour delineates configurational isomerism from conformational isomerism*?

A frequently cited definition for configurational isomers and, in particular the configuration at a stereocentre, is that they "cannot be interconverted without breaking and making bonds" with this usually taken in the context of the configuration at a tetrahedral carbon stereocentre. Such a definition, whilst reflecting much observed chemistry, is nonetheless biased by the chemistry of carbon and a thermal regime typical of human activities.

As shown later in Chapter 3 and Chapter 4, the distinction between conformational isomers and configurational isomers is arbitrary; it is largely "carbon-centric". Polytopal rearrangements (see Section A.33), which are extensively elaborated in this Thesis, can be defined in a general sense for all stereoisomerisation phenomena. A graphic summary of this is shown in Figure 1.1.

The definition of "stereogenic unit" (see Section A.44) is mathematically incorrect. Definition (2) reads "A chain of four non-coplanar atoms (or rigid groups) in a stable conformation, such that an imaginary or real (restricted) rotation (with a change of sign of the torsion angle) about the central bond leads to a stereoisomer". A mentioned earlier, "rotation" is more precisely called "torsion". Further "four non-coplanar" excludes torsion angles of 0° and 180°. The presumed intended definition should read "A chain of four atoms, no consecutive three of which are collinear…".

A final point concerns the term "inversion" in the context of stereoisomerism. "Inversion", in a mathematical sense, means a change of sign of a scalar quantity, a $\rightarrow$ -a, or for a vector, $(a_1, a_2, a_3…) \rightarrow (-a_1, -a_2, -a_3…)$. Both of these concepts play a role in their use in stereoisomerism.

In chemistry it is common practice to refer to "inversion of configuration" at a tetrahedral centre. This term arising out of the older phenomenological terms "optical inversion" and "Walden's Inversion" where "inversion" was adopted directly from its mathematically correct use in describing the change in sign of optical polarisation-rotation measurements.[34-36] This concept of "inversion of configuration" is further directly applicable to trigonal pyramidal ML_3 centres where the lone-pair acts as a fourth "ligand" thus

emulating a tetrahedral species. In this case, "pyramidal inversion" means the conversion of one stereoisomer to the other.

In the context of geometry, the "inversion" of a vector representing atomic Cartesian coordinates for a molecule is $(x,y,z) \rightarrow (-x,-y,-z)$. Similarly defined is "reflexion", a process which, in the demonstrative case of reflexion in a Cartesian plane, involves only a single Cartesian coordinate changing sign whilst the other two remain unchanged, *e.g.*, for reflexion in the *xy* plane: $(x,y,z) \rightarrow (x,y,-z)$. The "inversion" of configuration as described above for ML_4 and ML_5 species is correctly described as a reflexion and not a geometric inversion. This distinction is critical if one wishes to digitally represent and manipulate aspects of stereoisomerism or stereoisomerisation in databases, drawing and structure manipulating software, and quantum-chemical calculations. Group theory, which forms the basis of much molecular-structure classification including, for example, the ubiquitous X-ray crystallographic data, explicitly uses the mathematical definitions only.

This terminological difficulty is not just associated with interchanges at chiral centres but also applies to ML_2. Historically, the conversion of a bent structure through its linear form to the reflexed bent structure has been known as "bond-angle inversion". This usage parallels that for "inversion of configuration" and "pyramidal inversion". Similarly, the mathematical operator involved in these cases are also all "reflexions" and not "inversions". However, reflexion and inversion, in the mathematical sense, are both processes that could be of relevance in ML_2 systems, so the significance of the ambiguity of terminology is particularly important.

In this thesis, I use only the strict mathematical expression "reflexion" to describe the interchange of these types of isomers and not "inversion" as commonly used by organic Chemists. IUPAC defines inversion, when used as a stand-alone term, as a geometric property involving a centre of inversion (see Section A.22).

2.4   New and revised terms introduced in this thesis

The following terms, introduced in this work, are formally defined here in a manner analogous to IUPAC's terminology.

2.4.1   akamptisomerism

A type of stereoisomerism defined by the relationships between akamptisomers.

2.4.2   akamptisomerisation

Stereoisomerisation process that converts between a pair of akamptisomers.   The process may occur by any mechanism.

2.4.3   akamptisomers

Stereoisomer pair related to each other by bond-angle reflexion only.   Term constructed in analogy to that for atropisomers (see Section A.1). From Ancient Greek ἀ- (á-, privative) + κάμπτω (kámpto, "bend") + ἴσος (ísos, "equal") + μέρος (méros, "part"). "Akamptisomeric" is the adjectival form.

2.4.4   autakamptisomers

An akamptisomeric pair that are structurally degenerate.   Any interconversion between autakamptisomers represents a degenerate rearrangement (see Section A.12). "Autakamptisomeric" is the adjectival form.

2.4.5   bond-angle reflexion (BAR)

A bond angle $\theta < 180°$ changing to $360° - \theta$ as measured in a fixed reference frame. In real molecules this may only be approximately true depending upon additional interactions. BAR may *contextually* describe either a local geometric operator or a physical motion (mechanism).   A BAR mechanism is an $R_{st}^{c}1$ mechanism associated with akamptisomerism and involves a transition state/intermediate conforming to a linear polytope.
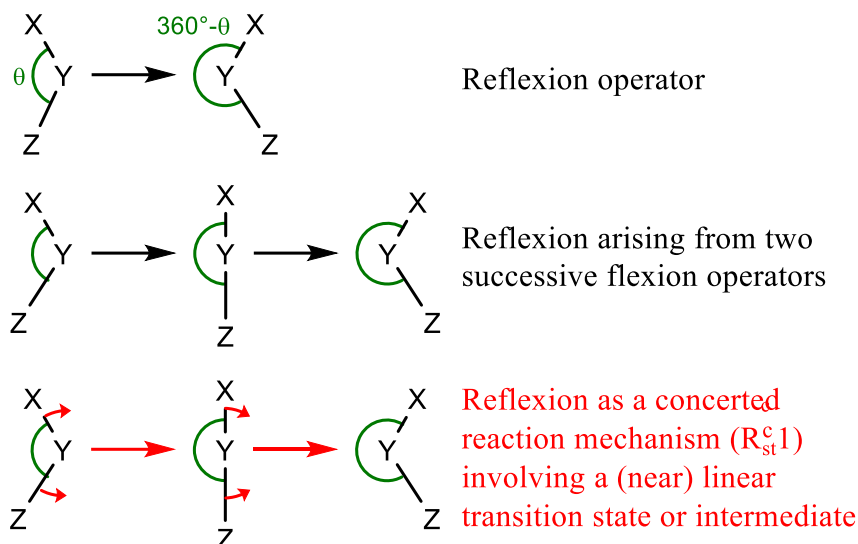
Reflexion operator

Reflexion arising from two successive flexion operators

Reflexion as a concerted reaction mechanism ($R_{st}^c 1$) involving a (near) linear transition state or intermediate

At least one additional connected atom beyond *each* end of the specified bond angle is necessary to embed (see Section 2.4.9) the BAR within the frame of reference, hence five consecutively connected atoms A–X–Y–Z–B, where both A–X–Y and Y–Z–B are not collinear, are required to define a BAR operation or process.

Reflexion operator

Reflexion arising from two successive flexion operators

Reflexion as a concerted reaction mechanism ($R_{st}^c 1$) involving a (near) linear transition state or intermediate
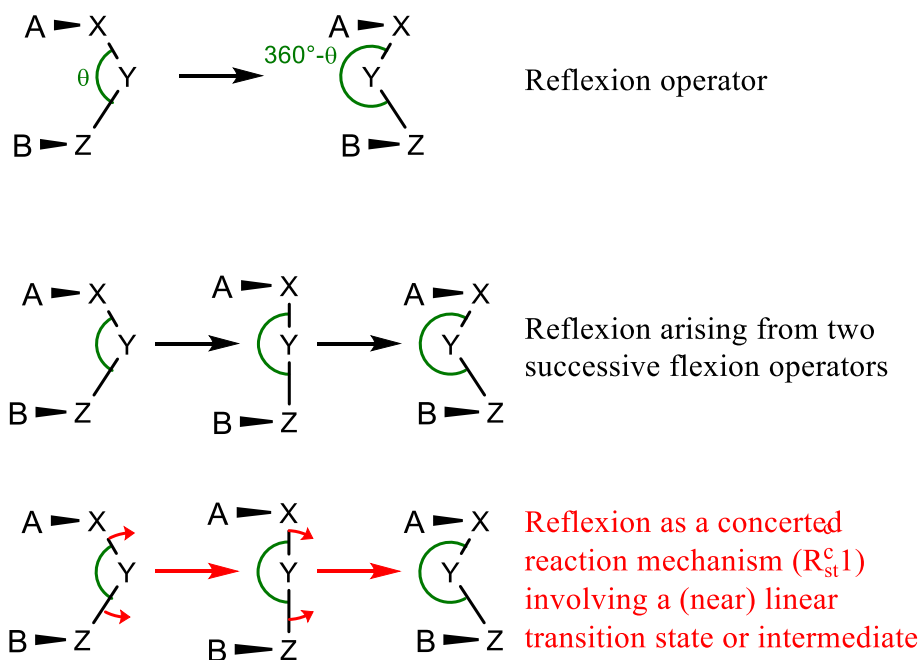
2.4.6   *bond-angle flexion*

A local geometric operator applied to a bond angle X–Y–Z whereby a bond angle $\theta <$ 180° changes to 180°, or the reverse change. Two flexions in succession and the *same direction* result in a reflexion operator (See Bond-Angle Reflexion).

2.4.7   concise polytope set

For the Polytope Formalism of stereoisomerism and a given polytope family ML_n, the concise polytope set are all polytope species for which on the associated graph, no adjacent species share the same generic symmetry point group. A concise set can be generated from a more extensive non-concise set by resolving all adjacent polytope set members with the same generic symmetry point group into a single representative polytope.

2.4.8   demesne

When a Polytope Formalism model is used to describe the relationships between properties, behaviours and objects that comprise a general set of physical phenomena then these are the *demesne* ([pronounced dɪˈmeɪn](#), plural: *demesnes*) of that specific implementation of the Polytope Formalism. For example, Chapter 3 describes the Polytope Formalism for the demesne of *stereoisomerism*.

A property of a demesne is that, for its allowed operators, it exhibits *set closure*. Further, set members that are semi-terminal define key boundaries for the demesne and indicate the existence of a neighbouring or enclosing demesne. For example, the demesne of stereoisomerism exists within the demesne of isomerism.

2.4.9   embedding

In the context of an *n*-coordinate, single atom stereocentre represented generally as ML_n, an embedding is where all ligating atoms form part of an extensively connected semi-rigid network that provides a frame of reference *external* to the stereocentre. Embedding restores rotational invariance to an ML_n stereocentre.

2.4.10 generic polytope graph

Within the context of the Polytope Formalism of stereoisomerism, a *generic polytope graph* is an undirected or directed graph showing the traverses between members of a set of polytope genera for a ML_n polytope family,

2.4.11 hula-twist-like mechanism

A stereotropic mechanism associated with akamptisomerisation that consists of coupled torsions. The bond angle θ does not change significantly throughout the process.
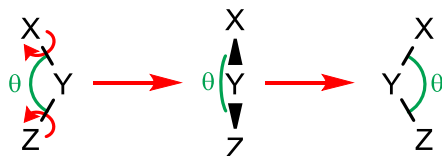
At least one additional connected atom beyond each end of the specified bond angle is necessary to embed the motion within an external frame of reference, hence five atoms A–X–Y–Z–B are required to define the hula-twist-like mechanism in terms of independent A–X–Y–Z and X–Y–Z–B torsions, where A–X–Y, X–Y–Z, and Y–Z–B are all not collinear.

This term is derived in reference to geometrical changes present during the hula-twist (HT) mechanism as defined in the IUPAC "Gold Book"[4] (See Appendix A Section A.19)

2.4.12 local geometric operator

A geometric operator applied to a specified arrangement of atoms within a molecular entity. The internal coordinates describing all other atoms in the molecular entity remain unaltered by the local geometrical operator.

2.4.13 polytope class

In the context of the Polytope Formalism, a polytope class is a high-level taxonomic grouping of all families united by a high-level characteristic. For stereoisomerism, the class is ML_n for all n. For constitutional isomerism, the class is the number of atoms whose bond topology is permuted. This would be described as the *N*-atom bond-topology class. Other descriptors can be applied to constitutional isomerism where simplifying principles are implemented, *e.g.*, for the $\mathcal{S}_m\mathcal{B}_n$ partitioning described in Chapter 6, $\mathcal{S}_m\mathcal{B}_n$ is the class.

2.4.14 polytope family

In the context of the Polytope Formalism, the polytope family is a taxonomic grouping of genera (and their constituent species) that are closely related by some uniform character. For stereoisomerism, ML_n designates the polytope family where n is the coordination number of centre M. M can represent any element. L can represent any single element or member of a set of different elements. For constitutional isomerism, the polytope family, or bond-topology family, are groups of genera that have the same number of bond-topology elements.

2.4.15 polytope genus

In the context of the Polytope Formalism, the polytope genus describes the general form or pattern of a configuration. For stereoisomerism, this equates to a polytope form without any vertex index assignment and IUPAC inorganic nomenclatural polyhedral symbols[3] (or new ones constructed in an analogous manner, as needed) can be used to describe the polytope genus. For constitutional isomerism, the polytope genus, also called the "bond-topology" genus, comprises all symmetry related species with the same general arrangement of bonds.

2.4.16 polytopal rearrangement

The IUPAC definition (see Appendix A Section A.33) defines a polytopal rearrangement as representing a *stereoisomerisation process*. This implies a change between stereoisomers with such isomers necessitating they be local minima on the associated potential energy surface (PES) deep enough to support zero-point vibration. The definition is expanded here to include rearrangements between any arrangement of atoms corresponding to critical points on the PES. A polytopal traverse to an adjacent structure necessarily involves a change in the saddle-point order. For example, a single traverse from a local minimum to a first order saddle point. The corollary of this is that any rearrangements between critical points of the same saddle-point order necessitates an even number of traverses.

2.4.17 polytope species

In the context of the Polytope Formalism, any polytope featuring a unique configuration. For stereoisomerism, this equates to a unique form and vertex index assignment. For constitutional isomerism, this equates to a unique configuration of bonds.

For stereoisomerism, IUPAC inorganic nomenclatural polyhedral configuration symbols[3] can be used to label the polytope species if the vertex indices are used in the same manner as Cahn-Ingold-Prelog priorities.

2.4.18 polytopal traverse

A change (either as a geometric operator or a continuous motion) between *adjacent* polytopal forms within a polytope family. The endpoints of a polytopal traverse can be

characterised by, usually low energy (including imaginary) and large amplitude, normal vibrational modes.

2.4.19 pontation

A molecular design concept whereby the ring-fusion C–C bond of a polyaromatic or related conjugated molecule is replaced by C–X–C bridging group resulting in a bridged annulene. This is referred to as X-pontation, *e.g.*, CH$_2$-pontation of naphthalene gives the methylene-bridged [10]annulene:
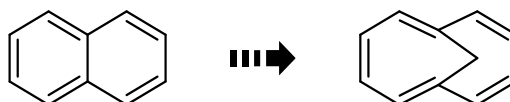
where the special "pontation arrow" used here specifically indicates the alteration of structure brought about through pontation.

From Latin *pons* ("bridge") + *ation* (verb suffix). *Pontate* is the participle of pontation and *pontated* is the past participle.

2.4.20 $R_{st}^c 1$

Chemical reaction mechanism symbol denoting a concerted unimolecular stereotropic rearrangement elementary reaction mechanism. Being a stereotropic process, no bonds are broken or made during the $R_{st}^c 1$ reaction. An $R_{st}^c 1$ mechanism can correspond to a single traverse between two critical points on a potential energy surface or can be used to describe two contiguous traverses as would be the case, for example, between two local minima on a potential energy surface. **R** = rearrangement, **c** = concerted, **st** = stereotropic, **1** = unimolecular. Symbol constructed in an analogous manner to S$_N$2, E1, *etc*.

e.g., "inversion" (reflexion) of tertiary amine

(*R**)   trigonal planar transition structure   (*S**)
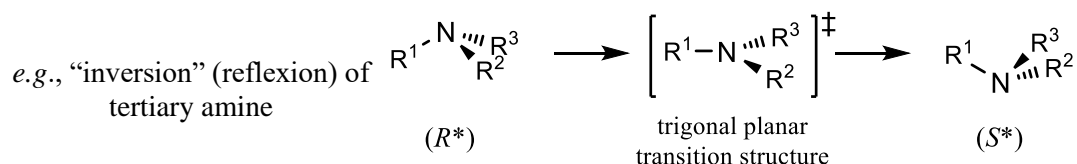
2.4.21 $R_{to}^{c}1$

Chemical reaction mechanism symbol denoting a concerted unimolecular topotropic rearrangement elementary reaction mechanism. Being a *topotropic* process, the bonding connectivity must change during the $R_{to}^{c}1$ reaction.
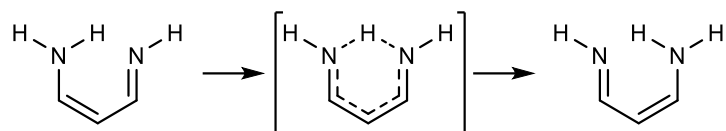
2.4.22 seed polytope

Within the context of the Polytope Formalism of stereoisomerism, for a given polytope family ML_n the *seed* polytope genus is the polytope genus of highest symmetry, *e.g.*, for ML_2 it is *L*-2, for ML_3 it is *TP*-3, for ML_4 it is *T*-4, and for ML_5 it is *TBPY*-5. For a seed polytope genus, each of the corresponding species can be referred to as seed polytope species.

2.4.23 semi-terminal property and semi-terminal polytope

Within the context of the Polytope Formalism of stereoisomerism, a polytope genus or species is semi-terminal if one direction of at least one non-dissociating vibrational normal mode does not lead to a polytope within the set of those under consideration. Semi-terminal polytopes delineate the boundaries of the set with respect to set-closure. The directed graph representation of this is a self-loop.

2.4.24 specific polytope graph

Within the context of the Polytope Formalism of stereoisomerism, a *specific polytope graph* is a directed graph showing all traverse connections between the full set of polytope species belonging to a set of polytope genera for a ML_n polytope family. The requirement for the graph to be directed is to accommodate polytope species with semi-terminal properties.

2.4.25 stereotropic

Adjective pertaining to a change in stereoisomeric form and does not involve bond breaking and bond making. From Ancient Greek στερεός (stereós, "solid") + τροπικός (tropikós, "of or pertaining to a change). Not to be confused with stereotopic.

e.g., stereotropic reaction
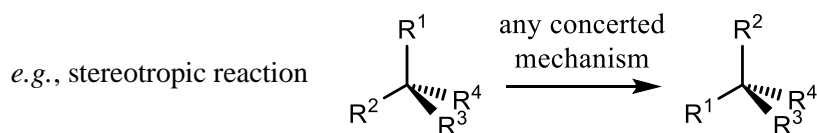
2.4.26 strepsisomerisation

A $R_{to}^{c}1$ isodesmic bond-walk mechanism by which a central polytopic group bonded within a macrocyclic ligand leads to successive ligation isomers as this central group rotates coherently within the macrocycle cavity. From Ancient Greek στρέψις (strépsis, "a turning [inward]") + ἴσος (ísos, "equal") + μέρος (méros, "part") + "isation" (verbal suffix for a process).

The following example shows the four step strepsisomerisation of a peripherally substituted B(F)OB(F)-porphyrin.
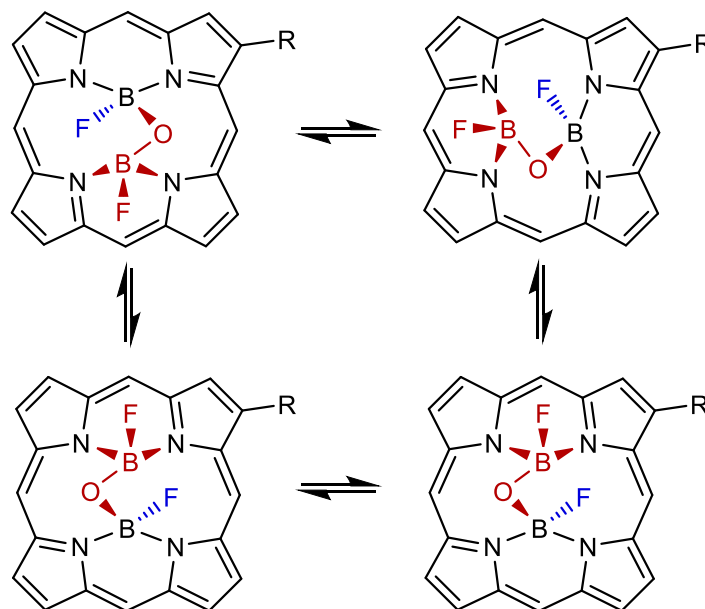
2.4.27 topology (molecular topology, bond topology)

In the context of a discrete molecular entity, topology means the bonding connectivity between the constituent atoms. The default meaning of "bonding" here is covalent bonding with bond order ignored. The term may be extended to refer to other types of bonding such as hydrogen bonding with such details needing to be explicitly noted.

2.4.28 topotropic

Adjective pertaining to a change in bonding topology of a chemical system. Constitutional isomerism is wholly concerned with topotropic differences. From Ancient Greek τόπος (tópos, "place, locality") + τροπικός (tropikós, "of or pertaining to a change).

2.4.29 torsion (bond), torsional angle, dihedral angle

In the context of molecular geometry, a bond torsion (or simply "torsion") describes a rotational angle or motion of atoms or groups about an axis defined by the bond. In practice, four contiguous atoms b^1–a^1–a^2–b^2 where no three consecutive atoms are collinear define the torsional angle (dihedral angle) between the two planes defined by the sets of points $\{b^1, a^1, a^2\}$ and $\{a^1, a^2, b^2\}$.

3   Polytopal Formalism of stereochemistry at a single stereocentre

3.1   Summary

This Chapter introduces conceptual framework capable of describing the full scope of stereoisomerism and associated concerted unimolecular rearrangement processes ($R_{st}^{c}1$) arising at a single atomic centre of arbitrary coordination number. The conceptual framework is called the Polytopal Formalism of stereochemistry and builds upon the Muetterties polytopal-rearrangement model of stereoisomerisation developed from 1969 onwards, in large part, to account for fluxional behaviour observed in molecules such as PF_5. Whilst the polytopal-rearrangement model of stereoisomerisation, which is described in detail, provides useful insights for both conformational isomerism and configurational isomerism, it has a notable deficiency in not being capable of describing torsions about bonds – the very process at the heart of all classical conformational isomerism that so dominates organic chemistry.

In seeking to address this, and other deficiencies, the Polytopal Formalism, as presented here, is developed through the formal inclusion of such torsions to the existing polytopal-rearrangement model. Further, the inclusion of torsions allows for the concept of "embedding" that provides an external frame of reference for the spatial arrangements. The result is a conceptual framework is capable, *in principle*, of comprehensively describing both the full scope of geometries and the geometric changes available at a single atomic centre of arbitrary coordination number.

Armed with this, centres with low coordination number are examined in detail and it is revealed that the ML_2 case presents unique properties. The Polytope Formalism demonstrates how these novel properties can be manifested in real molecules with significant implications for Chemistry. The ML_3 case is also described in detail indicating the generality of the formalism.

3.2 Introduction

This Chapter focuses on stereoisomerism and, more broadly, on the different arrangements that bonded atoms can take up about a central atom. In this chapter and due to the mathematical treatment presented, this is called a "stereocentre" regardless of whether it is a stereogenic unit (see Section A.44) as formally defined by IUPAC. In geometric terms, stereoisomerism can be divided into two general classes: torsional stereoisomerism (see Section A.48) where the different stereoisomers only differ by a specific dihedral angle, hence requiring at least three consecutive non-collinear atoms, and polytopal-rearrangement stereoisomers that can be defined as the different ways a set of ligating atoms or molecular fragments can be arranged around a single central atom. On first inspection, these two classes may appear to be distinct and unrelated. Further, the condition that a dihedral angle only has meaning when the constituent atoms are not collinear may appear innocuous. Much of the forthcoming development is concerned with understanding the consequences of three atoms becoming collinear, in terms of how polytopes are defined and how they describe the associated stereoisomerism.

A general model describing torsional isomerism is straightforward in just requiring a single torsion angle φ to be defined about an axis formed by two bonded atoms a^1 and a^2 of the non-collinear group of atoms b^1–a^1–a^2–b^2. A geometric operator can be defined to change φ by an arbitrary amount. Further, a concerted unimolecular stereotropic rearrangement mechanism ($R_{st}^c 1$ – see Section 2.4.20) can be associated with this torsional isomerism that uses the same definition of φ as the reaction coordinate.

Polytopal-rearrangement stereoisomerism at a single central atom is, however, more complicated and is described in detail in Section 3.2.1. An essential aspect of the polytopal-rearrangement model is that, for an atom of coordination number *n*, it treats the associated stereoisomerism as a highly reduced system with the central atom and its *n* ligating atoms reduced to point-like objects and, historically, the geometry idealised to a high symmetry object called a polytope. This model describes, *in principle*, all the distinct stereoisomerically meaningful ways the ligating "atom" points can be arranged around a central "atom" point. These geometries correspond to arrangements of atoms representing stereoisomers (potential energy local minima) and transition structures (saddle points of varying orders). Associated $R_{st}^c 1$ mechanisms describe the transformation of one polytope to another.

Together, these two separate models can describe all stereoisomerism at a single stereocentre. In this Chapter it is shown that both can be combined into a single mathematically based model called the Polytope Formalism of stereochemistry. This formalism can, *in principle*, provide a complete and systematic exploration of the full scope of stereoisomerism at a single atomic centre. Significantly, it is not subject to anomalies attributable to singularities that occur when three or more atoms involved in a torsion become collinear, or any other similar effect related to polytope identification. It is from the perspective of this formalism that the discovery of bond-angle reflection isomerism, called akamptisomerism and described in Chapter 4, was revealed to represent a new fundamental type of stereoisomerism.

3.2.1 *A modern perspective on the polytopal-rearrangement model*

Despite the significant body of literature on the polytopal-rearrangements model of stereoisomerisation, a single clear and comprehensive description is lacking. The following reviews some of the history of the model and attempts to describe it in sufficient detail for the reader to understand the Polytope Formalism of stereoisomerism on which it is based. To facilitate a clear and precise description, new terminology, deemed essential, is introduced to help organise the conceptual elements into a clear framework.

From the perspective of coordination number, organic molecules exhibit only few distinct basic geometries at an atomic centre, despite the vast multitude of known organic compounds. Typically, the coordination number is limited to four or fewer except for some exotic or fleeting species. In contrast, inorganic compounds exhibit a far richer diversity of both coordination numbers and basic geometric forms. Consequently, the challenges of description, classification, and nomenclature for any stereochemistry arising in higher coordination number species are far greater.

For much of its history, chemistry has dealt with emerging novel stereochemical phenomena in a case-by-case manner, as needs demanded. In the 20th century, research in chemistry surged, especially after World War II with powerful investigational tools such as X-ray crystallography and NMR becoming widespread. This revealed an increasingly complex diversity of stereochemical structural and dynamical phenomena exhibited by inorganic species. Given all this diversity and, in particular, Berry's 1960 hypothesised

pseudorotation mechanism to explain the fluxional behaviour in PF_5,[37] there was a clear need for some unifying approach to give order to the subject.

This led Muetterties and Guggenberger[26-29, 38-43] to pioneer the *polytopal-rearrangements model of stereoisomerisation* from the late 1960s onwards, initially in large part, to account for the observed "stereochemical nonrigidity" in numerous 5-coordinate phosphorus species[38, 43-48] including chelating ligands.[49-50] The early work also examined the case of 5-coordinate carbon species.[51] Further development of the model led to its generalisation to describe dynamical behaviour in 3-coordinate p-block[52] species; 4-coordinate s-block, [53-54] p-block[54-57], and d-block[54, 58-62] species; 5-coordinate d-block[63-70] and further work on p-block[71-73] species; 6-coordinate d-block[39, 74-79] and p-block[80] species; 7-coordinate general[42] and d-block[81-83] species; 8-coordinate[40, 82, 84-87] species; and the 9-coordinate[41] system. Mingos,[88] further, provides an in-depth look at the early concepts and application of the polytopal-rearrangements model of stereoisomerism from a structural correlation perspective for 5- to 8-coordinate systems. Due to its success and obvious utility in all these instances, this *conceptual model* was recognised and adopted in 1996 by IUPAC[89] for its *generality* in describing both geometry and non-dissociative stereoisomerisation processes.

The polytopal-rearrangements model of stereoisomerisation (the "Muetterties model") presents a physics-inspired mathematical treatment[26-29, 43] that separates out stereoisomerism, as being a property only of atomic spatial arrangements and geometric operators, from non-dissociative stereoisomerisation, the possible non-dissociative reactions that interconvert between stereoisomers. Further, it *defines* key "minimal distortion"[90] R_{st}^c1 pathways that must *always* be *conceptually available* for any specific set of stereoisomers. Structure and transformative reactions become simple, independent concepts that are, nevertheless, intrinsically related to each other.

The essential elements in this picture are *polytopes*, the fundamental types of idealised geometrical arrangements of bonded point-like "atoms" that can be arranged as objects of one, two, or three spatial dimensions as linear forms, regular polygons and convex polyhedra, respectively. These polytopes are typically used as *simplified representations* for real molecular systems of coordination number n that conform to the generalised formula ML_n. Traditionally, Muetterties[38, 40-41] called such systems the n-atom *family*. In this work, ML_n is called the polytope *family* (see Section 2.4.14). Within this context, "L" represents the atom

bonded to "M", whether it is part of a larger molecular fragment or not, and each of the "L" atoms can be different in the real molecule. Further, "M" can represent any element, not just metals.

Within this model, all M–L bonds are treated equally, and bond order has no meaning. The nature of the bonding also has no meaning; it could be covalent, ionic, van der Waals, *etc.* – so long as it meets the requirement that it supports zero-point vibration (see Section A.31). Further, as the model is focussed on *non-dissociative* phenomena, any changes in M–L bond lengths are deemed inconsequential, *i.e.*, the connectivity or *bonding topology* (see Section 2.4.27) remains constant.

There are several different ways of graphically representing a polytope, each having strengths and weaknesses. The various IUPAC "Red Book", section IR.9.3.2, recommendations for inorganic nomenclature depict common polyhedra to describe coordination geometries and are reproduced in Appendix B.[3] Figure 3.1a shows bonding to a tetrahedral centre as typically represented in 2D using the wedge-hash convention[91] for indicating stereochemistry when drawing molecular structures. This is typical for organic molecules and for inorganic molecules with coordination numbers of 6 or fewer. In Figure 3.1b only the edges (in blue) of the defining tetrahedron are indicated, making the overall structure of the polytope (the tetrahedron) the obvious feature. This "edge-only" style is usually only common when depicting inorganic species of high coordination number where, otherwise, a multitude of radiating bonds from the centre would be visually confusing. For smaller coordination numbers, this edge-only approach can cause a different type of confusion where, for example, in the case of a "square planar" polytope it could be mistaken for cyclobutane. In Figure 3.1c, both styles are combined with distinct colours differentiating

| (a) | (b) | (c) | (d) |
| M-L bonds | polytope edges | bonds and edges | 3D form |
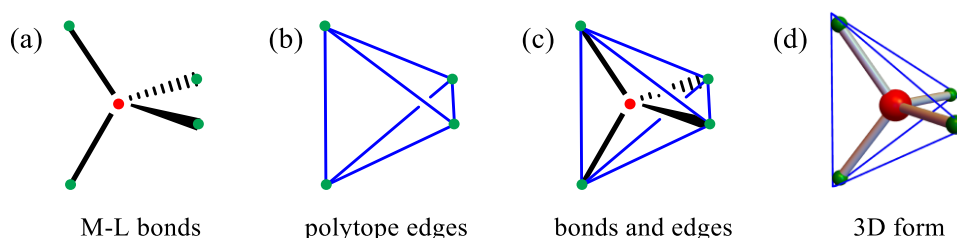
Figure 3.1   Equivalent representations of the tetrahedral polytope genus, *T*-4, for ML$_4$.   a) Typical 2D representation using the wedge-hash convention[89] with only the M–L bonds showing.   b) Representation showing only the edges of the polytope. c) Representation combining both (a) and (b). d) 3D representation of (c). The red dots and sphere represent the central atom "M" and the green dots and spheres represent the ligating atoms "L".

the respective bond and edge features. Figure 3.1d shows a more realistic 3D representation of Figure 3.1c with the 3D visual cues helping to make the spatial relationships clearer for a wide diversity of cases. This will be the style used throughout this chapter.

In this work there is an important conceptual distinction between the *general shape* of a polytope and when that same shape has a specific arrangement (configuration) of *distinct* indices assigned to the polytope's vertices. When discussing *only* the shape, the polytope is referred to as a polytope *genus* (plural: *genera*; adjectival form: *generic*. See Section 2.4.15), for example, Figure 3.2b shows the tetrahedral genus and its IUPAC nomenclatural polyhedral symbol *T*-4 (see Appendix B: IUPAC "Red Book" polyhedral symbols and structures.). The *T*-4 genus belongs to the ML_4 *family* (Figure 3.2a). When vertex indices are specifically assigned, the polytope is referred to as a polytope *species* (see Section 2.4.17). Figure 3.2c shows the two distinct ways of assigning nondegenerate indices to a tetrahedron. If the indices are used in the same way as the Cahn-Ingold-Prelog (CIP) priorities, then chemical nomenclatural (*T*-4-*S*) and (*T*-4-*R*) polyhedral configuration symbols (IUPAC "Red Book" [3], section IR.9.3.3) can be assigned to each species and provide a convenient and unique labelling. It is due to the use of nondegenerate indices being assigned to the polytope vertex "atoms" that the term stereocentre is used throughout this Chapter.

| (a) | (b) | (c) |

ML_4

Tetrahedral *T*-4 (*T*-4-*S*) (*T*-4-*R*)

Family **Genus** **Species**
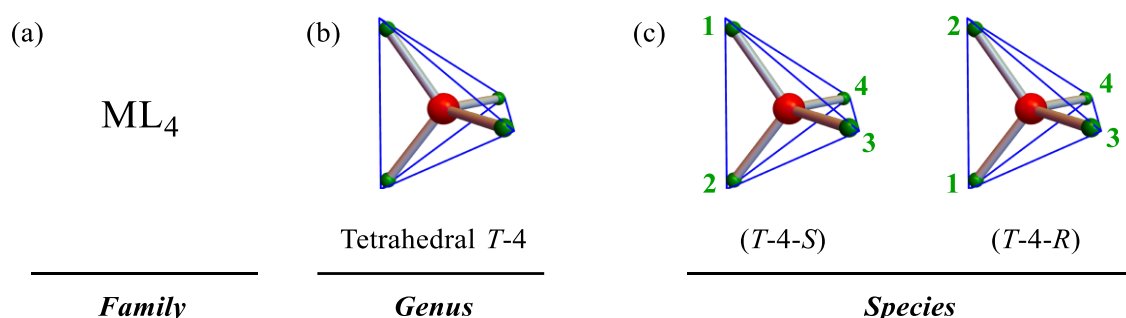
Figure 3.2 Distinctions between the polytope *family*, *genus* and *species* concepts. (a) ML_4 is the polytope *family*. (b) The tetrahedral *genus T*-4 only refers to the general shape. (c) The two distinct ways of indexing the vertices correspond to distinct polytope *species*. Chemical nomenclatural (*T*-4-*S*) and (*T*-4-*R*) polyhedral configuration symbols[3] are assigned to each species using the vertex indices to represent CIP priorities,

Within the Muetterties model, the distinction between genus and species is important when cataloguing different forms and when discussing symmetries. Polytope genera will typically exhibit a high degree of symmetry. For example, the tetrahedral *genus* shown in Figure 3.2b exhibits the T_d symmetry point group. The two tetrahedral *species* shown in Figure 3.2c, however, have their symmetry broken by the indices assigned, even if the polytope vertices represent ligands that are equivalent. When referring to a *species*, its

generic symmetry is that of the corresponding *genus*, for example, the generic symmetry of a species of (*T*-4-*S*) configuration is that of the *T*-4 genus, namely, T_d.

In this Chapter, genera and species are assigned descriptive symbols based on the IUPAC inorganic nomenclatural polyhedral symbols and polyhedral configuration symbols[3], respectively. Where existing symbols do not suffice, new ones are devised following the IUPAC-led principle. *Generic* symbols have the form *ABC-n*, where *ABC* is some letter or letter-number combination describing the general shape, *n* is the coordination number, all letters are capitalised and italicised, whereas all number are presented in plain text. Notably, there are no parentheses around the symbol. *Specific* symbols take the form (*ABC-n-x-y*) where *ABC* and *n* are for the genus, and *x*, and sometimes *y*, are used to indicate a specific arrangements of vertex indices. In Figure 3.2, the symbols assigned are indicative.

An illustrative 4-coordinate example is shown in Figure 3.3 for the (*S*) enantiomer **1a** of a chiral organosilane,[92] where the silicon atom exhibits tetrahedral coordination.
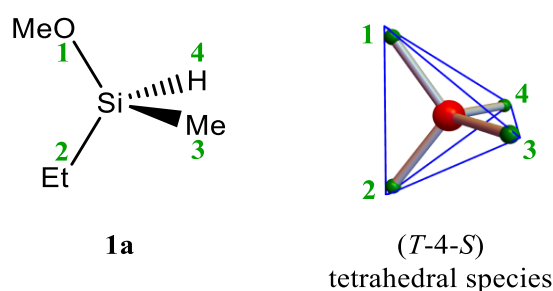
Figure 3.3 The organosilane[92] **1a** and the corresponding polytopal representation of the stereochemistry at the Si atom as a tetrahedral polytopal species within the ML_4 family. The CIP priorities of the ligands to Si are used as the vertex indices. The red sphere represents the Si atom and the smaller green spheres each of the four ligating atoms from the substituent groups O (OMe), C (Et). C (Me), and H. The blue lines describe the edges of the tetrahedron.

This simplification of a real molecule to this minimalist model representation also requires that all the "L" points now be considered "physically identical" yet assigned a *unique* index, *i.e.*, the details of their chemical nature are no longer important, but the "L" points can be tracked by their indices. In the Muetterties model, the polytope is idealised as the highest symmetry version. For the example **1a** shown in Figure 3.3, the symmetry of the *generic* tetrahedral polytope is the T_d point group regardless of the precise geometry of the real molecule or that the substituents on the Si atom are distinct. In Figure 3.3, each vertex has been assigned a unique index (green coloured numerals). The indexing-system choice

can be arbitrary but, for the example of **1a**, the CIP priority rules are implemented for convenience.

In the traditional implementation of the model and for a given polytope genus, a simple classical (molecular) mechanics-like model for this geometry is made and a vibrational normal-mode analysis performed. Traditionally, the low energy (including imaginary valued), large amplitude vibrational modes were recognised as corresponding to initial reaction-coordinate vectors along some feasible $R^c_{st}1$ path leading to a *different* polytope *genus* within the same ML_n *family*.[29] For example, Figure 3.4 depicts the tetrahedral polytope representation of **1a** undergoing the deformations arising from a normal vibrational mode comprised of a pair of orthogonal antisymmetric bond-angle flexional motions. The corresponding *generic* symmetry is T_d for which this motion has *e* symmetry. One direction of the transition vector describing this deformation is congruent with one endpoint along the $R^c_{st}1$ reaction coordinate leading to a square planar polytope (having the generic *SP*-4 polyhedral symbol), here specifically representing the structure **2a** with absolute configuration (*SP*-4-2). As was the case for the *T*-4 genus, the highest symmetry *generic* polytope representation of **2a** is used, here being the square planar polytope having the D_{4h} intrinsic symmetry point group.

| **1a** | (*T*-4-*S*) | (*SP*-4-2) | **2a** |
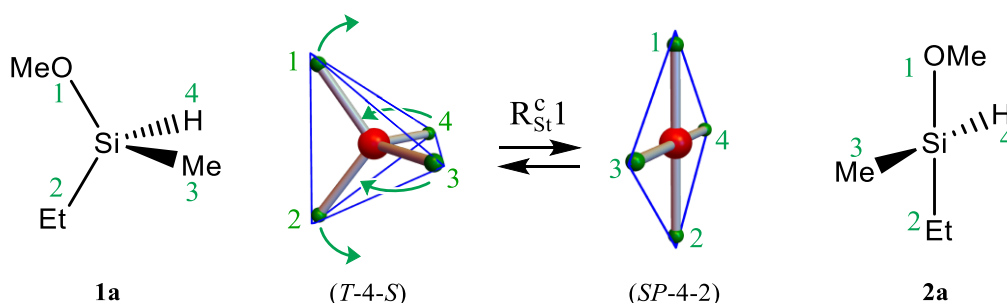
Figure 3.4 Structure **1a** is represented as a tetrahedral polytope species with vertex indices as marked. Structure **2a** is represented as a square planar polytope species with the vertices as marked. For the tetrahedral and square planar polytopes, vibrational normal modes of *e* and b_{2u} symmetry, respectively, are congruent with the corresponding end points of the $R^c_{st}1$ process. This motion is characterised as a pair of orthogonal bond-angle symmetric flexions as indicated by the green arrows.

Similarly, for the square planar polytope species in Figure 3.4 representing **2a**, one direction of the transition vector describing the motion of the unique vibrational mode (of symmetry for the *SP*-4 genus) is congruent with the other endpoint of this same $R^c_{st}1$ mechanism. As the generic symmetry here is D_{4h}, the symmetry of this vibrational mode is b_{2u}. A "minimal distortion" single-step process like the one depicted in Figure 3.4 is called a polytopal "traverse" by Muetterties[29] and is formally defined in Section 2.4.18.

37

For the tetrahedral example in Figure 3.4, the orthogonal bond angles undergoing antisymmetric flexion are (1–M–2) and (3–M–4) – or more concisely written as (12,34) – and are represented by the green arrows in Figure 3.4. In Figure 3.5a, these same bond-angle antisymmetric flexional motions for the *SP*-4 arrangement are shown continuing in the *same direction* leading, again, to a *T*-4 arrangement. The different directions of the normal vibration mode transition vector are congruent with forward and backward $R_{st}^c 1$ mechanisms, each leading to a *T*-4 species. Examination of the configuration of the vertex indices for the two *T*-4 species shows that they are different.

Given that *all* vertices of the *generic T*-4 polytope are equivalent, however, any of the three possible *combinations of pairs* of vertices can be chosen to undergo this type of orthogonal antisymmetric double bond-angle flexions with the resulting *SP*-4 structures shown in Figure 3.5b. Each of the three possibilities will lead to a *different SP*-4 arrangement of the ligands and each correspondingly assigned a distinct nomenclatural polyhedral configuration symbol using CIP priorities.

Figure 3.5b shows that structure **1a**, having the (*S*) absolute configuration, can be converted by the *same* general $R_{st}^c 1$ mechanism to **1b**, having the (*R*) absolute configuration, *via* each of the three different possible *SP*-4 geometries: **2a**, **2b** and **2c**. The same general $R_{st}^c 1$ mechanism amounts to what is traditionally called an "inversion" of configuration of a tetrahedral species. The *mathematically* correct term describing this process is a *reflection* of configuration.

(a)

antisymmetric paired flexions of (12,34)

e → b_{2u}

$R_{st}^c 1$

Tetrahedral

antisymmetric paired flexions of (12,34)

b_{2u} → e

$R_{st}^c 1$

Square Planar

Tetrahedral

Priority sequence 1, 2, 3, 4
CIP priorities OMe > Et > Me > H

2b = (*SP*-4-3)

(13,24)

(13,24)

(b)

1a = (*T*-4-*S*)

(12,34)

2a = (*SP*-4-2)

(12,34)

1b = (*T*-4-*R*)

(14,23)

(14,23)

2c = (*SP*-4-4)

(c)

T-4   *SP*-4

(d)

(*SP*-4-3)

(*SP*-4-2)   (*T*-4-*R*)

(*T*-4-*S*)

(*SP*-4-4)
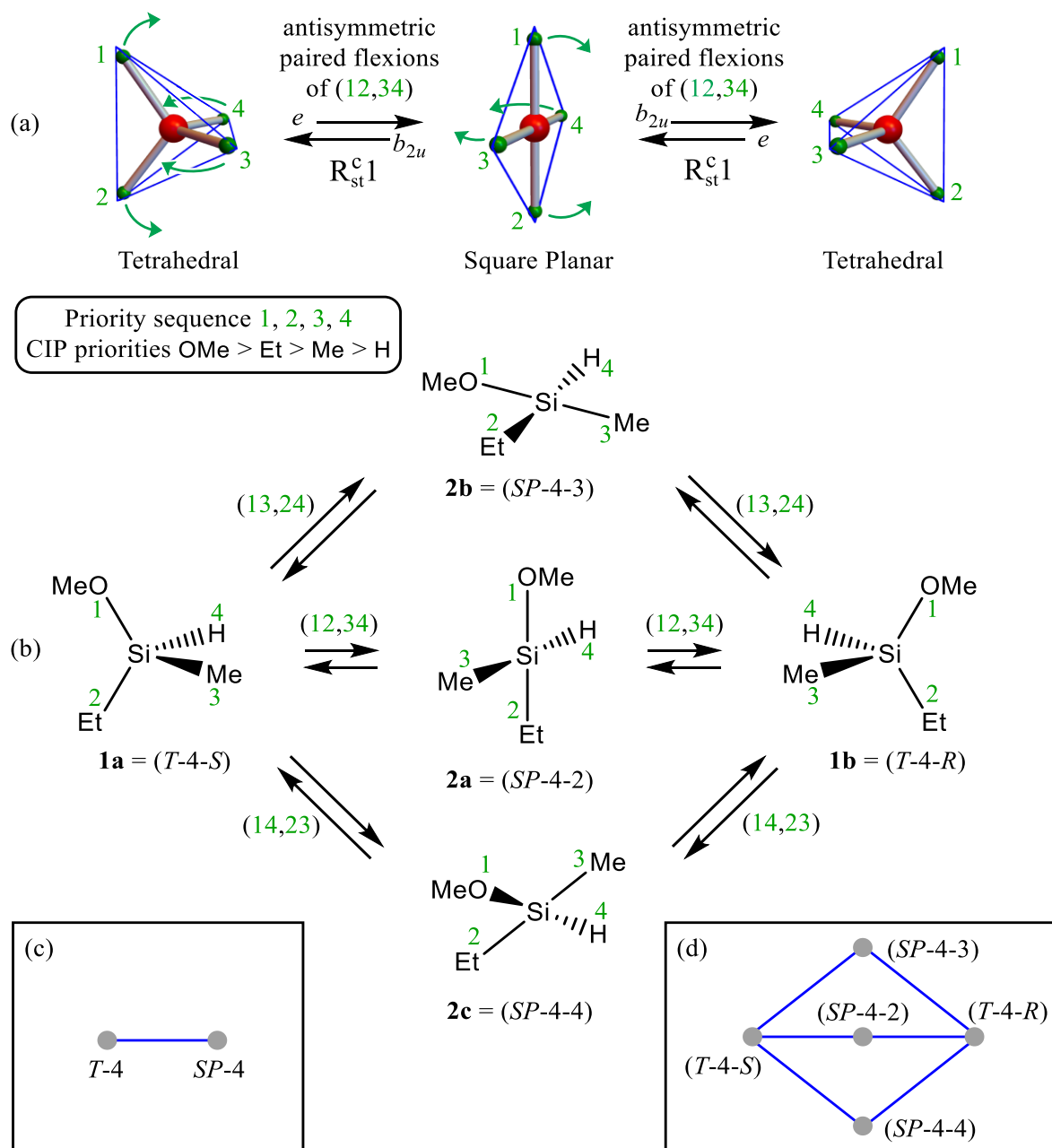
Figure 3.5 Polytopal rearrangements of tetrahedral and square planar polytopes. (a) Transformation of one tetrahedral polytope species to the other *via* a square planar polytope species. For each of the two tetrahedral polytopes on the left and right, the vibrational normal modes of *e* symmetry involving flexion of orthogonal bond angles 1–M–2 and 3–M–4, or simply as (12,34), are congruent with the beginning of the $R_{st}^c 1$ processes leading to the square planar polytope in the centre. Similarly, for the square planar polytope, each direction of the vibrational normal mode of b_{2u} symmetry acting upon the two angles (12,34) is congruent with an endpoint of each of these same $R_{st}^c 1$ processes. (b) Details of the rearrangement of **1a** to **1b** showing the three different choices of pairs of bond angles giving distinct square planar species **2a**, **2b**, and **2c**. Standard IUPAC nomenclatural polyhedral configuration symbols[3] are listed for each species. (c) The polytope *generic* graph representation of the scheme in (b). (d) The graph showing the relationship between each polytope species for *T*-4 and *SP*-4 using the polyhedral configuration symbols used as labels.

In the language of the polytopal-rearrangement model of stereoisomerisation, this path analysis shows *all* the possible ways that the *T*-4 species **1a** can be transformed into the *T*-4 species **1b** *via* polytopal rearrangements involving only *SP*-4 intermediates. Further, it is important to note the immediate correlation between the nomenclatural configuration requirements that have been established[3] and the vertex index permutations for each polytope genus. There are only two vertex index permutations for *T*-4 and hence only two polyhedral configuration symbols: (*T*-4-*R*) and (*T*-4-*S*). Similarly, there are only three vertex index permutations for *SP*-4 and hence only three corresponding polyhedral configuration symbols: (*SP*-4-2), (*SP*-4-3), and (*SP*-4-4). Figure 3.5b shows how these nomenclatural symbols arise by the geometric perturbations described.

Figure 3.5c shows what, within the mathematical discipline of Graph Theory, is called the *graph* of polytopal rearrangements of *T*-4 and *SP*-4 polytope genera for the ML_4 family. This is referred to here as the *generic polytope graph* and is formally defined in Section 2.4.10. As only two genera are included, this graph only shows that these polytope genera are *adjacent*. Figure 3.5d shows the complete graph of all the polytope species belonging to the *T*-4 and *SP*-4 genera and utilises the corresponding IUPAC nomenclatural polyhedral configuration symbols[3] using the vertex indices as CIP priorities.

In this worked example of **1a** → **2** → **1b** as shown in Figure 3.5, compounds **1** having *T*-4 coordination geometries represent PES local minima and compounds **2** having *SP*-4 coordination geometries represent first order saddle point transition structures. This same analysis and resulting graph equally apply to compounds where the *SP*-4 geometries are potential energy surface (PES) local minima and the *T*-4 geometries are transition structures, for example as in some Pt(II) complexes[93]. It is also important to note that, whilst the example detailed here deals with a 4-coordinate centre featuring four distinct substituents, the same analysis is equally applicable to systems where some, or indeed all, of the ligand atoms or groups are *equivalent*. The use of assigned unique vertex indices provides for the tracking of ligand atoms as if they were distinct.

An obvious question arises: *are there other polytopal forms within an ML_n family that are possible or need to be considered*? Historically, the answer would depend upon what the purpose was for invoking the polytopal-rearrangement model. In the case of fluxional molecules, on which most of the current literature is focussed,[42-44, 46-48, 63, 72-73, 77, 80, 85-87, 94] typically only one or, at most, two additional polytope genera, in addition to that describing

the lowest energy geometry, are needed to be considered. Later, this Chapter describes, in principle, a fully comprehensive picture of potential stereoisomeric forms and processes. Consequently, *all* possible polytopes require consideration and a formal definition of "what constitutes a polytope" introduced.

The literature does not directly address the question of how to define a polytope from first principles. Typically, the starting point is the empirical observation of stable molecular geometries and reference to Valence Shell Electron Pair Repulsion[95] (VSEPR) Theory models. Whilst this starting point has direct practical application, it nonetheless avoids the critical question. Later, in Section 3.3, this question is addressed directly.

Staying with the ML_4 family for now, at least two additional polytope genera are recognised from VSEPR-like models: the "M"-apical square pyramidal polytope genus (*SPY*-4) and the see-saw polytope genus (*SS*-4) as shown in Figure 3.6a and Figure 3.6b. Additional planar non-VSEPR-like polytope genera are introduced here: the *half hexagonal planar* polytope genus (given the *HHP*-4 polyhedral symbol) shown in Figure 3.6c, derived from a vibrational normal mode of the *SS*-4 polytope genus of a_2 symmetry and named for its shape; the *gyfu planar* polytope genus (given the *GYP*-4 polyhedral symbol) shown in Figure 3.6d, derived from a vibrational normal mode of the *SP*-4 polytope genus of b_{2g} symmetry and named after the Anglo-Saxon rune of the same shape; and the *algiz planar* polytope genus (given the *ALP*-4 polyhedral symbol) shown in Figure 3.6e, also derived from a vibrational normal mode of the *SP*-4 polytope genus of e_u symmetry and named after the Elder Futhark rune of the same shape. Each of these additional polytopes embody a set of stereochemical relationships distinct from the known ML_4 VSEPR-like polytopes. Whilst no *stable* molecules featuring these unusual 4-coordinate geometries are currently recognised, these genera, generated by the same vibrational-mode-following principle described by Muetterties, serve as evidence for the existence of additional forms beyond those traditionally considered.

Additional VSEPR-like polytopes:

(a)   (b)

SPY-4
"M"-apical square
pyramidal polytope genus

SS-4
See-saw
polytope genus

Some planar non-VSEPR-like polytopes:

(c)   (d)   (e)

HHP-4
Half hexagonal planar
polytope genus

GYP-4
Gyfu polytope genus
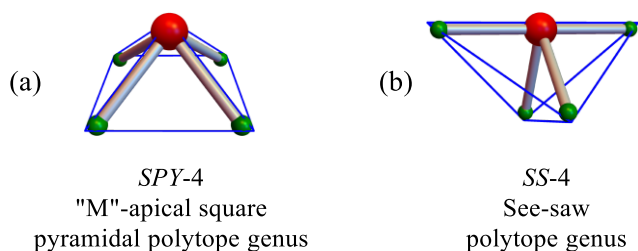
ALP-4
Algiz polytope genus
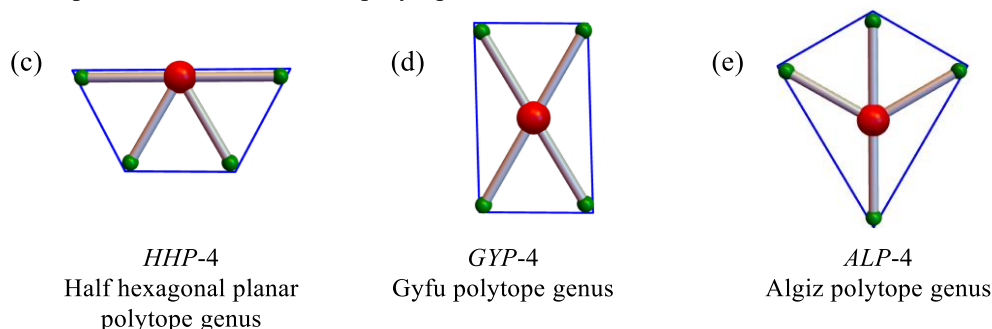
Figure 3.6   Some additional known ML_4 polytopes. Top row: Additional VSEPR-like polytopes; (a) the "M"-apical square pyramidal (*SPY*-4) and (b) the see-saw (*SS*-4) genera. Bottom row: Some planar non-VSEPR-like polytopes; (c) the half-hexagonal planar (*HHP*-4) polytope, gyfu planar (*GYP*-4) polytope, and algiz planar (*ALP*-4) polytope genera.

A detailed analysis, like the approach depicted in Figure 3.5 can be made for the ML_4 family including these additional polytope genera. A full listing of vertex index permutations for each polytope genus gives the total number of polytope species. Using the vertex indices as CIP priorities, the standard IUPAC nomenclatural polyhedral configuration symbols,[3] augmented by analogously defined ones for the *HHP*-4, *GYP*-4, and *ALP*-4 genera, can be assigned to each. A full tabulation of the 41 polytope species is given in Table 3.1 for these seven polytope genera of the ML_4 family as discussed in this chapter.

Not all polytope genera are accessible by a single-step traverse from a different form. In the language of Graph Theory, the genera are not *adjacent*. For example, to change a *T*-4 polytope to its nearest *SPY*-4 polytope necessitates a two-step traverse *via* a *SS*-4 polytope. This is because there is no vibrational normal mode (or linear combination of modes of the same symmetry) that transforms a *T*-4 form directly to a *SPY*-4 form. The results of working through this path analysis and defining *all* possible minimal distortion traverses between polytopes in the set can then be represented as a graph. Such a graph showing the relationships between all species in the VSEPR-like polytope genera (*T*-4, *SP*-4, *SPY*-4 and *SS*-4) are shown in Figure 3.7.

Table 3.1 Full tabulation of all 41 polytope species comprising tetrahedral, square planar, "M"-apical square pyramidal, see-saw, half hexagonal planar, gyfu planar, and algiz planar polytope genera of the ML$_4$ family. The polyhedral symbol,[3] generic symmetry point group and number of symmetry operations for each polytope genus are given. Using vertex indices as CIP priorities, the standard IUPAC polyhedral configuration symbols[3] are given and provide convenient and distinct polytope labels. No standard polyhedral configuration symbols currently exist for the half hexagon planar, gyfu planar, or algiz planar with those listed here in blue being constructed in an analogous manner to those already defined by IUPAC. Chirality and total species for each polytope genus are indicated.

Polytopal form:	Tetrahedral	Square planar	"M"-apical square pyramidal	See-saw	Half hexagonal	Gyfu planar	Algiz planar
Generic polyhedral symbol	*T*-4	*SP*-4	*SPY*-4	*SS*-4	*HHP*-4	*GYP*-4	*ALP*-4
Generic point group	T_d	D_{4h}	D_{4v}	C_{2v}	C_{2v}	D_{2h}	C_{2v}
Generic symm. operations	24	16	8	4	4	8	4
Species configuration symbols [a]	(*T*-4-*R*) (*T*-4-*S*)	(*SP*-4-2) (*SP*-4-3) (*SP*-4-4)	(*SPY*-4-2-*C*) (*SPY*-4-2-*A*) (*SPY*-4-3-*C*) (*SPY*-4-3-*A*) (*SPY*-4-4-*C*) (*SPY*-4-4-*A*)	(*SS*-4-12-*C*) (*SS*-4-12-*A*) (*SS*-4-13-*C*) (*SS*-4-13-*A*) (*SS*-4-14-*C*) (*SS*-4-14-*A*) (*SS*-4-23-*C*) (*SS*-4-23-*A*) (*SS*-4-24-*C*) (*SS*-4-24-*A*) (*SS*-4-34-*C*) (*SS*-4-34-*A*)	(*HHP*-4-12) [b] (*HHP*-4-13) [b] (*HHP*-4-14) [b] (*HHP*-4-23) [b] (*HHP*-4-24) [b] (*HHP*-4-34) [b]	(*GYP*-4-12-3) [c] (*GYP*-4-12-4) [c] (*GYP*-4-13-2) [c] (*GYP*-4-13-4) [c] (*GYP*-4-14-2) [c] (*GYP*-4-14-3) [c]	(*ALP*-4-12) [d] (*ALP*-4-13) [d] (*ALP*-4-14) [d] (*ALP*-4-23) [d] (*ALP*-4-24) [d] (*ALP*-4-34) [d]
Species chirality	chiral	nonchiral	chiral	chiral	nonchiral	nonchiral	nonchiral
Totals	2	3	6	12	6	6	6

a: Polytope species nomenclatural polyhedral configuration symbols[3] under the assumption of distinct vertices. *b*: (*HHP*-4-*ij*) where *i*–M–*j* is the linear grouping and *i* < *j*. c: (*GYP*-4-*ij-k*) with *i* = 1, *j* is *trans* to *i*, *k* is closest adjacent vertex to *i*. *d*: (*ALP*-4-*ij*) with *i* the "post" of the rune away from the clustered vertices and with *j trans* to *i* and *i* < *j*.
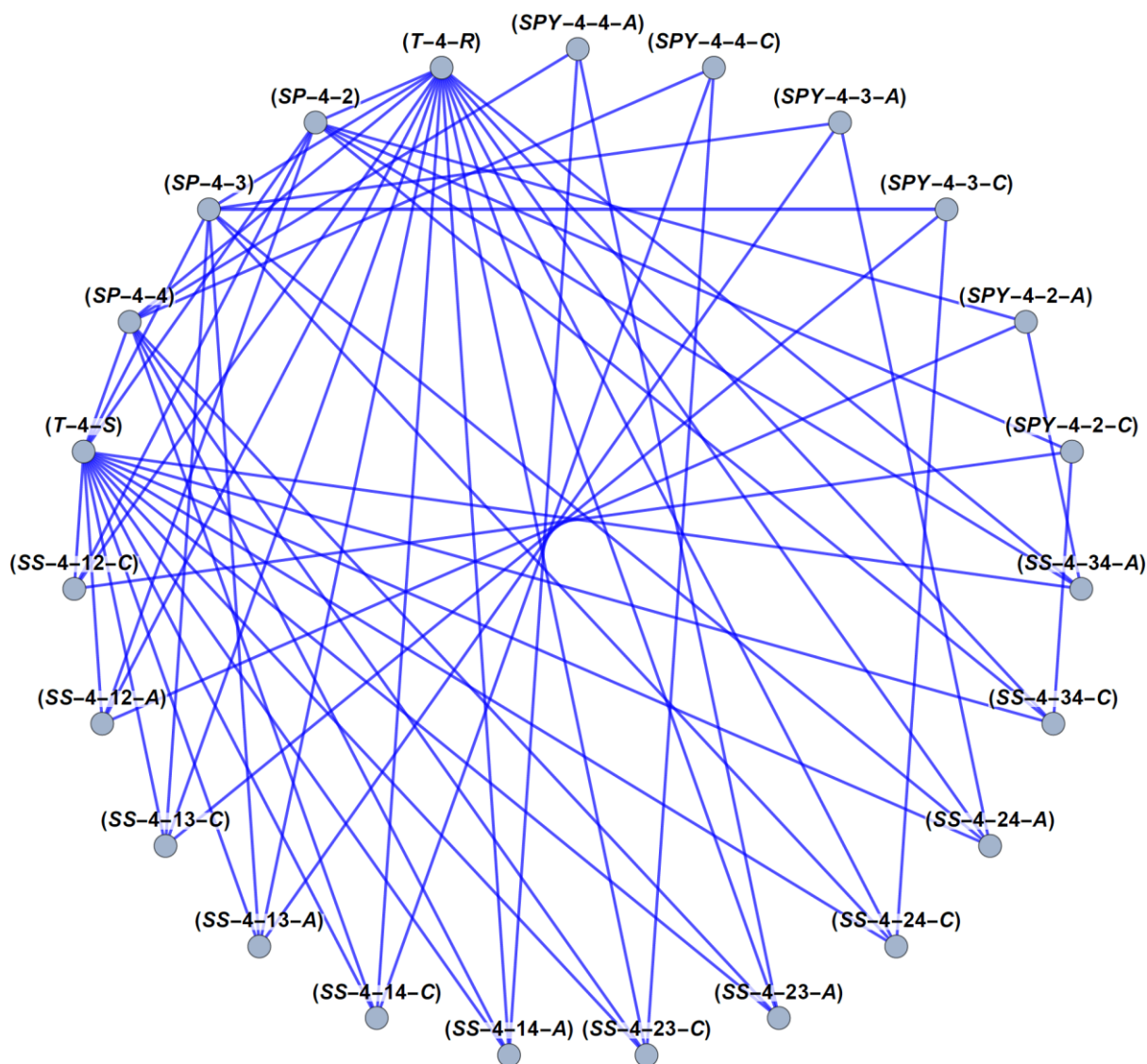
Figure 3.7 Graph showing all $R_{st}^c 1$ traverses for the full set of tetrahedral, square planar, "M"-apical square pyramidal, and see-saw polytope species of the ML_4 family. The labels are the standard IUPAC polyhedral configuration symbols.[3]

The path analysis and resulting graph for all species belonging to the considered set of polytope genera maps out the *full scope* of geometries (the distinct *spatial* arrangements and permutations of vertex indices for each), as well as and how these geometries transform (the rearrangements) into each other *via* $R_{st}^c 1$ processes. Put another way, within the limitations of the number of polytope genera in the set considered, this conceptual picture provides a *systematic* approach to listing the complete scope of possible stereoisomers and intermediates and provides non-dissociative concerted mechanisms for all interconversion reactions.

3.2.2 *"Closure" condition*

The Muetterties 1969 papers on the $ML_5{}^{38}$ and $ML_9{}^{26}$ families discuss the property of *closure* in reference to a set of polytopes. In this context, closure refers to *set closure* meaning that any $R_{st}^c 1$ operation on a member of a set *only* returns another member of that same set. In Graph Theory, closure has a different meaning.

Within the Muetterties implementation of the polytopal rearrangements model, typically only one geometric distortion operation transformation between two polytope genera was considered. Within this narrow setting, he defines[26] the closure conditions from the standpoint of stereoisomer permutations.

3.2.2.1 Applicability of the Muetterties model to real molecular systems

Within this discussion of idealised coordination geometries for a restricted set of polytopes, an important question arises: since the precise coordination geometries of real molecules often deviates from those arrangements described by the idealised polytopes, how appropriate is it to apply the polytopal-rearrangements model of stereoisomerisation?

Early work analysing numerous X-ray crystallographic structures of organic compounds and coordination complexes revealed that the deviations from ideal coordination geometries could be robustly interpreted as lying somewhere along a reaction coordinate between idealised forms. This principle of "Structural Correlation" by Bürgi and Dunitz[96-97] can be described as the geometric idiosyncrasies of real molecules appearing to fall along a reaction coordinate pathway on the potential energy hypersurface derived from idealised forms. This intimate relationship between static geometries and dynamical processes is a key concept which is discussed in more detail in Section 3.3.

Today, the appropriate quantum chemical methods can provide geometries, normal vibrational modes, and free energies for chemical species that faithfully match experimental data. Armed with such quantitative information, it is possible to reliably determine the kinetics for any $R_{st}^c 1$ process. For a complex web of interconnected polytope species, the tools of Graph Theory can be harnessed to formulate overall reaction rates involving numerous pathways. Further, if the graph representing a set of species and interconversion processes have energies and stationary point characteristics assigned to the graph vertices,

then the resulting mathematical object serves as a compact representation of the corresponding PES.

3.2.2.2 Relationships between geometries, vibrations and energy differences

Some general inferences can be made concerning the relationships between the magnitude of distortions occurring during an R_{st}^c1 process, the free energy change, and the frequency of the associated vibrational mode. In Figure 1.2 are shown orthogonal antisymmetric paired flexions for (a) the *T*-4 polytope of the ML$_4$ system and (b) the *TBPY*-5 polytope of ML$_5$. Despite these mechanisms being closely analogous, in the ML$_4$ case, the magnitude of the displacements interconverting the *T*-4 and *SP*-4 polytopes are relatively large as, typically, are the potential-energy changes with the energy of the corresponding vibration being high. In contrast, for the ML$_5$ case, the magnitude of the displacements interconverting the *TBPY*-5 and the *SPY*-5 polytopes are relatively small as are the potential-energy changes and corresponding vibration energy. This underpins why *T*-4 geometries are generally configurationally stable whilst *TBPY*-5 geometries often exhibit fluxionality.

(a)

| antisymmetric paired flexions of (12,34) | antisymmetric paired flexions of (12,34) |

(*T*-4-*S*)
Tetrahedral

(*SP*-4-12)
Square planar

(*T*-4-*R*)
Tetrahedral

(b)

(*TBPY*-5-12-*A*)
Trigonal bipyramidal

(*SPY*-5-12-*C*)
Trigonal bipyramidal

(*TBPY*-5-34-*C*)
Trigonal bipyramidal
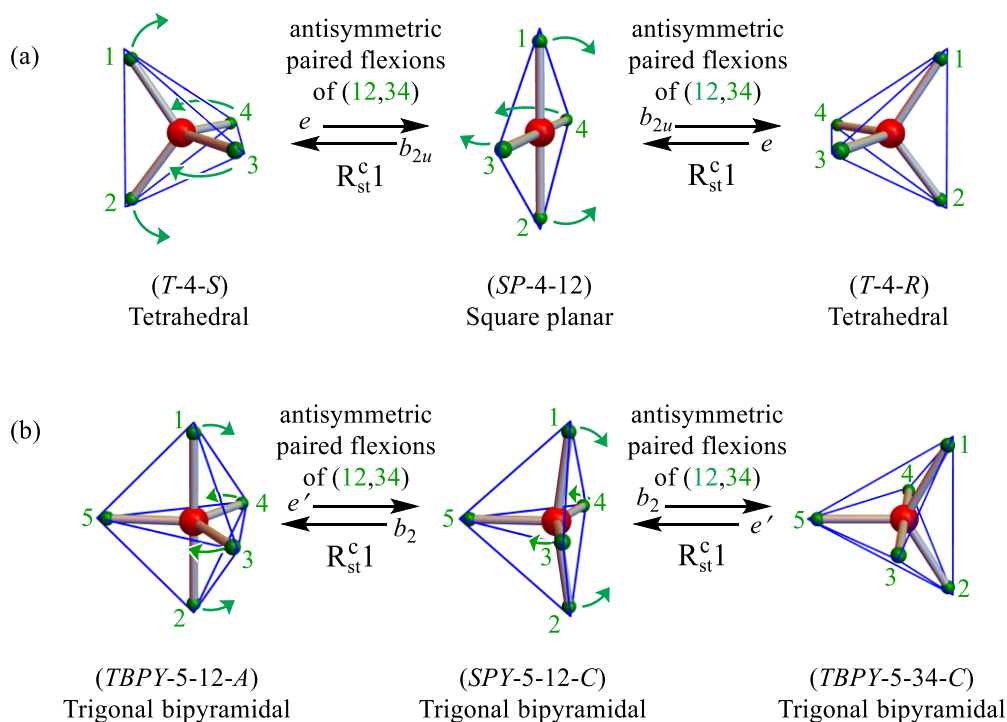
Figure 3.8 Direct comparison of antisymmetric paired orthogonal flexions of bond angles (12,34) for (a) the *T*-4 and *SP*-4 genera in the ML$_4$ polytope family, and (b) the *TBPY*-5 and *SPY*-5 genera in the ML$_5$ polytope family. In (a) for the ML$_4$ family, the distortions are comparatively large and are generally associated with large R_{st}^c1 reaction barriers. For the analogous process in the ML$_5$ family, the distortions are comparatively small as too are the corresponding R_{st}^c1 reaction barriers. The process in (a) is a unimolecular "inversion" of configuration of a tetrahedral centre and (b) is the Berry pseudorotation mechanism (see Section A.34).

3.2.2.3 <u>Problematic issues with the polytopal-rearrangements model</u>

Notwithstanding the success of the Muetterties model in rationalising the observed dynamical behaviour of stereocentres and apart from the lack of an *a priori* definition for what constitutes a polytope there exist several important shortcomings.

In all prior implementations of the polytopal-rearrangements model, the number of polytope genera considered was always few and typically two. Explicit reference to only considering the "low energy/frequency, large amplitude"[29, 38, 43] vibrational modes provided a practical and physically reasonable criterion for solving the types of problems of interest, such as fluxionality and shape distortion,[42-44, 46-48, 63, 72-73, 77, 80, 85-87, 94] to the inorganic chemistry community. Implementing this criterion, however, introduced a degree of arbitrariness with further development stalling. Whilst the underlying principle of following non-dissociative vibrational modes as reaction coordinates to generate different polytopes was set as the core of the model, the model itself was never fully and rigorously developed.

Whilst Muetterties explicitly recognised that geometric transformations between different polytope genera correlate with non-dissociative vibrational normal modes, his analysis being focussed on practical problems such as the Berry pseudorotation mechanism, only considered the *single direction* of a ground state vibrational motion that led to the polytope genus corresponding to the transition structure of interest. An unbiased implementation of this vibration-defined transformation principle would need to consider *both directions* of the vibrational motion displacement vector of interest. In the case where there are degenerate vibrations, this needs to be amended to *all directions* of these vibrations. Depending upon the ML_n family, symmetry properties of the polytope genus, and the symmetry of a vibration, different directions of the displacement vector could lead to distinct polytope genera.

A further obvious issue with the Muetterties model is that it is biased towards a small subset of non-dissociative vibrations. If the aim is to describe a fuller scope of stereoisomeric phenomena, then there is the implicit need to include all such vibrational modes.

As noted earlier, the polytopal-rearrangements model cannot describe isomerism arising from a torsion about a bond (torsional stereoisomerism, see Section A.48). As this property is the basis for all "classical conformational" stereoisomerism and stereoisomerisation processes, the vast majority of stereoisomerism that is important to

organic chemistry and biochemistry is absent from the model. Whilst it is true that torsional stereoisomerism can be described separately, the need for two distinct approaches is philosophically unsatisfying.

With much of this discussion as to "idealised forms", the question arises as to what "idealised" geometry actually means? For high symmetry geometries, such as tetrahedral or octahedral, the answer is obvious but for polytopes of lower symmetry the question becomes pressing. If the intention is to describe the relationships between a small number of well-defined polytope genera within a family, then some precise details may not be necessary. For example, when considering the relationships between the VSEPR-like polytope genera (*T*-4, *SP*-4, *SPY*-4, and *SS*-4), the precise angles and the apex of the *SPY*-4 genus or the angle between the "legs" of the *SS*-4 genus are not important. If, however, the intention is to locate some experimentally measured non-ideal static geometry along a pathway between two "ideal polytope genera" then precise and complete geometric definitions of those pathway-defining polytope genera are critical. In a rigorous treatment, any molecular structures should be unambiguously assigned to some polytope, without such issues arising.

A factor that may have contributed to the limited development of the polytopal-rearrangements model is that to do with scaling. As n increases for the ML_n families, the number of polytope genera is expected to rise *superexponentially*. Further, the permutations of vertex indices for each genus, *i.e.*, the number of polytope species for a genus, is also expected to scale superexponentially. Even with a precise definition for what constitutes a polytope, the scaling issue imposes limits to what is practical. At least for the common low coordination numbers, the problem could be expected to be tractable.

The IUPAC "Gold Book" definition for polytopal rearrangement (see Section A.33) also mentions rearrangements of cluster compounds such as polyhedral boranes. Mingos also discusses alkali metal systems,[88] but in these cases – best described as "clusters" – the rearrangements involve *changes in connectivity* and thus falls outside of the purview of this discussion limited to stereoisomerism where, by definition, the connectivity must be preserved.

3.3   Critical aspects needed for a complete description of stereoisomerism

In this Chapter, the aim is to develop a comprehensive descriptive model of stereoisomerism at a single stereocentre of arbitrary coordination number. Such a general approach, based upon a rigorous mathematical treatment can be reasonably expected to provide not just the description of stereoisomerism and mapping out allowed $R_{st}^{c}1$ processes, but also a *systematic* and *complete* classification framework. As outlined in Chapters 1 and 2, the current state of stereoisomerism is an ad hoc collection of treatments that have developed phenomenologically over the history of chemistry.

Addressing the deficiencies of the Muetterties model as detailed in Section 3.2.2.3, development of the Polytope Formalism of stereoisomerism requires the following:

(1) To formally accommodate bond torsions for the M–L bonds and hence describe torsional isomerism.

(2) To allow constraining ligand – ligand relationships to be defined.

(3) To formally include all non-dissociative vibrational modes.

(4) To follow all directions of the non-dissociative vibrational modes.

These are discussed below in detail.

3.3.1 *To formally accommodate bond torsions for the M–L bonds and hence describe torsional isomerism*

The simplest mathematical way to include the *property* of torsional relationships (either as geometric transformations or continuous motions) about a bond "M–L" for arbitrary ML_n, whilst still keeping the model minimalist and given the existing polytopal-rearrangements model, is to ascribe the property of "*orientability*" to each vertex point. Here, orientability is an *abstract mathematical property* and does not define a torsion in the traditional sense but defines the *capacity* for these traditional torsions. Orientability, in this context, can be thought of as a way of concisely representing the chemical *structure* of a ligand to "M" beyond the "L" atom. To represent this property graphically and to give an intuitive sense, the vertices of polytopes having the property of *orientability* are now represented as chequered spheres. This is shown in Figure 3.9 for the *A*-2 or "bent", 2-coordinate polytope genus.
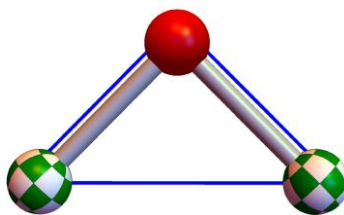
Figure 3.9 Depiction of the *A*-2 or "bent", 2-coordinate polytope genus with ligands having the property of *orientability*.

The property of *orientability* of a ligand can be thought of as a succinct representation of molecular connectivity *beyond* the ligating atom that allows for the traditional definition of a torsion without having to explicitly include these additional atoms. For example, in Figure 3.9, a ligand "atom" vertex may feature additional substituents that allow for a torsion to be defined; a^1–L^i–M–L^j, where no three or more consecutive coordinates are collinear.

If a quantum chemical model is used in the analysis, then in practice the orientable L "atoms" need to be modelled as simple, small, and sensible chemical groups that preserve the overall symmetry of the system. For example, in Figure 3.10a the ML_5 trigonal bipyramidal *TBPY*-5 genus is modelled as a substituted phosphorus species **3** comprised of axial, eclipsed methyl groups and trigonal-planar equatorial amino groups. The *TBPY*-5 genus having the same D_{3h} symmetry point group of **3** is maintained as either conformations **3a** or **3b**. The ML_2 linear *L*-2 polytope genus exhibits the $D_{\infty h}$ symmetry point group which cannot be *physically represented* without reducing the symmetry to a finite order. In Figure 3.10b, the model shown is the *eclipsed* conformation of linear dimethyl ether exhibiting the D_{3h} symmetry point group as an approximation to $D_{\infty h}$.
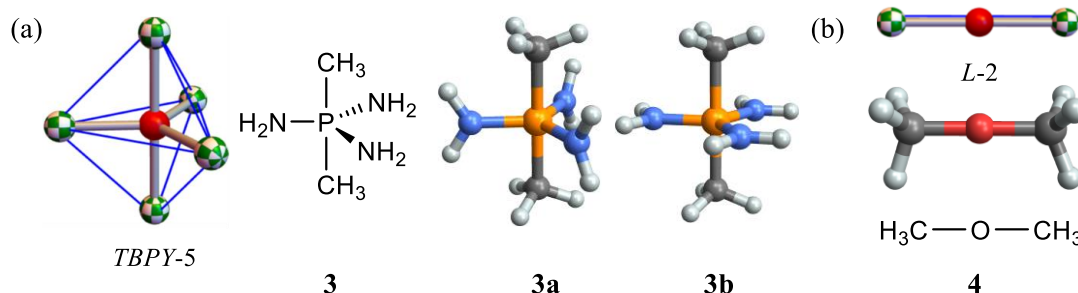
Figure 3.10 Practical implementations of polytopes featuring orientable ligands. (a) The *TBPY*-5 polytope genus with orientable ligands implemented as the phosphorus species indicated. As the *TBPY*-5 polytope genus exhibits the D_{3h} symmetry point group then the substituents need to preserve this whilst still being capable of describing M–L torsions. Here, two different orientations of the equatorial amino groups achieve this. (b) For the special case of the *L*-2 polytope genus, the $D_{\infty h}$ symmetry point group cannot be implemented as a chemical species with orientable ligands without reducing the symmetry to a finite order. The model shown here is the *eclipsed* conformation of linear dimethyl ether exhibiting the D_{3h} symmetry point group.

The polytopes shown in Table 3.2 and Table 3.3 have their L "atoms" depicted as being orientable.

3.3.2 To allow constraining ligand – ligand relationships to be defined

With the orientability property of the ligand "atoms" standing in for structure beyond the ligating atoms, this provides a convenient way of "coupling" a set of ligating "atoms" into a constrained group. If a set of *i* ligand "atoms" are coupled, this has the effect of making them into an *i*-dentate chelating group. From an internal coordinate perspective, for a set of coupled ligand "atoms", all the internal coordinates between the ligating atoms are preserved or only undergo small changes as the stereocentre undergoes non-dissociative rearrangements. Any rearrangement that would lead to a disruption of this coupling would be suppressed and hence not appear in the associated graph depicting the $R_{st}^c 1$ pathways.

If all *n* ligand "atoms" are coupled, then rotational invariance for the entire molecule is restored. Such a situation is referred to here as "embedding" since it places the stereocentre within a "fixed" external frame of reference (see Section 2.4.9). This is an important concept that is discussed further in several contexts throughout the remainder of this Chapter.

3.3.3 Inclusion of all non-dissociative vibrations and vibrational displacement vector directions

Nominally, vibrations can be partitioned into those that only change bond angles, torsional angles, *etc*., and those that lead to bond dissociation. Those that lead to dissociation describe constitution isomerism or other processes and are hence outside the scope of stereoisomerism considered in this Chapter.

Requirements (3) and (4) describe the formal inclusion of all non-dissociative vibrational modes and to follow all directions of these. In doing so, all possible $R_{st}^c 1$ processes are defined from a physics-based approach. It is posited that for an ML_n family where no ligand – ligand constraints are imposed, all possible polytopes can form a closed set under the operations described by the $R_{st}^c 1$ processes.

A subtle but significant challenge arises where there are *degenerate vibrational normal modes of the same symmetry*. In such cases, all linear combinations of these

vibrations represent feasible processes, with an infinite set of solutions possible. This situation is a common feature for the highest symmetry polytope of a given family. Consequently, they are present in commonly encountered geometries like *TP*-3 and *T*-4. How this situation is dealt with becomes an important principle of the resulting Polytope Formalism of stereoisomerism.

3.4 Development of the Polytope Formalism

The following sections describe how these additional features are incorporated into the Muetterties model to produce the Polytope Formalism of stereoisomerism.

3.4.1 Accounting for all possible $R_{St}^{c}1$ processes within a selected subset of polytopes

Returning to the ML_4 polytope family, the worked examples given in Figure 3.5a and Figure 3.8a present two specific polytopes, (*T*-4-*S*) and (*SP*-4-12), each connected to the other by a vibration corresponding to orthogonal antisymmetric paired flexional motions of (12,34); this situation is replicated in Figure 3.11a. From the (*SP*-4-12) structure, this vibration of generic b_{2u} symmetry can be applied in two directions – designated here as $b_{2u}(+)$ and $b_{2u}(-)$, with each direction leading to a different *T*-4 configuration. Considering $R_{st}^{c}1$ reactions from *T*-4, the appropriate vibration, again corresponding to orthogonal antisymmetric paired flexional motions of (12,34), applied in one direction leads back to (*SP*-4-12), but when applied in the other direction leads to a polytope that is not included in the set being considered. In the Muetterties model, this motion is simply ignored. The consequence of this is that not all possible reactions are included on the associated reaction graphs and these graphs are not closed under all directions of vibrations defining the polytope transformations.

A way to overcome this incomplete accounting for vibrational displacements is to introduce the concept of the *semi-terminal* property (see Section 2.4.23). A polytope that is semi-terminal has at least one $R_{st}^{c}1$ reaction leading to a polytope that is excluded from the set under consideration, with this property indicated as a *self-loop* on a *directed graph*. The direction of the vibrational displacements that lead out of the set are returned and thus restore set closure. This notion is illustrated in Figure 3.11b, making explicit the nature of the chemical model being represented by the directed graph. The set of polytopes and vibrational

modes included in an analysis, along with other explicit details of the model, define the *demesne* (see Section 2.4.8) of the analysis.

With a more general approach, additional polytopes made by performing reactions in the *reverse* direction could be included within the set under consideration. This is illustrated in Figure 3.11c, with the new polytopes required representing a genus called a "tetragonally-elongated tetrahedron" and assigned absolute configurations (*TET*-4-2-*S*) and (*TET*-4-2-*R*). These configuration symbols are constructed following the form (*TET*-4-*i*-*X*), where *i* is the vertex index nearest to index 1 and *X* being *R* or *S* as for the *T*-4 genus. These polytopes also have critical vibrations, again corresponding to orthogonal antisymmetric paired flexional motions of (12,34), that convert back to their respective *T*-4 species. Hence the original problem of only considering motion in one direction is again encountered for the newly introduced polytope. Consequently, these *TET*-4 polytopes can now be considered as semi-terminal with respect to this now-enlarged set of polytopes under consideration. Importantly, the two *T*-4 species are no longer semi-terminal with respect to this enlarged set of polytopes and transformations.

Repeating the generalisation process by considering these vibrations (of generic a_1 symmetry) in the other direction leads to two L "atoms" moving on top of each other. This is a constitutional-isomerism process. In a general approach embodying both stereoisomerism and constitutional isomerism, the nature of the products would become important. Herein, only stereoisomerism is being considered and so such processes are not relevant to the analysis. The semi-terminal path shown in Figure 3.11c thus represents a "hard edge" to the demesne of stereoisomerism. Additionally, the symmetry point group of the *TET*-4 species shown does not change for this direction of the a_1 generic symmetric vibrational motions, merely increasing the degree of tetragonal elongation. The general stereoisomeric relationships between the vertices remain the same.
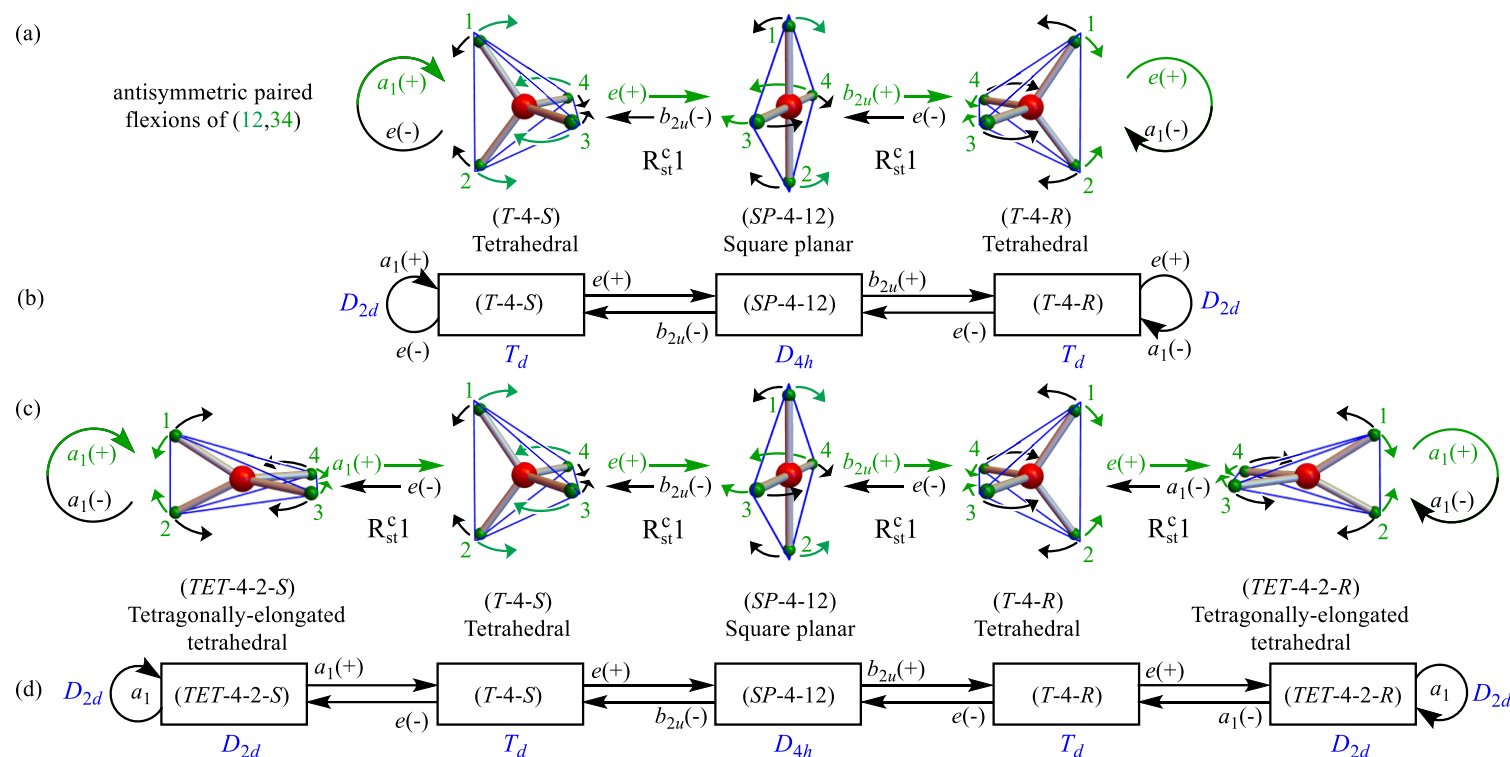
Figure 3.11 Effect of direction of the vibrational displacement vector on polytope rearrangements as demonstrated by an antisymmetric orthogonal pair of flexions for the ML$_4$ family. (a) Here only three polytopes are considered: the (*SP*-4-12) species (centre) and the two *T*-4 species (*T*-4-*S*) and (*T*-4-*R*). Each direction of the vibration of generic b_{2u} symmetry corresponding to antisymmetric paired flexional motions of (12,34), applied to the (*SP*-4-12) species leads to each of the *T*-4 species shown. For each of the *T*-4 species, analogous antisymmetric paired flexional motions have generic *e* symmetry. One direction of these vibrations returns both *T*-4 species to the (*SP*-4-12) species. The other direction of these vibrations leads to some other polytope of reduced symmetry not explicitly included in the set of polytopes considered here. Under these conditions, this property of the *T*-4 polytope is called *semi-terminal*. (b) The associated directed graph for (a) with generic symmetry point groups indicated. The semi-terminal properties of the *T*-4 species are indicated by self-loops. The purpose of self-loops is to account for all directions of motion for a set of considered vibrations. (c) The set of polytopes used in (a) can be expanded to now include those generated by these motions. Here, the new generic polytope is called a "tetragonally-elongated tetrahedron" exhibiting the D_{2d} symmetry point group with absolute configurations assigned as described in the main text. For the *TET*-4 species shown, analogous antisymmetric paired flexional motions have generic a_1 symmetry. One direction of motion for these a_1 vibrations returns the structures to the *T*-4 species whilst the other direction is again semi-terminal. (d) The directed graph for (c). In contrast to the situation in (a), in (c) the semi-terminal property of the *TET*-4 polytopes does not result from a change to a different polytope. The vibrational motions responsible for the semi-terminal property of *TET*-4 polytopes simply change the degree of tetragonal elongation, not the general stereoisomeric relationships of the vertices. Green vibrations and reaction arrows are in the forward direction, those in black are in the reverse.

The examples shown in Figure 3.11 focus on one general type of vibrational motion, albeit one that changes form slightly through a progression of different polytopes. Stereocentres embody many vibrational motions and expanded polytope sets can be envisaged in which an increasing number of vibrational motions, *i.e.,* $R_{st}^c 1$ mechanisms, are included. On the expanded graphs depicting this additional chemistry, the representation of excluded processes *via* depiction of semi-terminal polytopes will always be possible. For example, the graph in Figure 3.7 depicting rearrangements within the *T*-4, *SP*-4, *SS*-4, and *SPY*-4 genera can be modified to indicate which members are semi-terminal polytopes with respect to this set and hence account for *all possible* reaction mechanisms in a fully consistent manner. It would seem apparent that generic vibrational symmetry plays a key role in this development. Only motions that conform to one type of asymmetric vibration can contribute to an $R_{st}^c 1$ mechanisms.

3.4.2 *Determination of sets of polytopes closed under all possible associated $R_{St}^c 1$ reactions embodying all possible vibrational modes*

Generally applicable sets of polytopes can be obtained by the vibrational-mode-displacement-following method of Muetterties but *applied exhaustively* to an arbitrary ML_n stereocentre. To do this there are four main challenges:
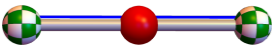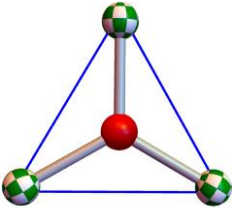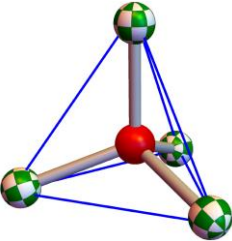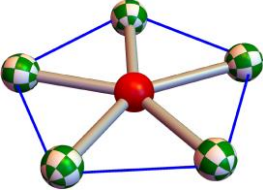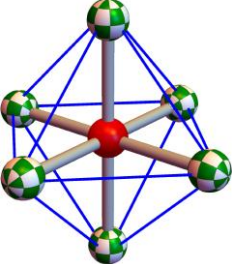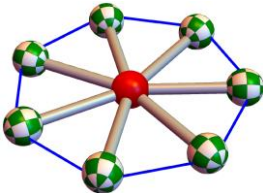
     (1) where to start,

     (2) how to deal with degenerate vibrations,

     (3) how to deal with polytopes of C_1 symmetry, and

     (4) polytopes with linear L–M–L′ angles
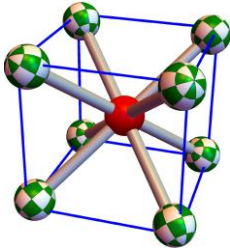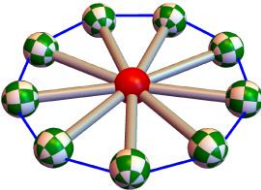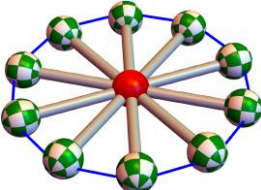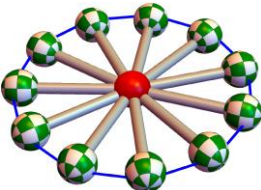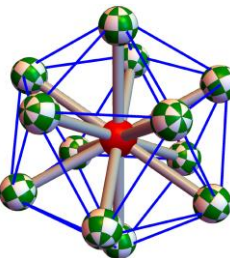
3.4.2.1 <u>Where to start</u>

Addressing point (1) is straight forward. Whilst, in principle, any polytope might serve as a starting point due to the requirement that a full set of unconstrained polytopes and associated $R_{st}^c 1$ processes exhibits set closure, in practice the *highest symmetry* polytope serves as the *seed polytope* (see Section 2.4.22) for "growing" the full set by the vibrational-mode-displacement-following method as described. For a polytope family ML_n, the highest symmetry polytope is unambiguously and uniquely defined. The highest-symmetry polytopes for up to ML_{12} are listed in Table 3.2 and provide an initial description for coordination numbers up to 12. Chemical species with higher coordination numbers are

unlikely due to the relative size of atoms and the steric constraints higher coordination numbers impose.

Table 3.2 Seed polytope genera for polytope families ML_n with n up to 12.
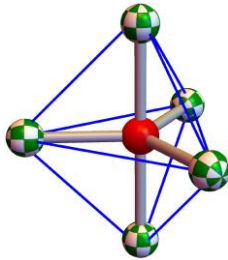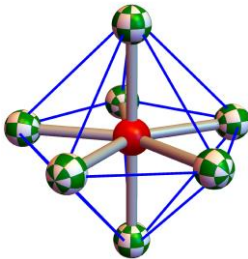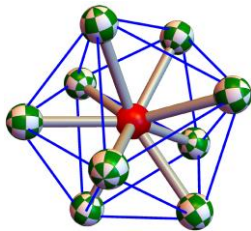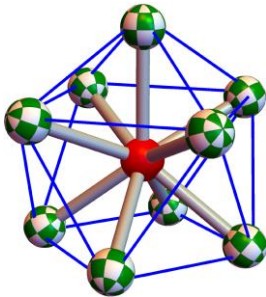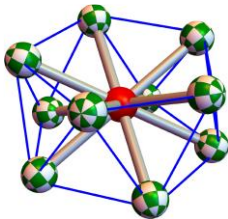
n	Polytope name	IUPAC polyhedral symbol [a]	Point group	Number symmetry operations	Image	Normalised vertex coordinates in $\mathbb{R}^3$ (See Appendix C)
0	Point	none	O(3)	∞		*null*
1	Linear interval	none	$C_{\infty v}$	∞		(1, 0, 0)
2	Linear	L-2	$D_{\infty h}$	∞		(1, 0, 0) (−1, 0, 0)
3	Trigonal plane	TP-3	D_{3h}	12		(1, 0, 0) (−0.5, −0.866, 0) (−0.5, 0.866, 0)
4	Tetrahedron	T-4	T_d	24		(0, 0, 1) (−0.471, −0.817, −0.333) (−0.471, 0.817, −0.333) (0.943, 0, −0.333)
5	Pentagonal plane	PP-5	D_{5h}	20		$(Cos\frac{2k\pi}{5}, Sin\frac{2k\pi}{5}, 0)$ with $k \in [1,2 \dots 5]$
6	Octahedron	OCT-6	O_h	48		(−1, 0, 0) (0 ,1, 0) (0, 0, −1) (0, 0, 1) (0, −1, 0) (1, 0, 0)
7	Heptagonal plane	HP-7	D_{7h}	28		$(Cos\frac{2k\pi}{7}, Sin\frac{2k\pi}{7}, 0)$ with $k \in [1,2 \dots 7]$
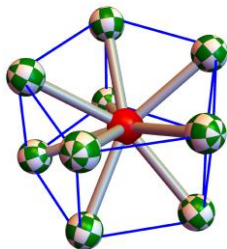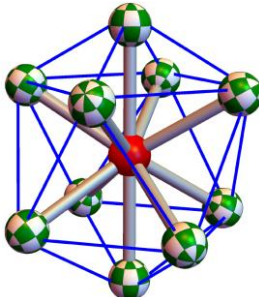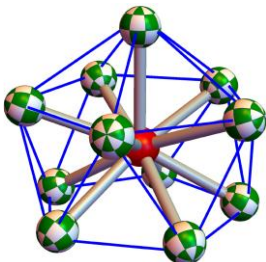
n	Polytope name	IUPAC polyhedral symbol [a]	Point group	Number symmetry operations	Image	Normalised vertex coordinates in $\mathbb{R}^3$ (See Appendix C)
8	Cube	*CU*-8	O_h	48		$(-0.577, -0.577, -0.577)$ $(-0.577, 0.577, 0.577)$ $(-0.577, 0.577, -0.577)$ $(-0.577, 0.577, 0.577)$ $(0.577, -0.577, -0.577)$ $(0.577, -0.577, 0.577)$ $(0.577, 0.577, -0.577)$ $(0.577, 0.577, 0.577)$
9	Nonagonal plane	*NP*-9	D_{9h}	36		$(Cos\frac{2k\pi}{9}, Sin\frac{2k\pi}{9}, 0)$ with $k \in [1,2 \ldots 9]$
10	Decagonal plane	*DP*-10	D_{10h}	40		$(Cos\frac{2k\pi}{10}, Sin\frac{2k\pi}{10}, 0)$ with $k \in [1,2 \ldots 10]$
11	Undecagonal plane	*UP*-11	D_{11h}	44		$(Cos\frac{2k\pi}{11}, Sin\frac{2k\pi}{11}, 0)$ with $k \in [1,2 \ldots 11]$
12	Icosahedron	*ICO*-12	I_h	120		$(0, 0, -1)$ $(0, 0, 1)$ $(-0.894, 0, -0.447)$ $(0.894, 0, 0.447)$ $(0.724, -0.526, -0.447)$ $(0.724, 0.526, -0.447)$ $(-0.724, -0.526, 0.447)$ $(-0.724, 0.526, 0.447)$ $0.276, -0.851, -0.447)$ $(-0.276, 0.851, -0.447)$ $(0.276, -0.851, 0.447)$ $(0.276, 0.851, 0.447)$

a: Symbols in blue not specified by IUPAC but introduced here and constructed in an analogous manner to the recognised IUPAC polyhedral symbols.

For ML$_n$ with $n > 3$, the Platonic solids are the only non-planar highest-symmetry polytopes. Hence these polytopes have often dominated applications of the polytopal-rearrangements model. Polytopes that better resemble expected molecular structures and may prove useful alternative seed polytopes, especially for higher coordination-numbers, are the VSEPR-related ones given in Table 3.3.

Table 3.3 Polytopes useful as alternative starting points.

n	Polytope name	IUPAC polyhedral symbol [a]	Point group	Number symmetry operations	Image	Normalised vertex coordinates in $\mathbb{R}^3$ (See Appendix C)
5	Trigonal bipyramid (Johnson 12)	*TBPY*-5	D_{3h}	12		(0, 0, −1) (0, 0, 1) (−0.5, −0.866, 0) (−0.5, 0.866, 0) (1, 0, 0)
7	Pentagonal bipyramid (Johnson 13)	*PBPY*-7	D_{5h}	20		(0, 0, −1) (0, 0, 1) (1, 0, 0) (0.309, −0.951, 0) (0.309, 0.951, 0) (−0.809, −0.588, 0) (−0.809, 0.588, 0)
9	Trigonal prism, square face tricapped (Johnson 51)	*TPRS*-9	D_{3h}	12		(−0.378, −0.655, −0.655) (−0.378, −0.655, 0.655) (−0.378, 0.655, −0.655) (−0.378, 0.655, 0.655) (0.756, 0, −0.655) (0.756, 0, 0.655) (−1, 0, 0) (0.5, −0.866, 0) (0.5, 0.866, 0)
9	Square antiprism, square face capped (Johnson 10)	*SAPRS*-9	C_{4v}	12		(−0.563, −0.563, −0.606) (−0.563, 0.563, −0.606) (0, 0, 1) (0, −0.919, 0.393) (0, 0.919, 0.393) (0.563, −0.563, −0.606) (0.563, 0.563, −0.606) (−0.919, 0, 0.393) (0.919, 0, 0.393)
10	Pentagonal antiprism	*PAPR*-10	D_{5d}	20		(−0.894, 0, −0.447) (0.894, 0, 0.447) (−0.724, −0.526, 0.447) (−0.724, 0.526, 0.447) (−0.276, −0.851, −0.447) (−0.276, 0.851, −0.447) (0.276, −0.851, 0.447) (0.276, 0.851, 0.447) (0.724, −0.526, -0.447) (0.724, 0.526, −0.447)

n	Polytope name	IUPAC polyhedral symbol [a]	Point group	Number symmetry operations	Image	Normalised vertex coordinates in $\mathbb{R}^3$ (See Appendix C)
10	Pentagonal prism	*PPR*-10	D_{5h}	20		(0.862, 0, −0.507) (0.862, 0, 0.507) (−0.697, −0.507, −0.507) (−0.697, −0.507, 0.507) (−0.697, 0.507, −0.507) (−0.697, 0.507, 0.507) (0.266, −0.820, −0.507) (0.266, −0.820, 0.507) (0.266, 0.820, −0.507) (0.266, 0.820, 0.507)
10	Square antiprism, square face bicapped (Johnson 17)	*SAPRS*-10	D_{4h}	16		(−0.608, −0.608, −0.511) (−0.608, 0.608, −0.511) (0, 0, −1) (0, 0, 1) (0, −0.860, 0.511) (0, 0.860, 0.511) (0.608, −0.608, −0.511) (0.608, 0.608, −0.511) (−0.86, 0, 0.511) (0.86, 0, 0.511)
11	Pentagonal antiprism, pentagon face capped (Johnson 11)	*PAPRP*-11	C_{5v}	10		(0, 0, 1) (−0.856, 0, −0.518) (0.93, 0, 0.367) (−0.264, −0.814, −0.518) (−0.264, 0.814, −0.518) (0.287, −0.885, 0.367) (0.287, 0.885, 0.367) (−0.752, −0.547, 0.367) (−0.752, 0.547, 0.367) (0.692, −0.503, −0.518) (0.692, 0.503, −0.518)

a: Symbols in blue not specified by IUPAC but introduced here and constructed in an analogous manner to the recognised IUPAC polyhedral symbols.

3.4.2.2 How to deal with degenerate vibrations of the same symmetry

As briefly mentioned in Section 3.3.3, degenerate vibrational normal modes pose a challenge for defining vibration displacement vectors as all linear combinations of such vibrations represent feasible processes, with an infinite set of solutions, possible. This situation will always arise for the seed polytopes listed in Table 3.2. Indeed, this is so in the worked example of the orthogonal antisymmetric paired flexional motions for *T*-4 described in Figure 3.4, Figure 3.5, and Figure 3.11, where the corresponding vibration has *e* symmetry and so doubly degenerate.

Inspiration for dealing with this problem comes from the known properties of the Jahn-Teller effect.[98] This specifies how structures supporting degenerate vibrations can

distort to reduce their symmetry. Interactions that lower symmetry are usually described in a hierarchy as first-order, second-order, *etc*., with each term in the series generating increasingly more subtle geometrical changes. Each order therefore generates sets of structures that need to be categorised in terms of polytopes. Throughout chemistry, mostly low-order interactions suffice to describe structures produced by the Jahn-Teller effect, demanding the availability of only a small set of polytopes. As all vibrational motions are to be included at a *consistent level*, the set of polytopes produced in this way is expected to be complete under the $R_{st}^c 1$ transformations supported by the specified order of the Jahn-Teller interaction. In practice, this means there is no single unique set of polytopes can fully characterise an ML_n family, but rather an infinite number of finite sets of increasing size arise, with each set being closed under its allowed transformations.

As chemical variations tend to be small, making useful the original Muetterties model that focused on only a few well-known structures of an ML_n family, it is likely that the smallest viable polytope set will be the most useful one. A property of the smallest set is that it includes polytopes representative of *every* symmetry point group that any stereocentre represented by the ML_n family could adopt. The smallest viable full-polytope set is called a *concise set* (see Section 2.4.7). The inclusion of low-symmetry structures is a significant departure from the Muetterties model, which assumed that all stable isomers have high symmetry or, where nearly so, were idealised to high-symmetry forms.

Determining the concise set, and indeed the procedures needed to create larger closed sets, would appear to be most naturally undertaken using a coordinate system that reflects the structure of the high-symmetry seed polytope. Examples are subsequently provided as to how this can be done for both ML_2 and ML_3.

In the concise set, all polytopes interconnected by $R_{st}^c 1$ operations have different point-group symmetry. Hence their point group, plus other known features of the available polytopes, facilitates the direct determination of which polytope any real stereocentre structure corresponds to. Other features of relevance concern internal relationships that differentiate between different types of structure with the same symmetry, plus external relationships associated with the embedding of the stereogenic unit in its environment. When this applies, there is no need for the Bürgi and Dunitz[97] correlation conjecture to be applied.

For example, Figure 3.5 depicts tetrahedral and square-planar polytopes of ML_4 in a polytope set interconnected by a single R_{st}^c1 process. The Bürgi and Dunitz principle allows any real structure to be expressed in terms of its similarity to these forms. In a concise polytope set, another polytope will exist depicting the lower symmetry of all intermediary structures. Real structures are then automatically assigned to either an original polytope or to this newly generated intermediate, depending on their symmetry. In the concise polytope set, R_{st}^c1 mechanisms connect the *T*-4 polytope to this new one, and this new one to the *SP*-4 polytope only. All polytopes do not need to correspond to stationary points on the molecular PES, and so it is possible, and indeed likely, that the traditional *T*-4 polytope and *SP*-4 polytope determine all isomers and transition states of interest in some situation. Nevertheless, the concise polytope set recognises the possibility that low-symmetry structures formed by high-order Jahn-Teller interactions, however uncommon, may represent the stable isomers instead.

In larger non-concise complete sets, polytopes with the same symmetry point group will be connected by R_{st}^c1 mechanisms. In this case, symmetry, polytope properties, and embedding will be insufficient to uniquely assign any real structure to one of the available polytopes. In this case, generalised forms of the Bürgi and Dunitz principle need to be developed to enable *a priori* structure assignment.

3.4.2.3 <u>How to deal with polytopes of C_1 symmetry</u>

If a polytope has high symmetry, then its vibrations may be categorised according to their symmetry. In many cases, this will be sufficient to identify all R_{st}^c1 mechanisms. When more than one vibration share the same symmetry, any linear combination of those vibrations could generate the anticipated R_{st}^c1 paths. Finding characteristic paths therefore becomes a challenge. For polytopes of C_1 symmetry, all vibrations need to be considered and appropriate linear combinations sought, which could present practical issues. An approach to solving this problem arises by recognising that R_{st}^c1 paths are bidirectional as demanded by the principle of micro-reversibility. Determination of the R_{st}^c1 path from a high-symmetry polytope to a lower-symmetry one immediately provides the return path. The constraint can then be introduced that all other R_{st}^c1 paths away from the low-symmetry structure must start orthogonal to this established path.

3.4.2.4 <u>Polytopes with linear L–M–L′ angles</u>

Many ML_n families feature high-symmetry seed polytopes with linear L–M–L relationships. This will become a critical feature of the associated compact polytope set. Also, polytopes with linear angles will pose significant issues for the embedding of the stereocentre inside an external frame of reference, as torsional angles embodying these atoms will be undefined owing to the associated singularity. Hence care will always need to be taken when describing the stereochemistry of these families.

Owing to the propensity of linear structures envisaged by VSEPR models, linear structures have been commonly envisaged as polytopes, even when their presence is not demanded by symmetry. A common example is the *SS*-4 see-saw polytope of ML_4, see Figure 3.6; another example is the "T-shaped" polytope of ML_3 that featured in Chapter 1, Figure 1.3. In these cases, there is no demand for a precisely linear L–M–L arrangement with a 180° angle representing just one example of possible structures of the *same symmetry*. As a structure of this type featuring a precisely linear L–M–L arrangement represents a single point along a continuum of forms in configuration space, the probability of a real chemical system having this precise geometry is zero. Even though the underlying singularity in the coordinate system of embedded stereocentres must always be considered, its effect will not be as impactful as is that for polytopes, like *L*-2, in which symmetry demands a linear angle.

3.5   The ML_2 family

Despite the systematic nature of the Muetterties model, it was never applied to the seemingly simple systems with low-coordination number, perhaps under the assumption that all possible stereochemistry was known and thus there was nothing new or interesting to discover. The vibrational-mode-displacement-following method for determining all possible polytopes for a family can now be applied to ML_2, presenting a unified picture. Even though ML_2 would appear to be simple and hence well defined, options are found for what is the best way to represent the ML_2 family of polytopes in different circumstances. Further, the low dimensionality of all ML_2 polytopes relative to $\mathbb{R}^3$ space provides significant challenges and opportunities for its embedding inside a larger chemical system.

3.5.1 Polytopes representing ML$_2$

The highest-symmetry polytope for ML$_2$ is the linear polytope *L*-2, which exhibits the *D*$_{\infty h}$ generic symmetry point group. The bending vibration is doubly degenerate and hence can be applied in an infinite number of ways, with *D*$_{\infty h}$ structures presenting singularities in coordinate-system definitions. Molecules of the form ML$_2$ have four vibrational motions and two rotational motions at *D*$_{\infty h}$ configurations, but three vibrations and three rotations otherwise. This feature stresses the critical nature of the singularity to the understanding of the properties of ML$_2$.

By whatever means, bending the bond angle L–M–L′ leads to structures with the *C*$_{2v}$ generic symmetry point group. Such bent structures have been historically associated with the *A*-2 polyhedral symbol, regardless of the value of the angle. This reflects the basis for these polyhedral symbols to name "stable" and characterisable structures, and not as they are used here to describe fine subtleties of structural relationships. The *L*-2 and *A*-2 polytopes form a concise set with its directed graph shown in Figure 3.12a. Any real ML$_2$ structure can be assigned to one of these concise polytopes based on symmetry information alone. In Figure 3.12b, an *A*-2 polytope is shown reflexing via an *L*-2 polytope to another *A*-2 polytope. The associated directed graph is shown in Figure 3.12c. There are an infinite number of ways in which the degenerate bending vibrations can combine when applied to *L*-2, making for an infinite number of possible $R_{st}^c 1$ mechanisms that can only be differentiated if the stereocentre is embedded into an external frame of reference. Regardless, all these mechanisms share the same basic feature in that bending takes *L*-2 to *A*-2 and on to constitutional isomerism at sufficiently small angles when the two ligands chemically interact.

Beyond this concise set, an expanded polytope set – also closed under all possible $R_{st}^c 1$ mechanisms – can be obtained by trifurcating *A*-2 into three structures all with the same symmetry featuring, *e.g.*, an obtuse angle, a right angle, and an acute angle. For labelling purposes, these are assigned the new polyhedral symbols *OA*-2, *RA*-2, and *AA*-2, respectively, and are shown in Figure 3.12d along with the *L*-2 polytope; their interconnecting $R_{st}^c 1$ mechanisms form the directed graph shown in Figure 3.12f. If, for some chemical system, stable stereoisomers corresponding to *OA*-2 and *AA*-2 could be made that could be interconverted *via* some transition state of intermediate geometry like *RA*-2, then it would be

essential to use this non-compact polytope set for the naming and storage of the associated structures.
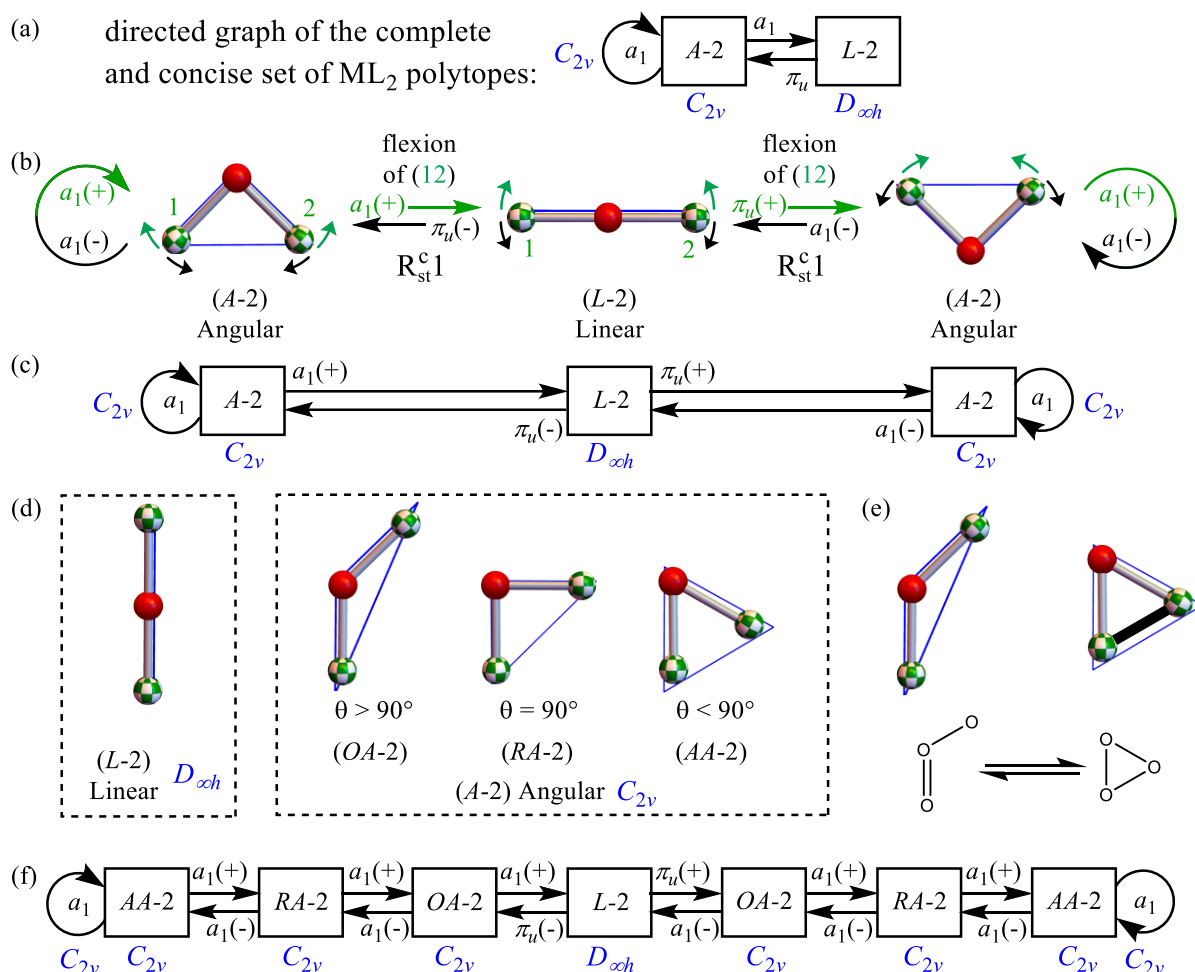
Figure 3.12 Polytopes of the ML$_2$ family. (a) The directed graph of the complete and concise set of ML$_2$ polytopes. (b) The vibrational mode – polytope traverse analysis for the ML$_2$ family utilising the simpler and concise *A*-2 form to demonstrate the relationships between a vibrational mode and the resulting structural changes. Both displacement vector directions of a single vibration from the orthogonal and degenerate pair or normal modes having generic π_u symmetry corresponding to the bond angle flexion of the *L*-2 linear form (centre) changes the polytope to an *A*-2 angular form. For an *A*-2 polytope, one direction of the displacement vector of the vibration of a_1 symmetry returns the structure to the *L*-2 polytope. The other direction of this same vibration does not lower the symmetry further, hence *A*-2 is a semi-terminal polytope with the self-loops indicated and shown in the associated directed graph (c). (d) The unique *L*-2 linear polytope and three angular polytopes produced through angle permutations described in the main text and as indicated here as obtuse angular *OA*-2, right angular *RA*-2, and acute angular *AA*-2. The three angular polytopes all exhibit the C_{2v} symmetry point group and can be described in a general sense as *A*-2 without reference to the angle categorisation. (e) In real molecular systems, when the bond angle is of the order of 60° or less bond topology changes can occur as shown here for the open form of ozone changing to the cyclo-form. Such a constitutional isomerism change at the ligand atoms has the effect of introducing a new external constraint to the ML$_2$ stereocentre that was absent in the other angular polytopes. (f) The directed graph showing the relationships between the four polytopes in (d) for a flexion of the bond angle. Due to the low dimensionality of the ML$_2$ system, in $\mathbb{R}^3$ the orthogonal and degenerate pair of π_u symmetry vibrations for *L*-2 do not have an absolute orientation. And each *A-2* species is equivalent.

Nevertheless, it is doubtful that any chemical system could exhibit this stereoisomeric behaviour. As a bond angle decreases in a bent bi-coordinate system, the energy will typically rise until increasing L – L interactions lead to the formation of an L–L bond representing a constitutional change. This is highlighted in Figure 3.12e, taking ozone as an example. Clearly, it has an isomer that could be associated with *OA*-2, but the isomer that could be associated with *AA*-2 is more appropriately treated as a conformational isomer. Hence just the concise 2-element set {*L*-2, *A*-2} is sufficient to describe stereoisomerism in ozone.

3.5.2   *ML₂ embedding*

For the ML_2 family, the unique *L*-2 polytope is a one dimensional object and all angular polytopes are two dimensional. Consequently, any ML_2 polytope existing in $\mathbb{R}^3$ space cannot be oriented in an absolute sense without some reference to an additional system. The situation is particularly problematic for the *L*-2 case. Here, the only non-dissociative vibrational modes are a pair of degenerate and orthogonal normal modes of generic π_u symmetry. Without some external frame of reference there are an infinite set of solutions for the direction of these, consistent with the $D_{\infty h}$ generic symmetry point group. A general consequence of the low dimensionality of ML_2 polytopes is that IUPAC nomenclatural polyhedral *absolute configuration* symbols have not been developed.

On first inspection, it may seem that ML_2 has minimal scope of possibilities to offer stereochemistry and, indeed, this is true of simple open-chain molecules featuring an ML_2 stereocentre such as H_2O. The key to realising an increased scope of stereoisomeric possibilities derives from providing the ML_2 unit with an external frame of reference – achieved by coupling the orientable ligand atoms to a fixed (or approximately fixed) molecular entity. This is called "embedding" (see Section 2.4.9). For discrete molecular systems, this is most easily achieved by bonding the ML_2 unit into a cyclic structure with sufficient rigidity to isolate the now-distinct ML_2 orientations. For materials systems, an example would be to bridge two parallel slabs of material with a bicoordinate atom. In either case, motions of the ML_2 centre can now be described relative to the embedding. This principle is depicted in Figure 3.13, with a full description of available geometries and allowed motions given.
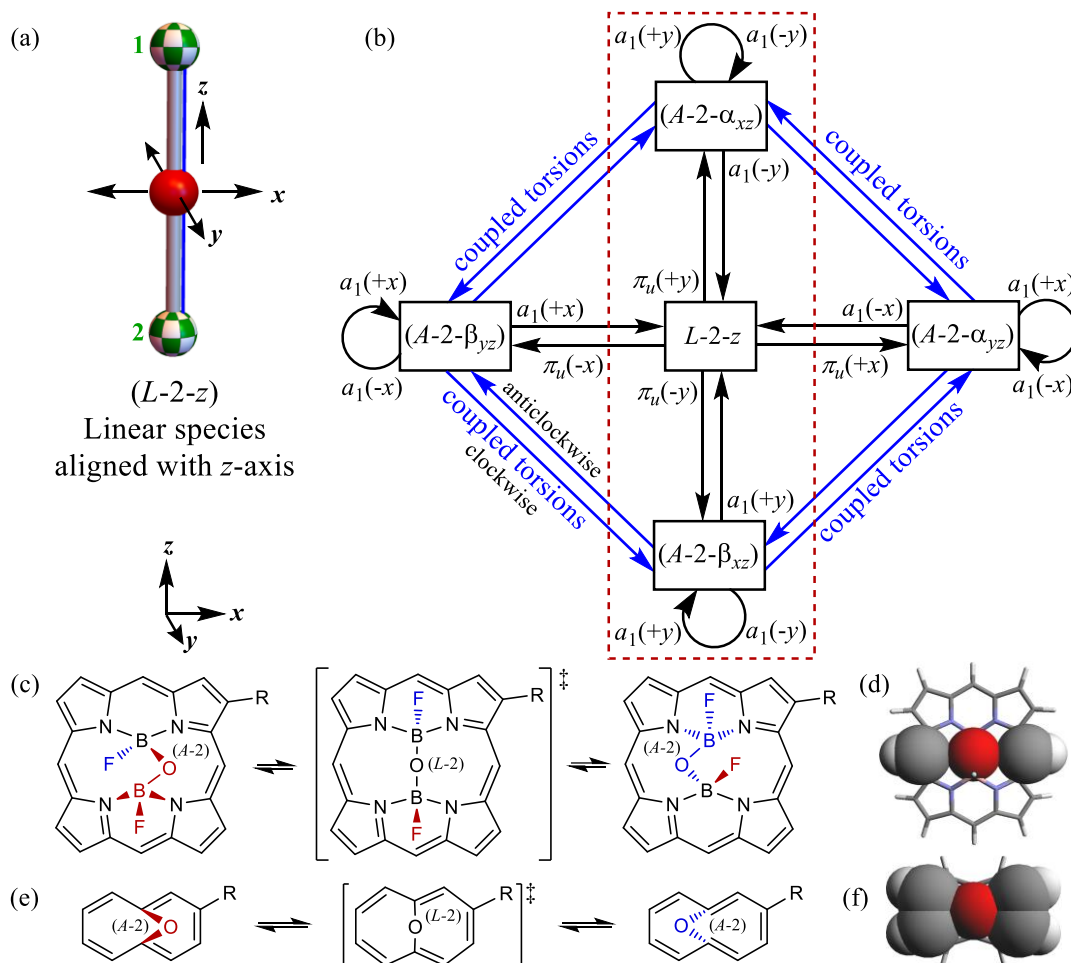
Figure 3.13 Full scope of stereoisomerism and associated $R_{st}^c 1$ stereoisomerisation of an embedded ML_2 unit. (a) The embedding of the linear ML_2 unit provides an external frame of reference. Here, the *L*-2 polytope species is aligned with the *z*-axis with *x*-axis horizontal and in the plane of the page, and the *y*-axis is normal to the plane of the page. With reference to these axes, the ML_2 unit is assigned the configuration symbol (*L*-2-*z*). (b) Directed graph representation indicating all possible motions of the M centre relative to the L "atoms". For (*L*-2-*z*), each direction of two degenerate orthogonal vibrational normal modes of π_u generic symmetry, aligned to the *x* and *y* directions are indicated changing the (*L*-2-*z*) polytope species to each of the angular species indicated. Displacement in the positive and negative directions are assigned α and β stereodescriptor modifiers, respectively, to define polyhedral absolute configuration symbols. The subscripts to these modifiers indicate the plane normal to the displacement vectors in each case. For each *A*-2 species, a flexional vibration of a_1 generic symmetry in the directions indicated either return to the *L*-2 polytope or remains *A*-2 and is hence semi-terminal. The directed graph edges indicated in blue together represent the coupled torsional clockwise (outer track) and anticlockwise (inner track) motions comprising hula-twist-like mechanisms. If further significant steric constraints are present, then one flexional mode can be suppressed. If flexional motion in the *x* direction is *fully* suppressed, then only out-of-plane flexional motion (in the *y* direction) is allowed, these being indicated by the enclosing red dashed box. Further, *all* coupled-torsions based hula-twist-like motions are also suppressed resulting in a strictly BAR geometric relationship and associated $R_{st}^c 1$ stereoisomerisation mechanism. This is the directed graph of the complete and concise set of embedded ML_2 polytopes. (c) Molecular embodiment of the embedding principle and suppression of one flexional mode in an asymmetrically substituted *transoid* B(F)OB(F)-porphyrin yielding *diastereomeric* BAR-related stereoisomers. (d) Key intramolecular interactions to the *L*-2 oxygen atoms of the (unsubstituted) transition structure shown in (c) suppressing motion in the *x*-direction. These are indicated by space-filling models of the atoms showing sub van der Waals contacts. (e) Embedding principle in an asymmetrically substituted 11-oxabicyclo[4.4.1]undeca-1,3,5,7,9-pentaene yielding *enantiomeric* BAR-related stereoisomers where, in each case, the *A*-2 species represent local minima and the *L*-2 species first-order transition structures. (f) Key intramolecular interactions to the *L*-2 oxygen atoms of the (unsubstituted) transition structure shown in (e) and similarly indicated as in (d). In (c) and (e), atoms and bonds indicated in red are above the plane of the page, those in blue are below.

In Figure 3.13a, an *L*-2 polytope is embedded with the Cartesian coordinate axes indicated. Figure 3.13b presents the directed graph showing all allowed motions for the concise ML_2 polytope set. An analogous graph could also be constructed using the expanded polytope set with the result appearing as three crossed and concentric diamonds instead of the single diamond shown in Figure 3.13b. To uniquely label all the embedded ML_2 polytopes, the existing polyhedral symbols are amended with reference to the coordinate axes. The linear form is (*L*-2-*z*) as it aligns with the *z*-axis. Each of the bond angle flexions arising from the orthogonal and degenerate vibrational normal modes of π_u generic symmetry are aligned with the *x*- and *y*-axes as an introduced convention. All directions for these displacement vectors lead to distinct *A*-2 polytopes. With the embedding, these now have absolute configurations. The configuration symbols are constructed to the following scheme: (*A*-2-X_{ab}) where X is either α or β depending on whether the relative displacement of M is in the positive or negative direction, respectively, and *ab* defines the Cartesian plane normal to the displacement.

For all the *A*-2 species, there exists a vibrational normal mode of a_1 generic symmetry that either returns the geometry to the *L*-2 polytope or simply further distorts the bond angle to values small enough to depict constitutional isomerism. Hence the *A*-2 polytopes are semi-terminal, as was described above in Section 3.5.1. An emergent property that arises from the embedding is that torsional motions for the orientable ligand atoms must now be coupled, resulting in a geometric change to an adjacent *A*-2 polytope. These are indicated by the blue directed edges in Figure 3.13b. A process comprised of such motions is called the hula-twist-like mechanism (see Section 2.4.11).

In Figure 3.13c and Figure 3.13e are two simple examples of an embedding bent X–O–X group inside rigid bicyclo molecules with B(F)–O–B(F) ligated to the inner cavity of a porphyrin and C–O–C bridge of the O-bridged [10]annulene, respectively. In each of these examples, the embedding is such that it completely suppresses motion of the O atom in the *x*-direction due to strong steric interactions as depicted in Figure 3.13d and Figure 3.13f. By suppressing all motion in the *x*-direction, it can be seen from the graph in Figure 3.13b that only   the (*A*-2-β_{xz}) ⇌ (*L*-2-*z*) ⇌ (*A*-2-α_{xz}) BAR processes (as indicted by the enclosing, red-dashed box) are now permitted. Significantly, the coupled torsions are fully suppressed, hence only the

BAR mechanism remains physically viable. The implications for stereochemistry of this type of embedding are profound and explored later in Chapters 4, 5, and 7.

3.6   The ML$_3$ family

The principles identified for ML$_2$ can also be applied to ML$_3$. For ML$_3$, the seed polytope (Table 3.2) is trigonal planar with D_{3h} symmetry. This stereocentre features six internal vibrational modes, of which three are bond-stretches leading to constitutional isomerism and three involve angle/torsion variations that lead to either stereoisomerism or constitutional isomerism. As only stereoisomerism is of current interest, the dissociative modes are neglected. Complete structural definition then requires specification of the three L–M–L bond angles and the improper torsional angle τ. For planar structures, $\tau = 0$ and the three bond angles sum to 360°; for non-planar structures, the three bond angles sum to less than 360°. Large-magnitude torsions and/or small bond angles force the ligands to overlap and hence depict constitutional isomerism.

To make a complete ML$_3$ polytope set for stereoisomerism, all three angle/torsion vibrations need to be considered in all directions. For the ML$_3$ seed polytope, the bending vibration is doubly degenerate and hence "all directions" includes all possible linear combination of the two components describing the mode specification and hence, like the case for L-2, presents an infinite number of possibilities. The specification of some finite set of polytopes therefore demands that this infinite range of possibilities be partitioned into chemically meaningful units. The most appropriate coordinate system for considering ML$_3$ is neither Cartesian coordinates nor internal coordinates comprised of bond and torsion angles. Instead, the torsion angle combined with a polar representation of the doubly degenerate bending vibration, (a, ϕ, τ), provides an unbiased minimalist description. With this specification, a is the magnitude of the distortion of the doubly degenerate mode, and ϕ is its phase.

The situation in which a Jahn-Teller distortion[98] occurs when a doubly degenerate electronic state interacts with a doubly degenerate vibration is common and well-studied. The complex network of structures that can arise directly reflect the polytopes of ML$_3$.

In effect, the torsional angle determines the nature of planar versus non-planar polytopes, whilst the three bent L–M–L angles combine to determine the polytopes that can

exist at each torsional value. For planar structures, a variety of angle patterns can occur: all three angles equal to 120°, two equal angles that are either less than or greater than 120°, one angle equal to 120°, and all angles unequal with either two less than 120° or else two greater than 120°. Each of these categories therefore can define polytope genera in an embedded environment where all discrete combinations of configurations are allowed. The genus with all angles equal, *G1P*-3, gives rise to only one species (*G1P*-3) – it is "monotypic"; 8 other genera could arise from this partitioning, each of which supports 3 species. Hence 25 planar embedded species are possible. The 24 species other than (*G1P*-3) differ from each other simply by 15° advancements in the polar angle ϕ for fixed a and τ. Distorting out of plane in each direction generates two symmetry-related sets of genera and species, making for a total of 18 possible genera and 75 possible species. These are listed in Table 3.4 using a typical value for both a and τ, highlighting a periodic cyclic path for $R_{st}^c 1$ reactions as a function of the angular variable ϕ. Coordinate files and generating code are found in **E_File_1**.

Inspection of Table 3.4 shows that groups of three polytope species with the same generic symmetry point group (C_s for planar polytopes, C_1 for non-planar ones) regularly appear. This polytope set is therefore not concise, rather mimicking the expanded polytope set for ML$_2$ that included the (*OA*-2), (*RA*-2), and (*AA*-2) species. A concise set of ML$_3$ polytopes can be obtained simply by deleting the (*G3xP*-3-*i*), (*G5xP*-3-*i*), (*G3xPY*-3-*i*-*y*), and (*G5xPY*-3-*i*-*y*) species from the table, where *x*, *y* = "*C*" or *"A"* and $i \in \{1, 2, 3\}$. This means that structures are sampled at 30° intervals in ϕ rather than at 15° intervals, with the larger set being that produced by trifurcation of all the (*G4xP*-3-*i*) and (*G4xPY*-3-*i*-*y*) species in the concise set. The concise set allows many features of the Jahn-Teller distortion to be represented and is expected to be adequate for most chemical scenarios; the extended set provides options for the description of more subtle chemical features, when required

Table 3.4 ML$_3$ polytope species ordered by torsional displacement τ, and then by their ϕ angle. Results are shown for an extended polytope set. The seed polytope (*G1P*-3) species is shaded blue. Adjacent polytope species of the same generic symmetry point group are shaded grey as single blocks. Species from genera *G3X*-3 and *G5X*-3, where X = *P* or *PY*, are not included in the concise set listed in Figure 3.14.

ϕ / °	species	generic point group	θ_{L1ML2} / °	θ_{L2ML3} / °	θ_{L3ML1} / °	τ / °
undefined	(*G1PY*-3-*C*)	C_{3v}	116.9	116.9	116.9	−30
0	(*G2PY*-3-1-*C*)	C_s	144.4	99.4	99.4	−30
15	(*G3APY*-3-1-*C*)	C_1	143.8	93.6	107.1	−30
30	(*G4APY*-3-1-*C*)	C_1	141.6	89.6	115.6	−30
45	(*G5APY*-3-1-*C*)	C_1	137.7	87.5	124.2	−30
60	(*G6PY*-3-1-*C*)	C_s	132.3	87.3	132.3	−30
75	(*G5CPY*-3-1-*C*)	C_1	125.6	88.9	139.1	−30
90	(*G4CPY*-3-1-*C*)	C_1	118.3	92.3	144.2	−30
105	(*G3CPY*-3-1-*C*)	C_1	110.7	97.3	147.5	−30
120	(*G2PY*-3-2-*C*)	C_s	103.6	103.6	148.6	−30
135	(*G3APY*-3-2-*C*)	C_1	97.3	110.7	147.5	−30
150	(*G4APY*-3-2-*C*)	C_1	92.3	118.3	144.2	−30
165	(*G5APY*-3-2-*C*)	C_1	88.9	125.6	139.1	−30
180	(*G6PY*-3-2-*C*)	C_s	87.3	132.3	132.3	−30
195	(*G5CPY*-3-2-*C*)	C_1	87.5	137.7	124.2	−30
210	(*G4CPY*-3-2-*C*)	C_1	89.6	141.6	115.6	−30
225	(*G3CPY*-3-2-*C*)	C_1	93.6	143.8	107.1	−30
240	(*G2PY*-3-3-*C*)	C_s	99.4	144.4	99.4	−30
255	(*G3APY*-3-3-*C*)	C_1	106.5	143.3	93.1	−30
270	(*G4APY*-3-3-*C*)	C_1	114.5	140.5	88.5	−30
285	(*G5APY*-3-3-*C*)	C_1	122.5	135.9	85.7	−30
300	(*G6PY*-3-3-*C*)	C_s	129.8	129.8	84.8	−30
315	(*G5CPY*-3-3-*C*)	C_1	135.9	122.5	85.7	−30
330	(*G4CPY*-3-3-*C*)	C_1	140.5	114.5	88.5	−30
345	(*G3CPY*-3-3-*C*)	C_1	143.3	106.5	93.1	−30
360	(*G1PY*-3-*C*)	C_s	144.4	99.4	99.4	−30
undefined	(*G1P*-3)	D_{3h}	120.0	120.0	120.0	0
0	(*G2P*-3-1)	C_{2v}	150.0	105.0	105.0	0
15	(*G3AP*-3-1)	C_s	149.0	98.8	112.2	0
30	(*G4AP*-3-1)	C_s	146.0	94.0	120.0	0
45	(*G5AP*-3-1)	C_s	141.2	91.0	127.8	0
60	(*G6P*-3-1)	C_{2v}	135.0	90.0	135.0	0
75	(*G5CP*-3-1)	C_s	127.8	91.0	141.2	0
90	(*G4CP*-3-1)	C_s	120.0	94.0	146.0	0
105	(*G3CP*-3-1)	C_s	112.2	98.8	149.0	0
120	(*G2P*-3-2)	C_{2v}	105.0	105.0	150.0	0
135	(*G3AP*-3-2)	C_s	98.8	112.2	149.0	0
150	(*G4AP*-3-2)	C_s	94.0	120.0	146.0	0
165	(*G5AP*-3-2)	C_s	91.0	127.8	141.2	0
180	(*G6P*-3-2)	C_{2v}	90.0	135.0	135.0	0
195	(*G5CP*-3-2)	C_s	91.0	141.2	127.8	0
210	(*G4CP*-3-2)	C_s	94.0	146.0	120.0	0
225	(*G3CP*-3-2)	C_s	98.8	149.0	112.2	0
240	(*G2P*-3-3)	C_{2v}	105.0	150.0	105.0	0
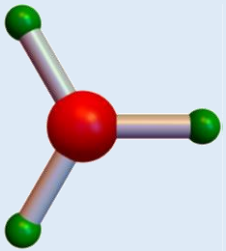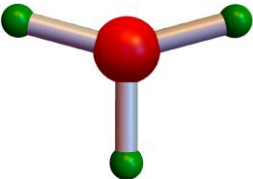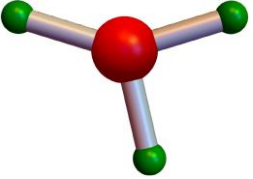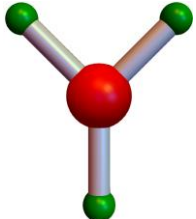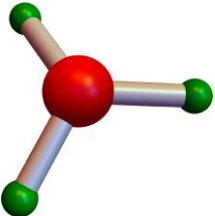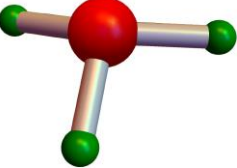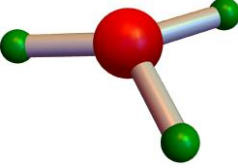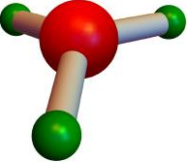
ϕ / °	species	generic point group	θ_{L1ML2} / °	θ_{L2ML3} / °	θ_{L3ML1} / °	τ / °
255	(*G3AP*-3-3)	C_s	112.2	149.0	98.8	0
270	(*G4AP*-3-3)	C_s	120.0	146.0	94.0	0
285	(*G5AP*-3-3)	C_s	127.8	141.2	91.0	0
300	(*G6P*-3-3)	C_{2v}	135.0	135.0	90.0	0
315	(*G5CP*-3-3)	C_s	141.2	127.8	91.0	0
330	(*G4CP*-3-3)	C_s	146.0	120.0	94.0	0
345	(*G3CP*-3-3)	C_s	149.0	112.2	98.8	0
360	(*G1P*-3)	C_{2v}	150.0	105.0	105.0	0
undefined	(*G1PY*-3-*A*)	C_{3v}	116.9	116.9	116.9	30
0	(*G2PY*-3-1-*A*)	C_s	144.4	99.4	99.4	30
15	(*G3APY*-3-1-*A*)	C_1	143.8	93.6	107.1	30
30	(*G4APY*-3-1-*A*)	C_1	141.6	89.6	115.6	30
45	(*G5APY*-3-1-*A*)	C_1	137.7	87.5	124.2	30
60	(*G6PY*-3-1-*A*)	C_s	132.3	87.3	132.3	30
75	(*G5CPY*-3-1-*A*)	C_1	125.6	88.9	139.1	30
90	(*G4CPY*-3-1-*A*)	C_1	118.3	92.3	144.2	30
105	(*G3CPY*-3-1-*A*)	C_1	110.7	97.3	147.5	30
120	(*G2PY*-3-2-*A*)	C_s	103.6	103.6	148.6	30
135	(*G3APY*-3-2-*A*)	C_1	97.3	110.7	147.5	30
150	(*G4APY*-3-2-*A*)	C_1	92.3	118.3	144.2	30
165	(*G5APY*-3-2-*A*)	C_1	88.9	125.6	139.1	30
180	(*G6PY*-3-2-*A*)	C_s	87.3	132.3	132.3	30
195	(*G5CPY*-3-2-*A*)	C_1	87.5	137.7	124.2	30
210	(*G4CPY*-3-2-*A*)	C_1	89.6	141.6	115.6	30
225	(*G3CPY*-3-2-*A*)	C_1	93.6	143.8	107.1	30
240	(*G2PY*-3-3-*A*)	C_s	99.4	144.4	99.4	30
255	(*G3APY*-3-3-*A*)	C_1	106.5	143.3	93.1	30
270	(*G4APY*-3-3-*A*)	C_1	114.5	140.5	88.5	30
285	(*G5APY*-3-3-*A*)	C_1	122.5	135.9	85.7	30
300	(*G6PY*-3-3-*A*)	C_s	129.8	129.8	84.8	30
315	(*G5CPY*-3-3-*A*)	C_1	135.9	122.5	85.7	30
330	(*G4CPY*-3-3-*A*)	C_1	140.5	114.5	88.5	30
345	(*G3CPY*-3-3-*A*)	C_1	143.3	106.5	93.1	30
360	(*G1PY*-3)	C_s	144.4	99.4	99.4	30

When embedded, the concise set embodies 10 genera and 39 species. These are shown listed in Figure 3.14, partitioned into planar structures represented by 5 genera and 13 species, as well as non-planar structures represented by 5 genera and 26 species. The non-planar genera can be conceived as being derived from the planar genera simply by distortion out of plane. The planar genera are: *G1P*-3 of D_{3h} generic symmetry, the "T-shaped" genus *G2P*-3 of C_{2v} generic symmetry, the "Y-shaped" genus *G6P*-3, also of C_{2v} generic symmetry, and the *G4AP*-3 and *G4CP*-3 genera of *Cs* generic symmetry. The two genera of C_{2v} symmetry are differentiated based on the equal L–M–L bond angles being either less than or greater than 120°, whereas the two genera of C_s symmetry are differentiated by their

71

embedding in the external environment, with "*A*" and "*C*" representing the rotational ordering of the ϕ angles with respect to an external reference. These distinctions allow for the species of any planar ML_3 stereocentre to be uniquely assigned. Hence the Bürgi and Dunitz structural correlation procedure is not required for structural assignments, by analogy to the result found already for the concise polytope set of ML_2. To apply the expanded polytope set, structural correlation would be required to differentiate between (*G*3*xP*-3-*i*), (*G*4*xP*-3-*i*), (*G*5*xP*-3-*i*), (*G*3*xPY*-3-*i*-*y*), (*G*4*xPY*-3-*i*-*y*), and (*G*5*xPY*-3-*i*-*y*)

The associated $R_{st}^c 1$ mechanisms that interconvert the 39 embedded concise set of species are depicted as a directed graph in Figure 3.15. As before, semi-terminal polytopes that present hard edges to the regime of stereoisomerism are indicated using cyclic arrows. The reaction mechanisms are classified as to whether they depict torsional (τ), radial (a), or phase (ϕ) reaction coordinates. The polytope species at the centre of the Jahn-Teller distortion, (*G*1*P*-3), (*G*1*PY*-3-*C*), and (*G*1*PY*-3-*A*) can react to form any species with the same torsional value through radial distortion and interconvert between these subsets by torsional distortion. All other polytope species can undergo precisely six reactions, some of which lead to constitutional isomerism as indicated in the figure. Normalised coordinates and generating Fortran code comprise **E_File_1**.

Figure 3.14 The concise set of ML$_3$ genera and constituent species. For each genus, the generic symmetry point group is indicated along with a full listing of species differentiated by their unique vertex index permutations. The monotypic highest-symmetry seed genus *G1P*-3 is indicated by the blue shading.

Planar genera:		"T-shaped"		"Y-shaped"
Generic point group:	D_{3h}	C_{2v}	C_s	C_s	C_{2v}
Genus symbol:	*G1P*-3	*G2P*-3	*G4AP*-3	*G4CP*-3	*G6AP*-3
Concise embedded species:	(*G1P*-3)	(*G2P*-3-1)	(*G4AP*-3-1)	(*G4CP*-3-1)	(*G6AP*-3-1)
		(*G2P*-3-2)	(*G4AP*-3-2)	(*G4CP*-3-2)	(*G6AP*-3-2)
		(*G2P*-3-3)	(*G4AP*-3-3)	(*G4CP*-3-3)	(*G6AP*-3-3)

Non-planar genera:				
Generic point group:	C_{3v}	C_s	C_1	C_1	C_s
Genus symbol:	*G1PY*-3	*G2PY*-3	*G4APY*-3	*G4CPY*-3	*G6APY*-3
Species:	(*G1PY*-3-*C*)	(*G2PY*-3-1-*C*)	(*G4APY*-3-1-*C*)	(*G4CPY*-3-1-*C*)	(*G6APY*-3-1-*C*)
		(*G2PY*-3-1-*A*)	(*G4APY*-3-1-*A*)	(*G4CPY*-3-1-*A*)	(*G6APY*-3-1-*A*)
	(*G1PY*-3-*A*)	(*G2PY*-3-2-*C*)	(*G4APY*-3-2-*C*)	(*G4CPY*-3-2-*C*)	(*G6APY*-3-2-*C*)
		(*G2PY*-3-2-*A*)	(*G4APY*-3-2-*A*)	(*G4CPY*-3-2-*A*)	(*G6APY*-3-2-*A*)
		(*G2PY*-3-3-*C*)	(*G4APY*-3-3-*C*)	(*G4APY*-3-3-*C*)	(*G6APY*-3-3-*C*)
		(*G2PY*-3-3-*A*)	(*G4APY*-3-3-*A*)	(*G4APY*-3-3-*A*)	(*G6APY*-3-3-*A*)

Figure 3.15 Directed graph representation of all species in the concise set of ML$_3$ polytopes connected by their R$_{st}^c$1 mchanisms. The colour-coded directed edges depict coordinate changes: blue are the torsional (τ) changes, magenta are the radial (a) changes, and red are the phase (ϕ) angle changes. Coordinates are found in **E_File_1**.

3.7   Conclusions

The polytopal-rearrangements model has been very successful in describing observed reactions in inorganic complexes, rationalising observed chemistry and dynamical behaviour in stereocentres often with significant size and complexity. This approach focused on the reactions, but in doing so the model also depicted iconic structures that could be identified as isomers (potential energy local minima) or transition states. IUPAC currently recognises the polytopal-rearrangements model's contributions to inorganic stereoisomerisation, but with less recognition of the associated stereoisomerism aspects. The geometries associated with the stereoisomerism were typically used as "inputs" into the model rather than consequences of and underlying *a priori* mathematical basis. The presentation here has reviewed key developments of the polytopal-rearrangement model, emphasising that it *implicitly* presents groundwork for a unified description to both stereoisomerism and stereoisomerisation. Indeed, its central tenet is to focus on sets of polytopes that are *closed* under the operation of the $R_{st}^c 1$ reactions involving all non-dissociative vibrational modes.

The general principles of the polytopal-rearrangements model were described in detail through a step-by-step process facilitated by introduced language needed to precisely and concisely describe key properties and relationships. The description focussed on the example of the ML_4 family as it is familiar to both organic and inorganic chemists alike and is sufficiently complex to demonstrate detail without being overly so. Whilst many polytope genera were described for this ML_4 system, it was clear that many additional low-symmetry genera would also be required for a complete set.

Based upon the main principles of the polytopal-rearrangements model of stereoisomerisation, the Polytope Formalism for stereoisomerism (and stereoisomerisation) is presented, embodying four principal advances. The first is the inclusion of torsions for the M–L bonds through the introduction of the "orientability" property of the L-vertex "atoms", this being critical for the description of all classical conformational isomerism.

The second advance was to formally include all non-dissociative vibrational modes without bias for the lowest energy modes. This allows for the description of high-energy reactions or high-energy species not previously included for the systems studied. Of note, however, what are high-energy species and/or barriers for one chemical system may turn out to be the isomers and barriers of critical interest in some other system. Hence this extension

provides for the generalisation of chemical understanding obtained in one context to apply to many more.

Further, *all directions* of the non-dissociative vibrational modes have been included and not simply one direction corresponding to a predetermined reaction pathway. For degenerate modes, this generalises to the requirement that all possible linear combinations of the mode components need to be included, giving rise to an infinite number of possibilities. Various ways of reducing this infinite number to finite representations are possible, with each representation leading to polytope sets that are closed under the operation of all included $R_{st}^c 1$ mechanisms interconnecting them. For ML_2 and ML_3, concise polytope sets were found by this process within which all $R_{st}^c 1$ mechanisms only interconvert polytopes of *different symmetry*. Also, it was found that some of the vibrational modes would lead to increasingly close ligand-ligand contact which, in a real chemical system, would lead to constitutional isomerism. To provide a way to account for these motions and maintain set closure, the concept of the semi-terminal property was introduced which, on a directed reaction graph, is represented as a self-loop.

The fourth advance is the principle of *embedding* which takes the orientability property of the L-vertex "atoms" and couples them together. This embedding principle implicitly embodies the semi-rigid molecular structure beyond the immediate ligand atoms and provides a frame of reference external to the stereocentre. This was demonstrated to be critically important for the ML_2 family. Due to the low dimensionality of ML_2 polytopes in $\mathbb{R}^3$ space, an external frame of reference is needed to manifest the full scope of ML_2 stereoisomerism and associated $R_{st}^c 1$ mechanism. This embedded ML_2 stereoisomerism scope was presented along with the notion that sufficient steric constraints from an embedding could suppress some motions. Full suppression of one direction of bond-angle flexion was shown to lead to the situation where the only remaining stereoisomerisation mechanism was a single bond-angle reflexion orthogonal to the suppressed motion. Importantly, the coupled torsional motions representing the "hula-twist-like" motions were also suppressed. Examples of molecular systems embodying these principles were given. The desire to understand this level of detail in the ML_2 system was the driving impetus for the research comprising this Thesis, following on from the discovery of compounds with ML_2 stereocentres that could not be described adequately using existing IUPAC chemical names and reaction specifications, as detailed in Chapter 4.

It is proposed that the concise polytope sets identified for ML_2 and ML_3 should be able to describe most stereocentres belonging to these families. Nevertheless, larger polytope sets allow for additional chemical variations to be described, when needed. A feature identified for both ML_2 and ML_3 families is that polytopes in a concise set can be trifurcated to make a larger polytope set. As concise polytope sets only support R_{st}^c1 mechanisms interconnecting polytopes of different symmetry, it is always possible to assign any real stereocentre to one of the polytopes based on just symmetry information combined with basic properties of the structure and its embedding. For larger polytope sets in which structures of the same symmetry are connected, it will always be necessary to introduce some structural correlation procedure to assign stereocentres to polytopes.

4 Experimental demonstration of bond-angle reflection stereoisomerism

4.1 Summary

Stereoisomerism is a fundamental chemical concept, reflecting the fact that the arrangement of atoms in a molecular entity has a profound influence on its chemical and physical properties. Here is described a previously unclassified fundamental form of stereoisomerism, through the synthesis of four resolved stereoisomers of a *transoid* B(F)−O−B(F)-quinoxalinoporphyrin. These comprise two pairs of enantiomers that manifest structural relationships not describable within existing IUPAC nomenclature and terminology. They undergo thermal diastereomeric interconversion over a barrier of 104±2 kJ mol^{-1}, which is termed "akamptisomerisation". Feasible interconversion processes between conceivable synthesis products and reaction intermediates were mapped out by density-functional theory calculations, identifying bond-angle reflexion (BAR) at a singly bonded atom as the reaction mechanism and an $R_{st}^c 1$ process. Necessary BAR-related stereodescriptors *parvo* and *amplo* are introduced.

4.2 Background

4.2.1 Overview of work performed prior to this Thesis and as part of this Thesis

A new form of stereoisomerism designated as "akamptisomerism" was described in the early stages of the work undertaken for this Thesis. Before the work described in this thesis, the synthesis of B(F)−O−B(F)-quinoxalinoporphyrins had been achieved at the University of Sydney. This work was undertaken to gain insight into the optoelectronic properties of the quinoxalino[2,3−*b'*]porphyrin structural motif, as exemplified by **1** (See Figure 4.1), and to validate theoretical methods used to model such properties.[99] The choice of these B(F)−O−B(F)-quinoxalinoporphyrins was based upon the understanding that they were expected to be stable chiral compounds amenable to chiroptical characterisation, in particular, using Circular Dichroism (CD) and Magnetic Circular Dichroism (MCD) spectroscopies. These compounds were to serve as model systems for molecular electronic

devices[100-127] and to assign the origins of analogous optoelectronic properties in some biological systems and biomimetic models.[100-110, 112, 128-141]

Prior to commencement of this Thesis, the initial synthesis was carried out by Dr Iain Blake, and resolution and isolation of compounds by the author of this Thesis, Peter Canfield, in the Crossley Laboratory. Nuclear Magnetic Resonance (NMR) Spectroscopy measurement and kinetics determinations were carried out by Dr Ian Luck. Chiroptical studies were carried out by Prof Elmars Krausz, Dr Zheng-Li Cai and the author of this Thesis, Peter Canfield. Modelling was carried out by Prof Jeffrey Reimers, Dr Rika Kobayashi and the author of this thesis, Peter Canfield.

At the outset of studies towards a PhD as described in this Thesis, there were still aspects requiring detailed investigation and rationalisation. In the initial stages of the research undertaken for this Thesis, these syntheses and chiral resolution were repeated by Prof Maxwell Crossley and the author of this Thesis. In the course of this Thesis work, structural assignments based upon the observed CD spectra, further NMR measurements and Density-Functional Theory (DFT) modelling was carried out and extensive development of the conceptual framework was carried out by the author of this Thesis, Peter Canfield, under the supervision of Prof Maxwell Crossley and associate supervisor Prof Jeffrey Reimers. This led to the publication of: "A new fundamental type of conformational isomerism", Peter J. Canfield, Iain M. Blake, Zheng-Li Cai, Ian J. Luck, Elmars Krausz, Rika Kobayashi, Jeffrey R. Reimers & Maxwell J. Crossley, Nature Chemistry volume 10, pages 615–624 (2018).[7] Subsequent to this publication, the implications pertaining to the basic understanding of isomerism and isomerisation were further developed and elaborated in the course of this doctoral candidature. This work is as yet unpublished, and it is also presented in this Thesis Chapter.

4.2.2 *The experimentally posed challenges addressed in this Thesis*

It was known[142-145] that the *transoid* B(F)OB(F) group could be ligated inside the C_{2v} point-group symmetry macrocycle (see Section A.28) cavity in four distinct ways with concomitant molecular distortions leading to chiral structures. This chirality was expected to produce measurable CD spectra in the visible region that could be used to help deconvolve the absorption bands into their different polarisation components. To this end, chiral stationary phase High-Performance Liquid Chromatography (HPLC) separation of the four

synthetic products had been conducted[7] yielding, in order of elution, fractions **Fr1**, **Fr2**, **Fr3** and **Fr4**. The four products were reasoned to have the stereoisomeric chemical structures shown in Figure 4.1a-b, structures that comprise two sets of enantiomers.

(a)

with **Fr1 − Fr4** being comprised of the putative *transoid* structures:

(b)

plan

side

(c)

1 = H$_2$-pqx

Ar = 3,5-di-*tert*-butylphenyl

4 × B$_2$OF$_2$pqx stereoisomers

Racemate "A"

Fr1 ⇌✗⇌ **Fr2**

Fr4 ⇌✗⇌ **Fr3**

Racemate "B"
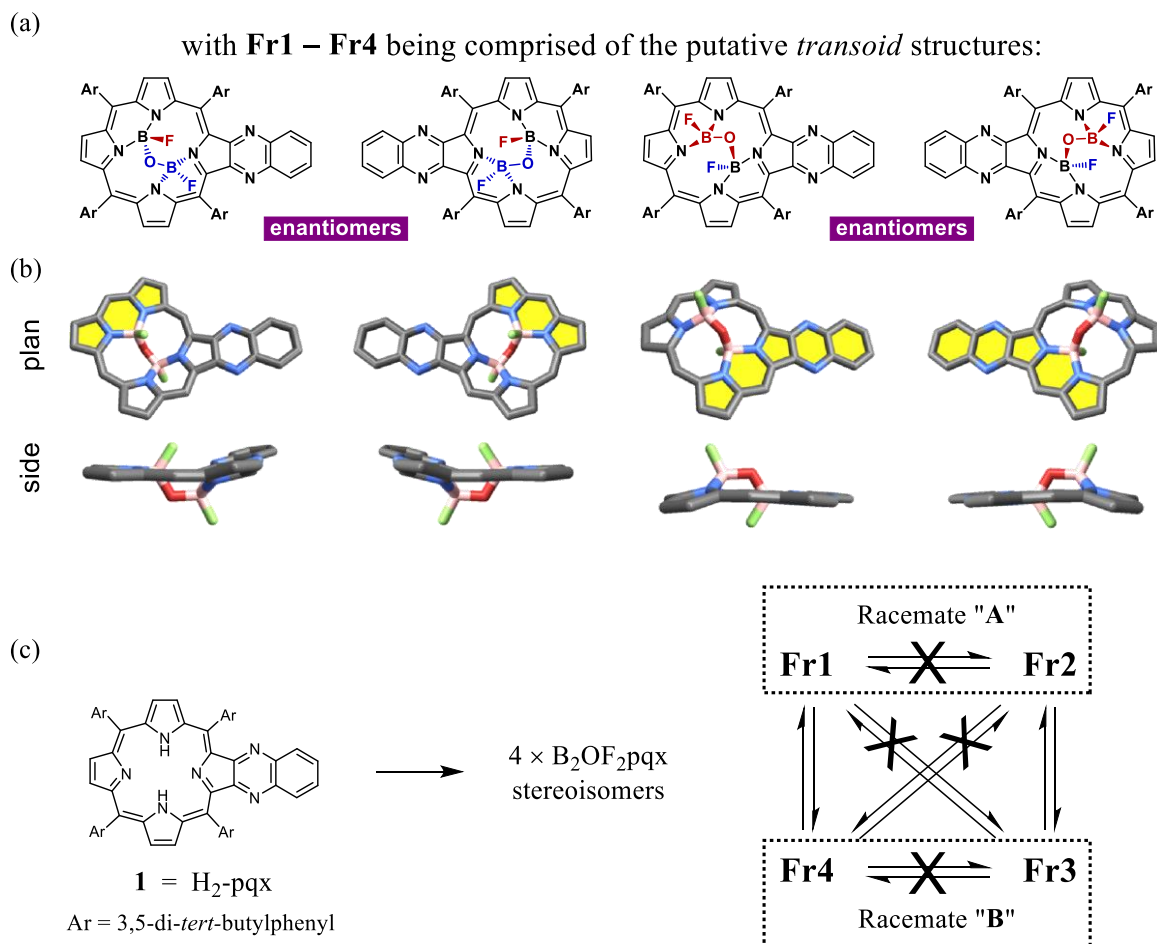
Figure 4.1 Stereoisomerisation of four putative *transoid* B(F)−O−B(F)-quinoxalinoporphyrin reaction products and their DFT-calculated structures.[7] (a) The 4 putative *transoid* ligation stereoisomers grouped as pairs of enantiomers. Atoms and bonds marked in red are above the quasi-plane of the porphyrin whilst those in blue are below it. (b) DFT-calculated structural models corresponding to (a) shown in plan and side view. Hydrogen atoms and the aryl side chains are omitted for clarity, with rings coplanar to a chelated in-plane boron atom shaded yellow to make clear the stereoisomeric differences. Atom colours: white- H, peach- B, grey- C, blue- N, red- O, green- F. (c) Observed stereoisomerisation pattern in simple solutions of the HPLC-separated fractions **Fr1 – Fr4**. The 2 racemates produced by flash-silica column chromatography are indicated as "A" and "B" within the dotted boxes.

During this HPLC separation process, there was the serendipitous discovery of an unexpected interconversion of pairs of isomers, specifically *only* between diastereomeric pairs (see Section A.13) **Fr1** and **Fr4**, and between **Fr2** and **Fr3** occurring on the order of hours at ambient temperatures. Racemisation interconversions within the enantiomeric pairs **Fr1** and **Fr2**, and **Fr3** and **Fr4** were not observed in these simple solutions. This pattern of isomerisation was shown in Figure 4.1c.

The kinetics of the isomerisation was determined[7] by tracking the distinct diastereomeric signals using variable temperature ^{19}F NMR measurements starting with enantiopure samples. It was shown[7] to follow first-order kinetics over an activation barrier of 104 ± 2 kJ mol^{-1}.

The challenges addressed in this Thesis Chapter concern the assignment of the four observed HPLC fractions to the four expected reaction products. This involved attempts to obtain single-crystal structural determinations, and the assignment of the observed NMR and CD spectra. A related challenge was the expansion of the possible-product molecular set to include options beyond the four compounds initially envisaged. Also, the HPLC fraction assignment needed to be consistent with the observed isomerisation reactions and their rates. Further, although multiple reaction products were expected, it was discovered that IUPAC rules did not provide convenient means to name the four compounds. The determination of a robust and general naming scheme therefore also presented a challenge, with broader implications concerning what these compounds revealed concerning the understanding of stereoisomerism and stereoisomerisation.

4.3 Determination of the composition of the observed compounds

4.3.1 Attempted X-ray crystallography studies

The most direct method for determination of the composition of fractions **Fr1** – **Fr4** is via X-ray crystallographic analyses of crystals. To this end, crystal growth was attempted for samples containing pure racemates. Whilst studies of mixtures would give results that are less informative than enantiopure fractions, these would nevertheless provide some structural information. These experiments, however, were hampered by observed and uncharacteristically high solubility of the molecules in all organic solvents, showing low solubility only in pure water. Eventually, crystals of sufficient quality needed for diffraction experiments were obtained from pure racemates. These were taken to the Australian Synchrotron for analysis. Supported measurement procedures at the Australian Synchrotron specify only a specific set of attachment methods for the sample crystals to the measurement probe, involving use of various glues. It was found that the crystals dissolved in all allowed glues.

Hence, the crystallography approach failed to yield useful information. Assignment of the observed fractions was therefore performed based on interpretation of NMR and CD data using a rationalised set of structural possibilities, driven by DFT calculations.

4.3.2 Structural possibilities – diversity of configurations and ligation modes of macrocycle-ligated B(F)OB(F) groups within quinoxalinoporphyrin

Without crystallographic information, it is essential that all possible chemical products of the reaction (see Figure 4.1c.) be considered when making structural assignments for the four observed product fractions.

The first consideration is the *configuration* of the B(F)OB(F) group. Figure 4.2 shows three dimensional (3D) structural examples of the *transoid* configuration (Figure 4.2a and the second structure of Figure 4.2f), an extended-*cisoid* configuration (Figure 4.2b, Figure 4.2d, Figure 4.2e and the first structure of Figure 4.2f), and the proposed hypothetical contracted-*cisoid* configuration shown in Figure 4.2c. Whilst the literature[142-145] indicates that porphyrin coordinated B(F)OB(F) groups adopting the *transoid* configuration are the likely isolated products from the syntheses (see Figure 4.2a), the extended-*cisoid* configuration (see Figure 4.2b) has also been observed in the related corrole system[146] (see Figure 4.2d) and recently in porphycenes[147] (see Figure 4.2e). Further, for the very closely related tetra-aza-porphyrin analogues; porphyrazines and phthalocyanines, both *transoid* and extended-*cisoid* configurations are *simultaneous products* from the syntheses.[148-149]

A key distinction between *transoid* and *cisoid* B(F)OB(F) configurations is that in the *transoid* case, the two boron atoms exhibit distinct geometric relationships to the macrocycle. For a *transoid* configuration, one boron is effectively "in-plane" with the half portion of the macrocycle chelating it, whilst the other boron atom is significantly "out-of-plane" with concomitant distortions to its chelating half of the macrocycle. For both the extended and contracted-*cisoid* configurations, both boron atoms exhibit the same relationships to their respective macrocycle chelating halves.
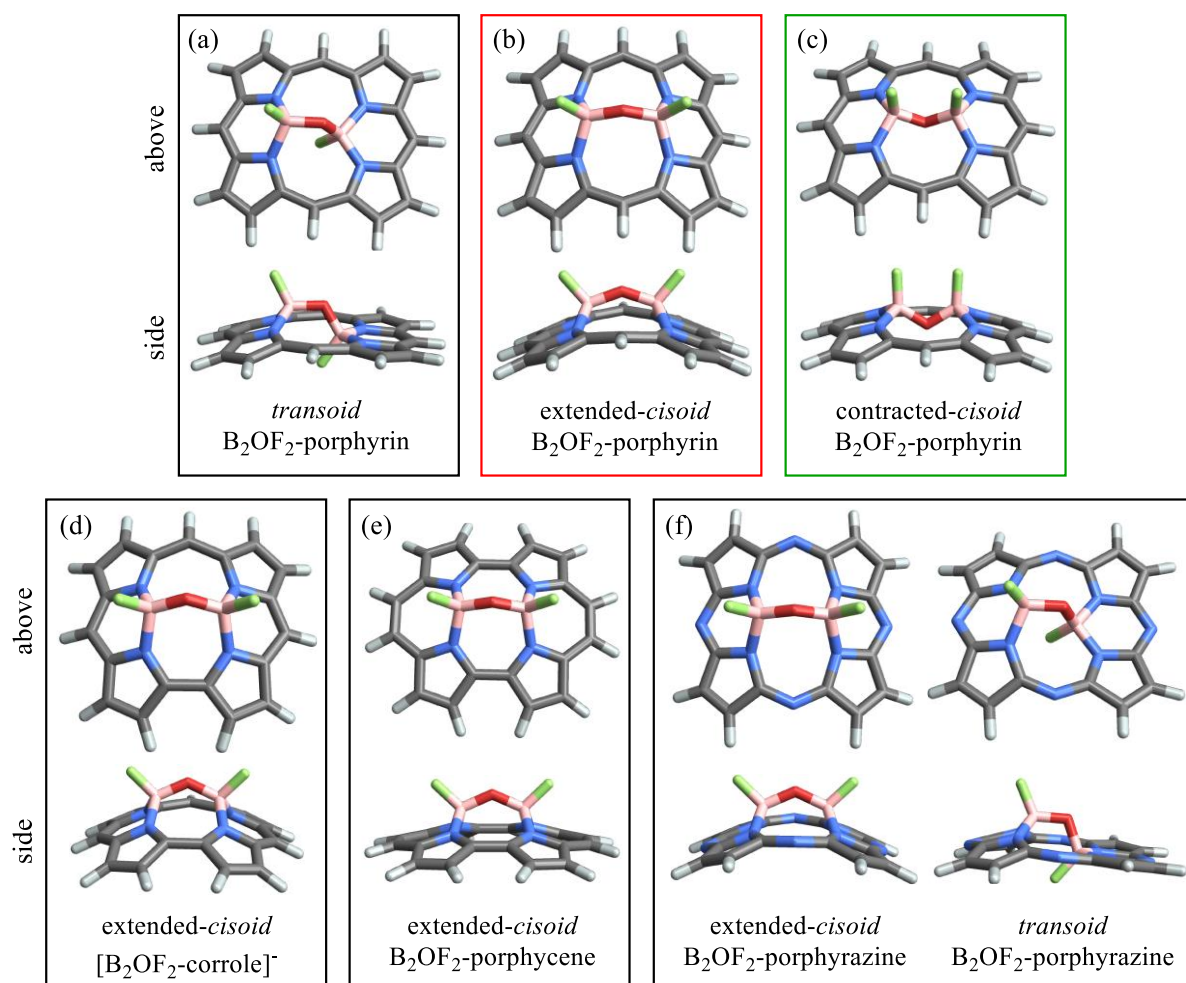
Figure 4.2 Possible structures of *transoid* and *cisoid* B(F)OB(F)-macrocycle complexes. All structures shown here feature unsubstituted macrocycles for clarity. (a) Two views of the *transoid* configuration of B(F)OB(F)-porphyrin. A *transoid* configuration is the exclusive isolated synthetic product for porphyrins. (b) Two views of an extended-*cisoid* configuration of B(F)OB(F)-porphyrin. Such a configuration has not been isolated from the reported syntheses. (c) Two views of a hypothetical contracted-*cisoid* configuration of B(F)OB(F)-porphyrin. Contracted-*cisoid* configurations have not been seen experimentally. (d) Two views of an extended-*cisoid* configuration of [B(F)OB(F)-corrole] anion. Syntheses on corroles exclusively yield this ligation mode and configuration. (e) Two views of an extended-*cisoid* configuration of B(F)OB(F)-porphycene. Synthesis on porphycenes exclusively yield this ligation mode and configuration. (f) Two views of each of the *transoid* and extended-*cisoid* configurations of B(F)OB(F)-porphyrazines. Both the *transoid* and *cisoid* configurations are simultaneous synthesis products. The closely related phthalocyanine macrocycle also yields the simultaneous *transoid* and *cisoid* configurations from the synthesis conditions. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

These distinctions have important consequences for the number of possible different ligation modes to 2,3-symmetrically disubstituted porphyrins like the quinoxalinoporphyrin **1**. According to the standard IUPAC recommendations for atom-locant numbering in macrocyclic rings[150], the quinoxalino-substituent to the porphyrin fixes the macrocycle inner-nitrogen locant numbers to those indicated in green in Figure 4.3.

Ar = 3,5-di-*tert*-butylphenyl

1^{2-} = pqx^{2-} B$_2$OF$_2$pqx
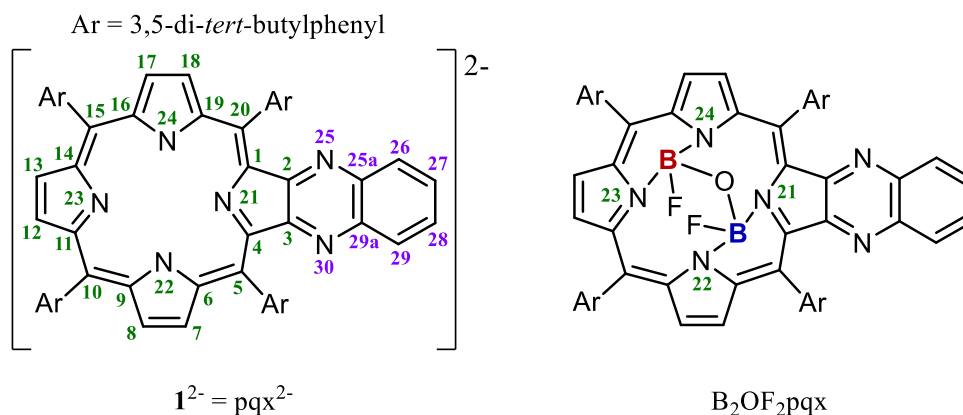
Figure 4.3 The dianion of **1**, denoted 1^{2-} or pqx^{2-}, exhibits the C_{2v} symmetry point group. The quinoxalino-substituent of the porphyrin fixes the inner nitrogen atom locant numbers to those indicated in green which follow the standard IUPAC numbering for porphyrins.[150] The locant numbers shown on the quinoxalino fragment are nonstandard but used for clarity. Regardless of the B(F)OB(F) configuration, all B$_2$OF$_2$pqx stereoisomers exhibit the C_1 symmetry point group and are chiral. The $\kappa^2 N^{21,22}$ boron atom is indicated by blue and the $\kappa^2 N^{23,24}$ boron atom indicated in red.

The dianion of **1** (1^{2-} or pqx^{2-}) exhibits the C_{2v} symmetry point group. A B(F)OB(F) group ligated within the central cavity of 1^{2-}, regardless of its configuration, leads to structures of the general formula "B$_2$OF$_2$pqx" (see Figure 4.3) all of which exhibit the C_1 symmetry point group. Subsequently, all B$_2$OF$_2$pqx stereoisomers are chiral. For the *transoid* B(F)OB(F) configurations, the unique in-plane boron can be located at either the nomenclaturally defined $\kappa^2 N^{21,22}$ boron atom position, as indicated by blue in Figure 4.3, or at the nomenclaturally defined $\kappa^2 N^{23,24}$ boron atom position, as indicated in red. Concomitantly and in each case, the unique out-of-plane boron atom of a *transoid* configuration will occupy the other position. Due to chirality, these two possibilities result in a total of four *transoid* B$_2$OF$_2$pqx stereoisomers.

In the case of *cisoid* B(F)OB(F) configurations, Figure 4.2 shows that there is a pair of enantiomeric extended-*cisoid* B$_2$OF$_2$pqx stereoisomers and a pair of contracted-*cisoid* B$_2$OF$_2$pqx stereoisomers for a total of four *cisoid* B$_2$OF$_2$pqx stereoisomers.

4.3.3 *Primary conclusion for scope of structural possibilities*

There are eight B$_2$OF$_2$pqx stereoisomeric possibilities in total. Figure 4.4 shows all eight of these stereoisomers as potential reaction products from the synthetic scheme indicated.[7] Compounds **2a**, **2b**, **3a**, and **3b** have *transoid* configurations, compounds **4a** and **4b** have extended-*cisoid* configurations, and compounds **5a** and **5b** have the contracted-*cisoid* configuration. Equilibrium reaction pathways are marked that are consistent with the primary experimental observation that enantiomerisation is not possible under ambient conditions for

84

simple solutions. **E_File_2** shows stereoisomers **2a**, **2b**, **3a**, and **3b** rotating together to make their geometric relationships clearer. Similarly, **E_File_3** shows **4a**, **4b**, **5a**, and **5b**.
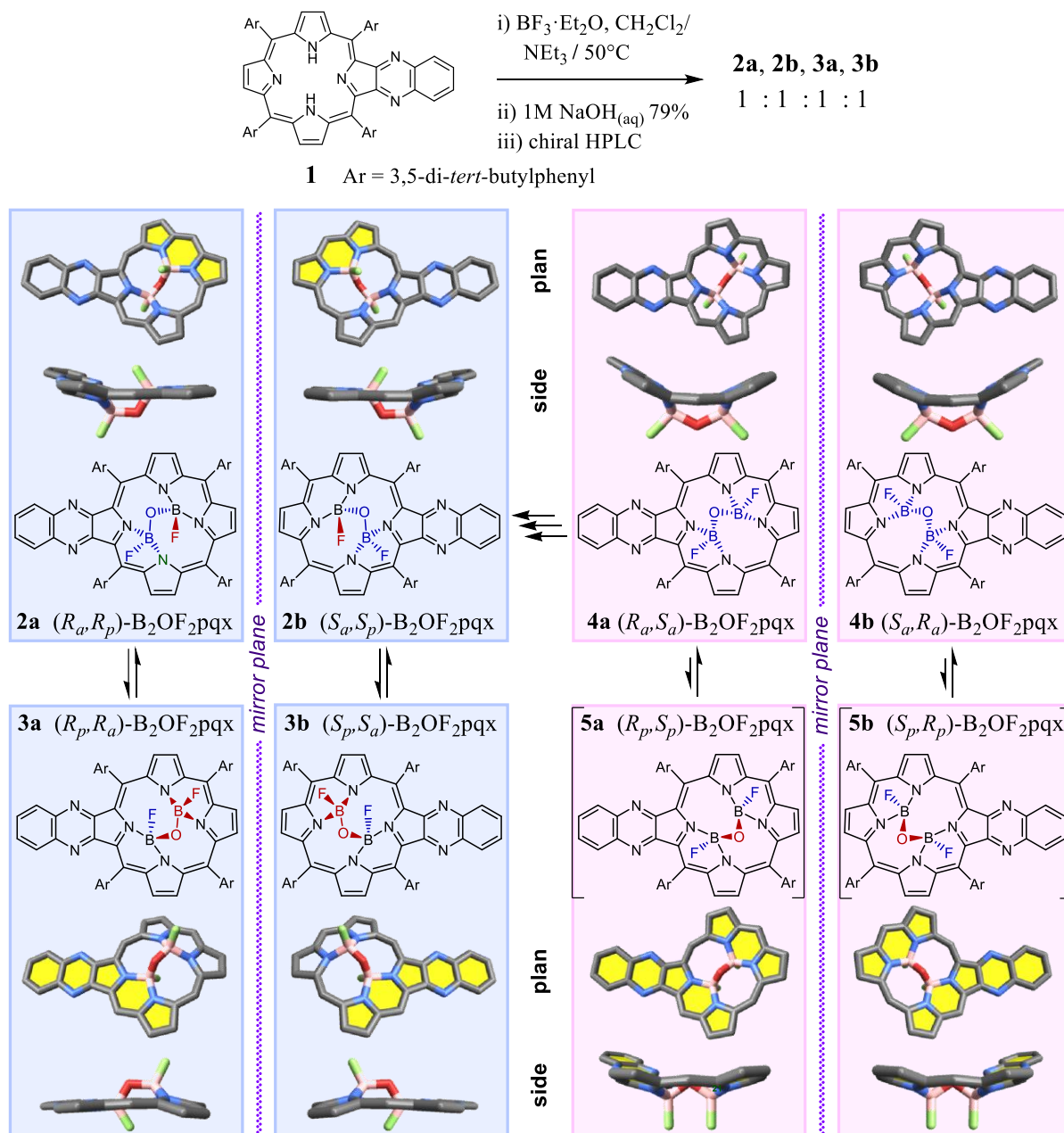
Figure 4.4 Synthetic scheme, structures and nomenclature of the eight possible stereoisomers produced by reaction of quinoxalinoporphyrin **1**. (a) Structure of **1** and reaction conditions. (b) Only **2a**, **2b**, **3a**, and **3b** (blue boxes) are isolated from the reported[7] synthesis, with DFT calculations suggesting **4a** and **4b** being much less stable thermodynamically, whilst **5a** and **5b** are unstable (pink boxes). Two dimensional representations of each structure are given. Enantiomers, for example, **2a** and **2b**, are related by reflection across the vertical mirror planes, as indicated, whilst akamptisomers, such as **2a** and **3a**, are related by BAR, moving the O through the effective plane of the macrocycle. All indicated stereochemistry is relative to the porphyrin ligand. The new stereodescriptors *parvo* (subscripted *p*) and *amplo* (subscripted *a*) are defined in Section 4.4.4. Atoms and bonds marked in red are above the quasi-plane of the porphyrin whilst those in blue are below it. Hydrogen atoms and the aryl side chains are omitted for clarity, with rings coplanar to a chelated *parvo* boron atom shaded yellow. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

4.3.4 Overview of isomerism and isomerisation as predicted by DFT

Free energies for model systems with *meso* Ar = H and Ar = 3,5-(di-*tert*-butyl)phenyl of B_2OF_2-pqx stereoisomers **2**, **3**, **4**, and **5** were calculated using DFT (see Methods Section 4.7.1.1). These results, as well as associated potential-energy surfaces, are indicated in Figure 4.5. These show that the *transoid* compounds **2** and **3**, within the precision of the method, are isoenergetic with $\Delta G = 0.0 \pm 0.2$ kJ mol^{-1} and have the lowest free energies of the different B(F)OB(F) configurations. The extended-*cisoid* compounds **4** are higher in energy by 30 kJ mol^{-1} and the contracted-*cisoid* compounds **5** is unstable undergoing barrierless conversion to **4**.

Three stereoisomerisation mechanisms "deemed" as feasible in simple organic-solvent solutions were considered and modelled using DFT. These mechanisms are the unimolecular BAR mechanism (see Sections 4.5.1 and 4.5.2), a unimolecular bond-walk mechanism called strepsisomerisation (see Section 2.4.26 for the formal definition and Section 4.5.3 for details specific to this Chapter), and a bimolecular process involving loss, rearrangement and subsequent addition of fluoride anion (see Section 4.5.4). These calculated reaction activation free energies are summarised in summarised Table 4.1. Only one reaction mechanism appropriate for the experimental conditions is envisaged, the akamptisomerisation reaction that occurs by a BAR $R_{st}^c 1$ over a calculated free-energy barrier of height 108 kJ mol^{-1}. This mechanism is depicted by the potential-energy surfaces shown in Figure 4.5 as a function of the B–O–B bond angle as the reaction coordinate. As the calculations do not indicate the feasibility of racemisation occurring, the results presented in that figure pertain to pairs of BAR-related stereoisomers only.

Table 4.1 Observed and B3LYP/6-31G+(d)-calculated reaction activation free energies in kJ mol^{-1}, for various types of isomerisation mechanisms in simple CH_2Cl_2 solutions. [a]

Environment	Process	Akamptisomerisation		Enantiomerisation	
		2a ⇌ 3a **3b ⇌ 2b**	**3a ⇌ 2a** **2b ⇌ 3b**	**2a ⇌ 2b**	**3a ⇌ 3b**
	Observed	*104 ± 2*	*104 ± 2*	N.O.[c]	N.O.[c]
	B−O−B BAR	108	108	N.A.[d]	N.A.[d]
CH_2Cl_2 solution	B(F)OB(F) bond-walk-rotation[b]	225	225	251	247
	F$^-$ loss[a]	284	285	220	220

[a] see Section 4.5.1, 4.5.3, and 4.5.4 for details, including pathways and considered alternatives.
[b] Strepsisomerisation – see Section 4.5.3.
[c] N.O. – Not observed. [d] N.A. – Not applicable.
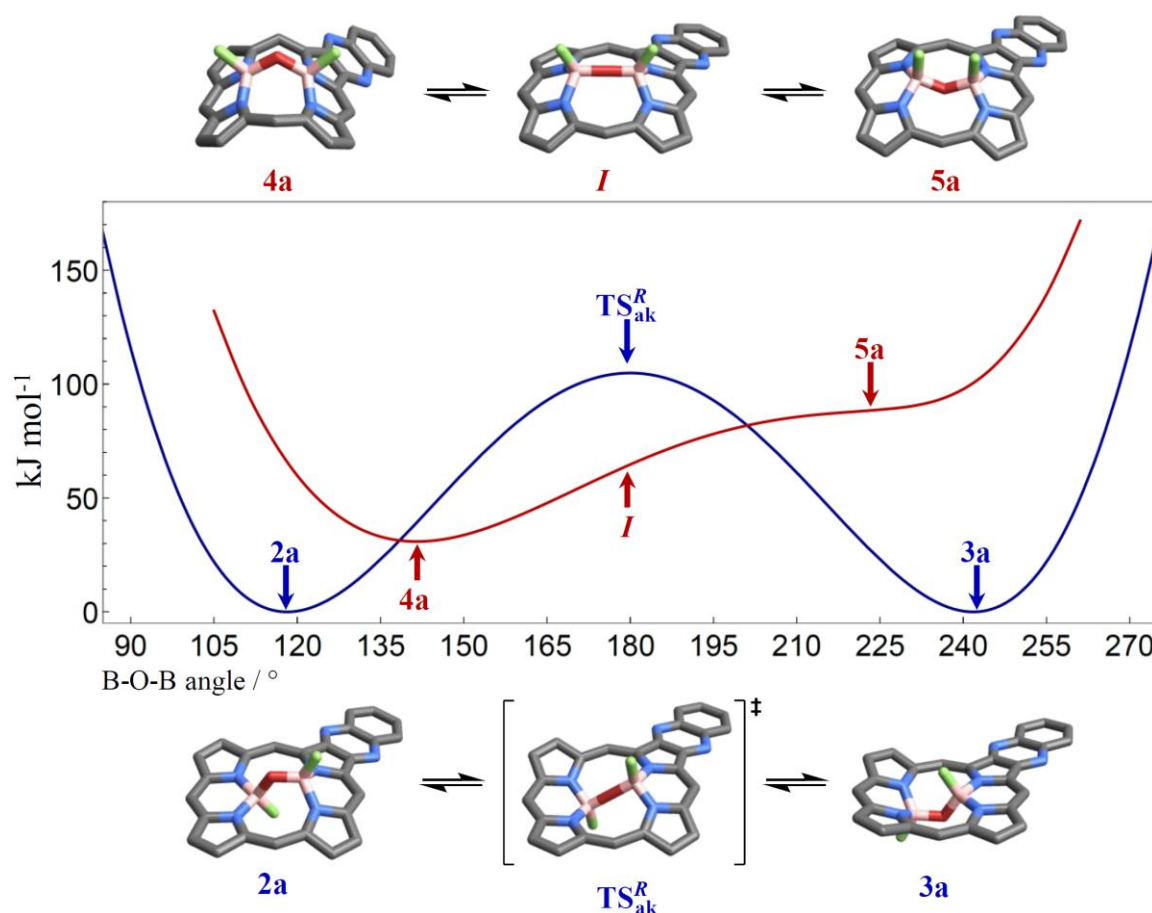
Figure 4.5 Energetics of BAR $R^c_{st}1$ mediated akamptisomerisation. DFT calculations show the metastable decay of the higher-energy contracted-*cisoid* compound **5a** to the lower energy extended-*cisoid* species **4a** through an intermediate *I* (shown in red, corresponding to the top structures) and the equilibrium over a BAR transition state TS^R_{ak} between *transoid* compounds **2a** and **3a** (shown in blue, bottom structures). The TS^R_{ak} structure features a linear B−O−B bond angle. The *meso* 3,5-di-*tert*-butylphenyl groups and hydrogens are omitted for clarity. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

Recapping the relevant prior experimental findings, under the conditions of the synthesis, the reaction products have been assumed to be under thermodynamic control. In the solid form, all four isolated fractions were found to be indefinitely stable at ambient temperatures. In solution (either CDCl₃ as used for NMR measurements or 99:1 *n*-hexane:2-propanol as used for the chiral-stationary phase HPLC resolution of stereoisomers), mild temperatures induced the specific diastereomeric interconversions as shown in Figure 4.1c. The rate constants for these processes as determined[7] using ¹⁹F NMR, led to the conclusion that the free-energy changes associated with each reaction are $\Delta G = 0.0 \pm 0.2$ kJ mol⁻¹, while the activation energies are $\Delta G^{\ddagger} = 104 \pm 2$ kJ mol⁻¹. No racemisation was observed under these conditions. The DFT calculations predict all of these critical features, indicating the likelihood of four equal-proportion or near-equal-proportion, near-equal energy products, a single reaction mechanism with a barrier of 108 kJ mol⁻¹, and no low-energy racemisation process.

Primary conclusion of the overview of isomerism and isomerisation

as predicted by DFT:

Only compounds **2a**, **2b**, **3a**, and **3b** need to be considered when assigning the isolated fractions **Fr1**, **Fr2**, **Fr3**, and **Fr4** to stereoisomeric structures.

4.3.5 DFT calculations used to assign the observed NMR spectra

4.3.5.1 <u>NMR structural assignments</u>

The experimental[7] ^{1}H, ^{13}C and ^{19}F chemical shifts and assignments for racemate (**Fr1** + **Fr2**) and racemate (**Fr3** + **Fr4**) based on homonuclear 2D NMR (dqf-COSY, NOESY) and heteronuclear 2D experiments (^{1}H-^{13}C HMBC, ^{1}H-^{13}C HSQC) are reproduced in Appendix D Table D.1, Table D.2, and Table D.3. The annotated 1D ^{1}H, ^{13}C and ^{19}F spectra of racemate (**Fr1** + **Fr2**) are reproduced in Appendix D Figure. D.2, Figure D.3, and Figure D.4, respectively and similarly for racemate (**Fr3** + **Fr4**) in Appendix D Figure D.5, Figure D.6, and Figure D.7.

Due to the conformational landscape and complex nature of the induced magnetic environments inherent in the B_2OF_2-pqx stereoisomers, DFT calculations of ^{1}H, ^{13}C and ^{19}F isotropic chemical shifts for the full systems (with Ar = 3,5-di-*tert*-butylphenyl; see Methods Section 4.7.1.5) were of additional utility in the interpretation of these experimental isotropic chemical shifts with the view of assigning structures to the isolated fractions **Fr1** – **Fr4**.

Chemical intuition suggested that the β-pyrrolic ^{1}H and aromatic ^{13}C centres, supported by the DFT NMR calculations, would be strongly diagnostic of structure (since enantiomers show the same NMR spectra and this method is not able to discriminate between **a** and **b**, but quite different spectra are predicted for **2** – **5**). The DFT-calculated isotropic β-pyrrolic ^{1}H, ^{19}F and ^{13}C chemical shifts for **2** and **3** corresponding to all experimentally assigned aromatic peaks (see Appendix D) of (**Fr1** + **Fr2**) and (**Fr1** + **Fr2**) are listed in Table 4.2Table 4.2 and with their experimental-calculated correlations shown in Figure 4.6. Good agreement is found only when **Fr1** and **Fr2** are assigned to racemate **2** and **Fr3** and **Fr4** are assigned to racemate **3**.

Table 4.2 DFT calculated β-pyrrolic ^{1}H, ^{19}F and ^{13}C isotropic chemical shifts corresponding to all experimentally assigned aromatic peaks for racemates **2** and **3**. Atom assignments match the locant numbers indicated in Figure 4.3.

Racemate **2** (as **2a**)		Racemate **3** (as **3a**)	
Position	**[1]H δ / ppm**	**Position**	**[1]H δ / ppm**
H7	9.12	H7	9.61
H8	9.06	H8	9.68
H12	9.06	H12	8.70
H13	9.54	H13	8.89
H17	9.82	H17	8.89
H18	10.03	H18	9.29
Position	**[19]F δ / ppm**	**Position**	**[19]F δ / ppm**
F^a	-175.0	F^a	-177.1
F^p	-203.9	F^p	-202.4
Position	**[13]C δ / ppm**	**Position**	**[13]C δ / ppm**
C5	134.43	C5	124.01
C6	154.84	C6	149.89
C7	127.04	C7	130.24
C8	142.43	C8	146.06
C9	154.97	C9	146.97
C10	135.26	C10	137.08
C11	133.25	C11	155.17
C12	146.00	C12	141.49
C13	133.25	C13	129.12
C14	147.69	C14	156.30
C15	129.63	C15	139.65
C16	145.53	C16	151.65
C17	131.94	C17	129.47
C18	144.82	C18	139.93
C19	152.33	C19	159.59
C20	129.95	C20	129.46
C26	138.29	C26	138.50
C27	135.53	C27	135.63
C28	136.01	C28	135.78
C29	137.48	C29	137.17
C25a	148.38	C25a	147.73
C29a	148.98	C29a	148.73
p-Ar5	109.00	p-Ar5	109.29
p-Ar10	110.50	p-Ar10	106.39
p-Ar15	108.64	p-Ar15	109.20
p-Ar20	105.22	p-Ar20	110.18
o-Ar5	118.05	o-Ar5	122.00
o-Ar5	119.68	o-Ar5	123.27
o-Ar10	126.64	o-Ar10	120.84
o-Ar10	127.92	o-Ar10	121.22
o-Ar15	117.92	o-Ar15	118.55
o-Ar15	122.55	o-Ar15	121.19
o-Ar20	125.29	o-Ar20	120.22
o-Ar20	128.38	o-Ar20	123.09
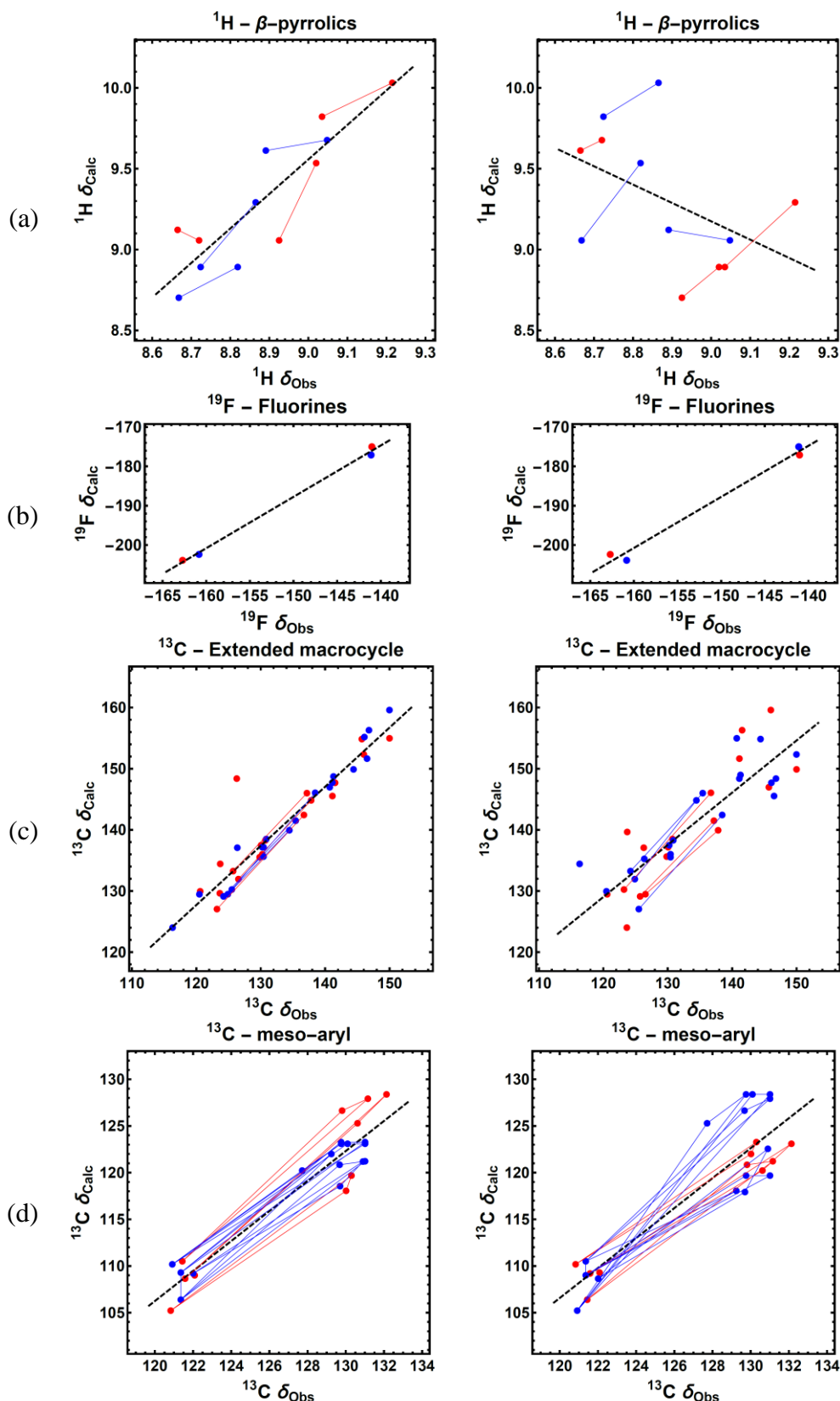
Figure 4.6 Correlations of experimental[7] and DFT calculated β-pyrrolic ¹H, ¹⁹F and aromatic ¹³C NMR isotropic chemical shifts, based on the possible alternative assignments of: Left- **Fr1** and **Fr2** to racemate **2**, and Right- **Fr1** and **Fr2** to racemate **3**. Red- data for molecules **2a** and **2b**, blue- data for molecules **3a** and **3b**; tie lines in (a) indicate vicinal β-pyrrolic proton pairs, (c) adjacent β-pyrrolic carbons and (d) the *para-* and two *ortho* carbons of individual *meso*-aryl groups. Each dashed line indicates a linear regression fit to the data points.

Primary conclusion of the NMR structural assignments:

(**Fr1** + **Fr2**) is assigned to racemate **2** and (**Fr3** + **Fr4**) is assigned to racemate **3**.

4.3.6 DFT calculations to assign the observed CD spectra

To determine chiral-structural details, DFT modelled properties were used to interpret the experimental[7] electronic transition chiroptical measurements. Electronic absorption and circular dichroism spectra were modelled using the optimised geometries for **2a** and **3a** only as their enantiomers give identical results except for the opposite sign of the calculated rotatory strengths. At least 8 electronic transitions were determined, the first two occurring in the Q-band region while the remainder are the Soret and nearby bands. See Methods Section 4.7.1.6 for technical details.

Results are given in Table 4.3 for **2a** and Table 4.4 for **3a**, featuring calculated excitation energies, electric-dipole oscillator strengths, rotatory strengths, the angle between the transition electric dipole vector and the direction made by projecting the B-B vector onto the plane best representing the four pyrrolic rings, and the "E-M" angle between the electric dipole moment and the magnetic dipole moment. The most noticeable common feature is that the E-M angles are all near 90°. Rotatory strengths are proportional to the cosine of this angle and hence are expected to be small and to be dominated by subtle effects that manipulate the perceived angle. This result reflects the geometrical property that chirality in these molecules is a subtle effect arising through the interplay of the distortions and asymmetries introduced by the *transoid* B(F)OB(F) group combined with the asymmetry introduced by mono-quinoxalino ring fusion.

Table 4.3 Calculated spectral properties of **2a**.

	method	CAM-B3LYP	CAM-B3LYP	CAM-B3LYP	CAM-B3LYP	ωB97XD	CAM-B3LYP	ωB97XD	CIS
	basis	6-31G*	6-31+G*	6-31+G*	6-31++G(2d,p)	6-31G*	6-31+G*	6-31+G*	6-31+G*
	solvent	none	none	CHCl$_3$	THF ε=7	none	none	none	none
	meso group[d]	dtp	dtp	dtp	dtp	dtp	H	H	H
0→1	Energy / cm^{-1}	16440	16250	16470	16120	16090	16930	16620	18370
	Osc. Str.	0.059	0.057	0.062	0.078	0.058	0.035	0.029	0.080
	Rot. Str.[a]	9	10	12	14	9	−9	−10	−40
	TM angle[b] / °	29	31	30		28	96	96	120
	E-M angle[c] / °	86	85	86		87	94	95	84
0→2	Energy / cm^{-1}	16740	16570	16790	16480	16460	17490	17140	19880
	Osc. Str.	0.009	0.011	0.012	0.024	0.009	0.009	0.010	0.085
	Rot. Str.[a]	0	1	1	2	1	−5	−4	−11
	TM angle[b] / °	7	6	13		10	63	109	57
	E-M angle[c] / °	87	88	87		86	100	98	105
0→3	Energy / cm^{-1}	24480	24110	24510	23380	24740	26100	26410	33050
	Osc. Str.	1.050	1.064	1.059	1.654	1.129	0.410	1.298	2.874
	Rot. Str.[a]	27	39	47	255	106	1	225	114
	TM angle[b] / °	86	93	83		73	97	35	16
	E-M angle[c] / °	86	85	83		70	90	29	92
0→4	Energy / cm^{-1}	24700	24320		23450	24950	26280	26660	34110
	Osc. Str.	1.255	1.314		1.635	1.241	1.399	0.707	0.886
	Rot. Str.[a]	−24	−20		−229	−99	110	−139	56
	TM angle[b] / °	10	10			162	34	30	48
	E-M angle[c] / °	91	91			96	81	100	73
0→5	Energy / cm^{-1}	26360	26270			26250	28540	28140	36310
	Osc. Str.	0.641	0.530			0.675	0.041	0.033	1.833
	Rot. Str.[a]	8	8			-6	−1	−6	−147
	TM angle[b] / °	138	137			38	142	144	62
	E-M angle[c] / °	88	87			93	94	102	91
0→6	Energy / cm^{-1}	28450	28210			28000	28920	28690	37640
	Osc. Str.	0.013	0.191			0.005	0.465	0.010	0.057
	Rot. Str.[a]	2	−13			3	−88	13	−17
	TM angle[b] / °	53	41			37	119	150	96
	E-M angle[c] / °	82	93			73	98	55	135
0→7	Energy / cm^{-1}		28710			29470	29260	29550	38340
	Osc. Str.		0.043			0.112	0.025	0.339	0.125
	Rot. Str.[a]		25			11	−6	−64	49
	TM angle[b] / °		145			135	47	61	57
	E-M angle[c] / °		47			88	92	98	94
0→8	Energy / cm^{-1}		29280			30060	30670	31030	40050
	Osc. Str.		0.004			0.036	0.058	0.017	0.109
	Rot. Str.[a]		2			25	53	34	-8
	TM angle[b] / °		140			139	35	50	79
	E-M angle[c] / °		77			65	73	78	122

a: 10^{-40} erg esu cm gauss^{-1}. b: angle between electronic transition moment and the B–B vector projected into the macrocyclic plane. c: angle between the electric dipole and magnetic dipole transition moments. d: dtp = 3,5-di-*tert*-butylphenyl.

Table 4.4 Calculated spectral properties of **3a**.

	method	CAM-B3LYP	CAM-B3LYP	CAM-B3LYP	CAM-B3LYP	ωB97XD	CAM-B3LYP	ωB97XD	CIS
	basis	6-31G*	6-31+G*	6-31+G*	6-31++G(2d,p)	6-31G*	6-31+G*	6-31+G*	6-31+G*
	solvent	none	none	CHCl$_3$	THF ε=7	none	none	none	none
	meso group[d]	dtp	dtp	dtp	dtp	dtp	H	H	H
0→1	Energy / cm^{-1}	16880	16690		16560	16620	17150	16900	18530
	Osc. Str.	0.006	0.008		0.017	0.005	0.042	0.039	0.113
	Rot. Str.[a]	9	10		14	9	14	13	−7
	TM angle[b] / °	83	95			82	51	51	49
	E-M angle[c] / °	85	84			85	87	87	89
0→2	Energy / cm^{-1}	17180	17030		16960	17040	18120	17960	22620
	Osc. Str.	0.064	0.063		0.098	0.065	0.001	0.001	0.050
	Rot. Str.[a]	−6	−7		−13	−6	1	1	−3
	TM angle[b] / °	24	25			25	27	45	45
	E-M angle[c] / °	99	101			100	77	71	48
0→3	Energy / cm^{-1}	24620	24240		23530	24870	26130	26350	32970
	Osc. Str.	1.143	1.164		1.701	1.226	1.147	1.252	3.032
	Rot. Str.[a]	−32	−34		−129	−60	−116	−141	−165
	TM angle[b] / °	76	74			107	13	14	19
	E-M angle[c] / °	98	98			105	105	107	117
0→4	Energy / cm^{-1}	25240	24840		24040	25600	26660	27270	35430
	Osc. Str.	1.413	1.392		1.727	1.505	0.609	0.775	0.982
	Rot. Str.[a]	58	52		122	94	13	59	−26
	TM angle[b] / °	9	9			169	106	99	88
	E-M angle[c] / °	85	85			83	88	85	70
0→5	Energy / cm^{-1}	26720	26700			26580	28960	28690	38100
	Osc. Str.	0.266	0.238			0.259	0.132	0.009	0.007
	Rot. Str.[a]	−54	−61			−56	72	7	153
	TM angle[b] / °	125	124			52	92	46	86
	E-M angle[c] / °	105	109			107	75	60	62
0→6	Energy / cm^{-1}	28480	28260			28240	29100	28780	37640
	Osc. Str.	0.044	0.199			0.001	0.009	0.004	0.057
	Rot. Str.[a]	25	29			1	20	23	-1
	TM angle[b] / °	14	12			121	35	97	104
	E-M angle[c] / °	79	84			76	29	5	88
0→7	Energy / cm^{-1}					29350	29650	29800	38760
	Osc. Str.					0.047	0.288	0.066	0.050
	Rot. Str.[a]					11	-54	-4	2
	TM angle[b] / °					150	102	86	126
	E-M angle[c] / °					89	99	94	79
0→8	Energy / cm^{-1}					29890	30260	30630	38900
	Osc. Str.					0.065	0.096	0.219	0.326
	Rot. Str.[a]					−40	−11	−39	−93
	TM angle[b] / °					139	154	114	113
	E-M angle[c] / °					110	96	127	92

a: 10^{-40} erg esu cm gauss^{-1}. b: angle between electronic transition moment and the B–B vector projected into the macrocyclic plane. c: angle between the electric dipole and magnetic dipole transition moments. d: dtp = 3,5-di-*tert*-butylphenyl.

The results of analysis, in combination of with the pattern of thermal isomerisation and NMR assignments, was used to assign absolute configurations **a** and **b** for each of **2** and **3** with the results summarised in Table 4.5 with **Fr1** assigned to **2b**, **Fr2** to **2a**, **Fr3** to **3a**, and **Fr4** to **3b**. There are only two feasible possibilities for the assignments of the four fractions, and only one of these results is in qualitative agreement between the observed and calculated CD spectroscopic properties.

Table 4.5 Assignment of structures to HPLC fractions by comparison of measured[7] and calculated ^{1}H and ^{19}F NMR, observed thermal isomerisation processes, and CD Spectra.

Fraction	NMR	Thermal isomerisation		Q-band intense peak~15600 cm^{-1}			Soret shoulder at ~22500 cm^{-1}		
		possibility 1	possibility 2	CD sign	CD strength	Ass.	CD sign	CD strength	Ass.
Fr1	**2**	**a**	**b**	–	strong	**2b**	+	weak	**2b**?
Fr2	**2**	**b**	**a**	+	strong	**2a**	-	weak	**2a**?
Fr3	**3**	**b**	**a**	–	weak	**3a**	-	strong	**3a**
Fr4	**3**	**a**	**b**	+	weak	**3b**	+	strong	**3b**

Primary conclusion of the structural assignments from
DFT interpreted chiroptical properties:

Fr1 is assigned to **2b**, **Fr2** to **2a**, **Fr3** to **3a**, and **Fr4** to **3b.**

4.4 Necessary new terminology and nomenclature

4.4.1 Terminology

New terminology is required to describe the observed stereoisomerism, *i.e.*, the geometric relationships between these stereoisomers, and the associated concerted unimolecular stereoisomerisation process ($R^c_{st}1$) that this work shows interconverts between the diastereomeric pairs (**2a** ⇌ **3a**), (**2b** ⇌ **3b**), and conceptually between diastereomeric pairs (**4a** ⇌ **5a**) and (**4b** ⇌ **5b**).

Whilst in Chapter 2, it was pointed out that stereoisomerism classifications should only describe geometric relationships, during the early stages of the work comprising this Thesis, common practice was followed where, for some types of conformational isomerism,

the terms mix their geometric relationship definition with reference to a real or conceptual mechanistic process.

The IUPAC definition for atropisomers (from Greek *atropos* meaning "without turning") meaning "*subclass of conformers which can be isolated as separate chemical species and which arise from restricted rotation about a single bond*" (see Section A.1) provided the template for the BAR analogue described here. Hence the term *akamptisomers* (from Greek *akamptos* meaning "inflexible, unbending", see Section 2.4.3) was coined to describe bond-angle reflectomers where the barrier to BAR was sufficiently high, or "hindered", to allow for practical isolation of the resultant stereoisomers. Further, the relationship between these akamptisomers was called *akamptisomerism* (see Section 2.4.1) in keeping with the IUPAC standard general definition for isomerism (see Section A.25). Similarly, any mechanism that can convert between akamptisomer pairs is called *akamptisomerisation* (see Section 2.4.2) with the associated $R_{st}^{c}1$ process being BAR.

4.4.2   The ML$_2$ stereogenic unit and their 2D depiction

This work revealed that an embedded ML$_2$ stereogenic unit can exhibit geometric relationships unforeseen by the existing IUPAC recommendations.[91] This is apparent when the ML$_2$ unit is embedded so as to provide an external frame of reference. Two limiting cases are shown in Figure 4.7. In Figure 4.7a is an example of a substituted 11-oxabicyclo(4.4.1)undeca-1,3,5,7,9-pentaene where the bridgehead L atoms are sp^2 carbon atoms and all bonds connecting to them form part of the embedding structure. Here, the existing IUPAC recommendations for the wedge-hash depiction of stereoisomeric relationships[91] and the IUPAC inorganic nomenclatural polyhedral symbols[3] prove sufficient to describe all the geometric relationships.
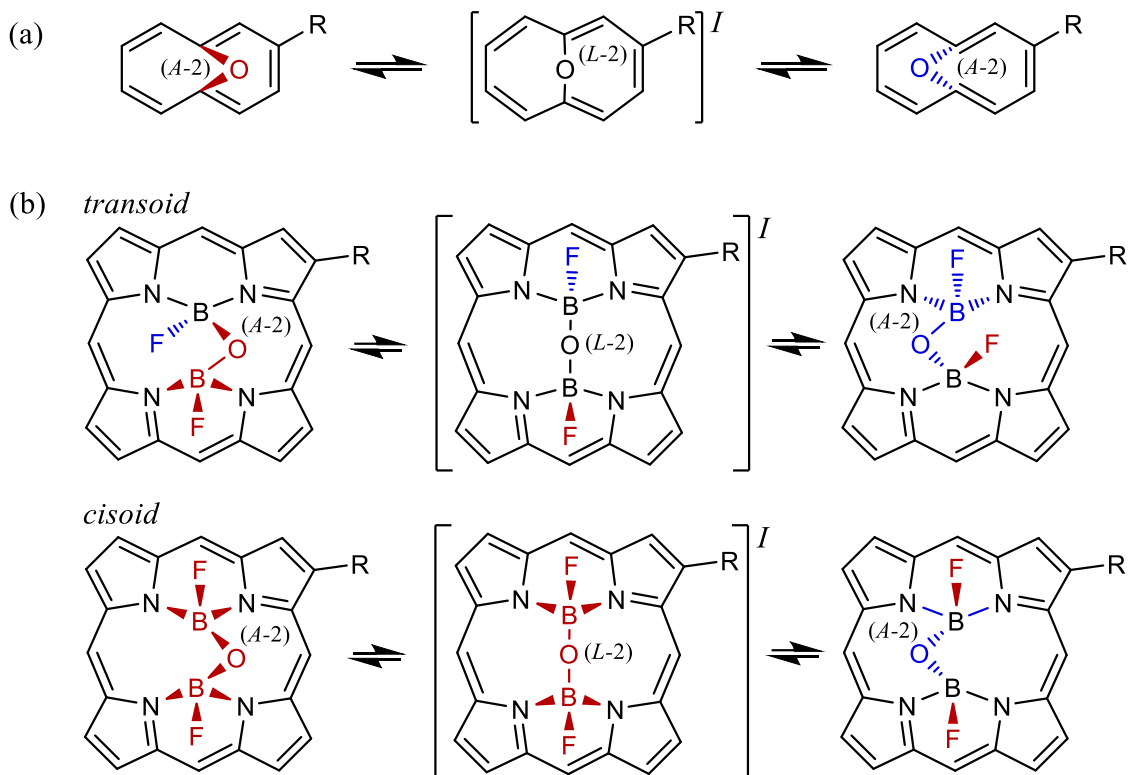
Figure 4.7 Usage of the wedge-hash system for the 2D depiction of the ML$_2$ stereogenic unit of BAR-related stereoisomers. (a) For a substituted 11-oxabicyclo(4.4.1)undeca-1,3,5,7,9-pentaene, the existing wedge-hash system[91] and inorganic polyhedral symbols[3] suffice to describe all stereoisomeric properties of the ML$_2$ stereogenic unit. (b) For the substituted porphyrin ligated B(F)OB(F) complexes the existing recommendation for the wedge-hash system fails. The non-standard usage shown here faithfully indicates the geometric relationships between M, L and the embedding porphyrin. Significantly, it is the presence of the non-embedded fluorine atoms that demands this. The intermediates marked as "*I*" need not be characterised as a transition-state structure. Atoms and bonds marked in red are above the quasi-plane of the porphyrin whilst those in blue are below it.

In Figure 4.7b are examples of substituted porphyrin ligated B(F)OB(F) complexes. In these cases, the bridgehead L atoms are sp^3 tetrahedral boron atoms that each feature a substituent that is not part of the embedding structure (the fluorine atoms). This work reveals that in such cases the existing wedge-hash system fails and that the entire ML$_2$ unit and its connections to the embedding structure *need to be considered as a single stereogenic unit*, requiring the non-standard wedge-hash usage as shown. This system, used throughout this Thesis, faithfully indicates all the necessary 3D geometric relationships between M, L and the embedding structure. A detailed examination of the implications that arise from the failings of standard IUPAC depictions of stereogenic units is discussed in Chapter 5.

4.4.3 *Deficits of current nomenclature for akamptisomeric systems*

The unusual nature, recent discovery, and (currently small) number of examples of compounds in this class has highlighted weaknesses, and indeed failings, of existing nomenclature. The nomenclature used in Chemical Abstracts (CAS) listings is ambiguous as it does not fully address the stereochemistry of these compounds. For example, B(F)OB(F) porphyrins are currently listed in CAS using additive systematic IUPAC nomenclature, *i.e.*, for the compound with CAS Registry Number 1022915-61-2, the name is difluoro-μ-oxo[μ-[tetraphenyl-21H,23H-porphinato(2-)-κN^{21},κN^{22}:κN^{23},κN^{24}]]diboron. In a similar fashion, the B(F)OB(F)-corrole complex, assigned CAS Registry Number 1013891-91-2 (Component: 1013965-96-2), is named [μ-[1,2,3,7,8,12,13,17,18,19-decadehydro-21,22-dihydro-5,10,15-tris(4-methylphenyl)corrinato(3-)-κN^{21},κN^{24}:κN^{22},κN^{23}]]difluoro-μ-oxodiborate(1-). However, from the primary publication, the first is a *transoid* compound[143] and the second an extended-*cisoid* compound[146], and the alternative forms of each are, in principle, feasible yet would be assigned the *same* name:

CAS number: 1022915-61-2	is (*parvo,amplo*)	but not (*amplo,amplo*)
H⁺ Ar = 4-methylphenyl CAS Number 1013891-91-2 (Component: 1013965-96-2)	is H⁺ Ar = 4-methylphenyl (*amplo,amplo*)	but not H⁺ Ar = 4-methylphenyl (*parvo,amplo*)
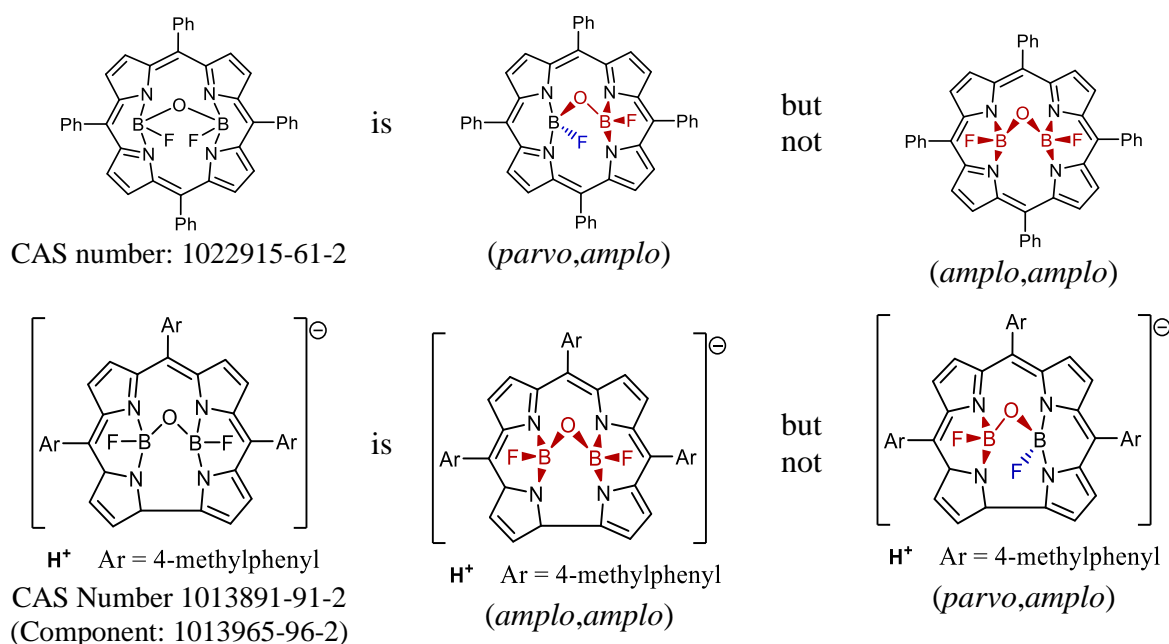
The most obvious difference between the two B(F)OB(F) configurations demonstrated are the positions of the boron atoms relative to the macrocycle. There does exist a system of α/β stereodescriptors for defining the different faces of a macrocycle[150] which could, in principle, be used to indicate the single protruding BF in the transoid case, or the two co-facially protruding BF groups in the cisoid case. This usage, however, is

somewhat lacking in that it introduces a new source of ambiguity in the case of porphyrinoid macrocycles as "β" is frequently used in a non-systematic way to indicate the β-pyrrolic positions.

Porphyrin complexes featuring the unsymmetrical bridge group B(OH)OB(Ph) prepared by Belcher *et al*.[143] demonstrate another shortcoming of existing nomenclature. Under the assumption of the bridge group adopting a *transoid* configuration, there is the question as to which, of the now distinct, boron atoms occupies the position that is coplanar with the macrocycle. This is demonstrated below for CAS Registry Number 1022915-75-8, which has been assigned the name hydroxy-μ-oxophenyl[μ-[5,10,15,20-tetrakis(4-methylphenyl)-21H,23H-porphinato(2-)-κN^{21},κN^{22}:κN^{23},κN^{24}]]diboron.

Ar = 4-methylphenyl	Ar = 4-methylphenyl	Ar = 4-methylphenyl
CAS number: 1022915-75-8	(*parvo,amplo*)	(*amplo,parvo*)
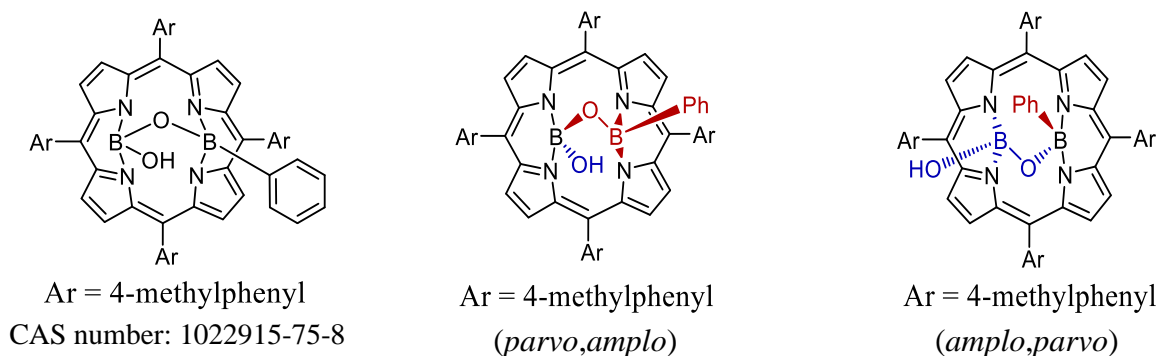
The low-symmetry B(F)OB(F) porphyrin isomers studied in this work further exemplify the shortcomings of existing nomenclature, especially when discussing the phenomenon of akamptisomerism. An ideal amendment to the systematic nomenclature would not only be to provide clear, concise, and unambiguous *names* but also make the *relationships* between stereoisomers clear and, preferably, obvious.

Additionally, no standard IUPAC descriptors[3] exist for akamptisomer nomenclature, although for some of these molecules existing descriptors could be utilised. The Klyne-Prelog system provides a crude description of torsional relationships, but it is not currently utilised within an IUPAC *nomenclatural* context, nor is it ideally fit for purpose here.

4.4.4   *Akamptisomer specific stereodescriptors: parvo and amplo*

Proposed here are universal stereodescriptors for bond-angle reflectomer and akamptisomer nomenclature, enabling easy implementation within computerised naming

schemes. These provide general molecular structural elements that are intuitively descriptive and can also be applied to akamptisomeric systems involving double bonds, if required.

The essential features of akamptisomers in a constraining (embedding) environment are:

(i)   angle bending around a bicoordinate atom M linked to atoms L^1 and L^2, forming an inner L^1-M-L^2 unit [for example, B(F)OB(F) for **2 – 5**], and

(ii)   this unit is constrained by bonds to its external embedding environment (macrocyclic ligand, polymer chain, matrix, *etc.*). See Chapter 3 Section 3.5.2.

Both L^1 and L^2 may connect to their embedding environment by more than one bond (*e.g.,* two B−N bonds each for **2 – 5** but with the fluorine atoms ignored as these are not connected to the embedding environment). The simplest representation for this is a single vector that averages all of the actual bond vectors to L^1. These vectors end at points labelled E^1 and E^2 that originate from L^1 and L^2, respectively (see Figure 4.8). Typically, E^1 and E^2 will lie very close to the L^1-M-L^2 plane, but, for generality, E^1 and E^2 are projected onto this plane. For porphyrins and other systems that exhibit a local reflection symmetry plane coincident with the L^1-M-L^2 plane (see Figure 4.2a–c, Figure 4.2e, and Figure 4.2f), E^1 and E^2 will lie very close to this L^1-M-L^2 plane. As a counter-example where this local reflection symmetry is not present such as in corroles (Figure 4.2d), the effect of the E^1 projection onto the L^1-M-L^2 plane is more significant.

To determine the configuration of L^i with respect to its external environment, the $L^i-E^i-E^j$ angle θ_i is calculated. This is then compared to the analogous angle θ_i' in the BAR stereoisomer. Here, an arbitrary numerical criterion is chosen for comparison of the angles; by testing against many B−O−B macrocycles complexes, including corroles (see Table 4.8), it was found that comparing the θ_i / θ_i' values to 2/3 and 3/2 appear to be a robust criterion. Accordingly, the stereodescriptors are defined as follows:

$$\theta_i\big/\theta_i' < {}^2\!/_3 \qquad\qquad L^i \text{ is assigned the stereodescriptor } parvo$$

$${}^2\!/_3 \leq \theta_i\big/\theta_i' \leq {}^3\!/_2 \qquad\qquad \text{no stereodescriptor applies}$$

$${}^3\!/_2 < \theta_i\big/\theta_i' \qquad\qquad L^i \text{ is assigned the stereodescriptor } amplo$$

with

> *parvo*: Latin; ablative, neuter, singular of *parvus*, meaning "small".
>
> *amplo*: Latin; ablative, neuter, singular of *amplus*, meaning "large, spacious".

A Mathematica script that automates the assignment of *parvo*/*amplo* stereodescriptors is provided in Appendix C.

Figure 4.8 shows the application of this scheme to **2a** – **5a**, also the BAR transition state $\mathbf{TS_{ak}^{R}}$ connecting **2a** with **3a,** and a defluorinated reaction intermediate, illustrating that this method leads to unique chemical names. When L^1 and L^2 are stereocentres, they are given their respective absolute configuration stereodescriptors to which the right subscripted labels "*amplo*" and "*parvo*" are added. If using standard IUPAC organic nomenclatural stereodescriptors, these will be either (*R*) or (*S*). If using standard IUPAC inorganic polyhedral absolute configuration stereodescriptors, these will be either (*T*-4-*R*) or (*T*-4-*S*). The labels for each atom L^i in the name are ordered according to the standard numbering system of the surrounding macrocycle. If there exists a choice for numbering, *parvo* is given priority over *amplo*. Abbreviated names are also provided in the figure. All **b** isomers are named similar to the **a** isomers shown, with all (*R*) and (*S*) labels interchanged. In Table 4.8, the names and angles are given for the 36 porphyrin and corrole structures and transition states mentioned in this work.

Compounds like **2** – **5** form two bonds from each of L^1 and L^2, to their environments, within a macroscopic ring system. Then (*parvo*) configurations generally correspond to the ligated centre being effectively co-planar ("*plano*") with its ligated half of the macrocycle, with (*amplo*) centres being distinctly out-of-plane ("*aplano*") with their ligated half of the macrocycle, thus the mnemonic:

$$\boldsymbol{p}arvo \approx \boldsymbol{p}lano \quad \text{and} \quad \boldsymbol{a}mplo \approx \boldsymbol{a}plano$$

can provide a more structure-intuitive guide.

The situation in which the L^1−M−L^2 angle is nearly linear is worthy of note. In this case, optimisation of the structure produced by BAR inversion usually results in a return to the original structure, making both θ_i / θ_i' = 1 and so no stereodescriptor is assigned. This result is both general and intuitive.

In the unusual cases of the 1,3-fluoride shift enantiomerisation transition structures, both the fluorine *and* the oxygen atoms occupy positions on opposite sides of the macrocycle, and each is bicoordinate with a bent geometry. Whilst it is physically impossible for both the B−O−B and B−F−B bond angles to undergo a R_{st}^c1 BAR, this simultaneous-dual reflection geometric relationship still holds. The outcome is that $\mathbf{TS}_{en}^{RS+}$ and $\mathbf{TS}_{en}^{SR+}$ constitute a pair of akamptisomers. The dual BAR in this case also interchanges the (*T*-4-*R*) and (*T*-4-*S*) polyhedral absolute-configuration stereodescriptors; something not seen in the *single* BAR cases.

4.4.4.1 <u>Nomenclature for transition-state structures encountered in this Chapter</u>

With the increasing ability to accurately model transition structures for chemical reactions, a natural consequence is the increasing need to unambiguously name these. Whilst a comprehensive treatment solving the full range of challenges faced for a nomenclatural system for transition-state structures in general requires further development, the challenges encountered in this work are far simpler.

In this Chapter all the transition structures differ from potential-energy local-minima structures only by either coordination geometry (in the case of linear bicoordinate oxygen) or coordination number (in the case of strepsisomerism). Consequently, these transition structures are readily described by implementing the existing standard IUPAC inorganic nomenclature (IUPAC "Red Book"[3]). The designation that a constructed name is a transition structure and not a potential-energy local minimum is readily indicated by simply adding the double dagger (‡) symbol to the end of the name. For example, the B−O−B BAR transition structures encountered here are characterised by a linear or near-linear bicoordinate oxygen, hence necessitating the use of the (*L*-2) symbol (see IR-9.3.2.1[3] and Appendix B).

For abbreviated names, the double dagger symbol can be used as is or superscripted as desired.
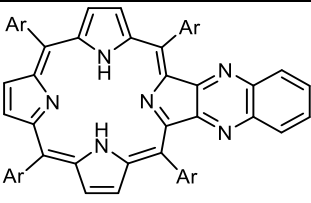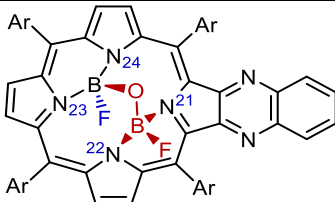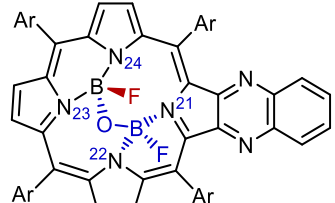
Schematics

Name	Annotated model	Isomer to name	BAR isomer	R_i assignment

2a
(R_a, R_p)-
B_2OF_2pqx

L^1 *amplo*
L^2 *parvo*

$\ldots 1\kappa^2 N^{21,22}{:}3\kappa^2 N^{23,24}\text{-}\{[1(T\text{-}4\text{-}R_{amplo}),3(T\text{-}4\text{-}R_{parvo})]\text{-}1,3\text{-difluoro-}1\lambda^4,3\lambda^4\text{-diboroxane}\}$

3a
(R_p, R_a)-
B_2OF_2pqx

L^1 *parvo*
L^2 *amplo*

$\ldots 1\kappa^2 N^{21,22}{:}3\kappa^2 N^{23,24}\text{-}\{[1(T\text{-}4\text{-}R_{parvo}),2(A\text{-}2),3(T\text{-}4\text{-}R_{amplo})]\text{-}1,3\text{-difluoro-}1\lambda^4,3\lambda^4\text{-diboroxane}\}$

4a
(R_a, S_a)-
B_2OF_2pqx

L^1 *amplo*
L^2 *amplo*

$\ldots 1\kappa^2 N^{21,22}{:}3\kappa^2 N^{23,24}\text{-}\{[1(T\text{-}4\text{-}R_{amplo}),2(A\text{-}2),3(T\text{-}4\text{-}S_{amplo})]\text{-}1,3\text{-difluoro-}1\lambda^4,3\lambda^4\text{-diboroxane}\}$

5a
(R_p, S_p)-
B_2OF_2pqx

L^1 *parvo*
L^2 *parvo*

$\ldots 1\kappa^2 N^{21,22}{:}3\kappa^2 N^{23,24}\text{-}\{[1(T\text{-}4\text{-}R_{parvo}),2(A\text{-}2),3(T\text{-}4\text{-}S_{parvo})]\text{-}1,3\text{-difluoro-}1\lambda^4,3\lambda^4\text{-diboroxane}\}$

TS_{ak}^R

L^1 *unlabelled*
L^2 *unlabelled*

$\ldots 1\kappa^2 N^{21,22}{:}3\kappa^2 N^{23,24}\text{-}\{[1\alpha(T\text{-}4\text{-}R),2(L\text{-}2),3\beta(T\text{-}4\text{-}R)]\text{-}1,3\text{-difluoro-}1\lambda^4,3\lambda^4\text{-diboroxane}\}\ddagger$

$(S_p, \_)$-
B_2OFpqx^+

L^1 *parvo*
L^2 *unlabelled*

$\ldots 1\kappa^2 N^{21,22}{:}3\kappa^2 N^{23,24}\text{-}\{[1(T\text{-}4\text{-}R_{parvo}),2(A\text{-}2),3(TP\text{-}3)]\text{-}1\text{-fluoro-}1\lambda^4\text{-diboroxanium}\}$
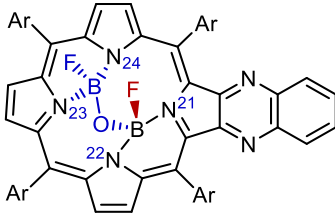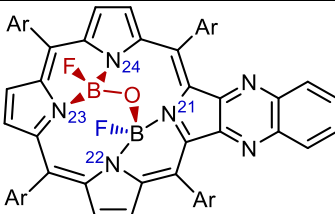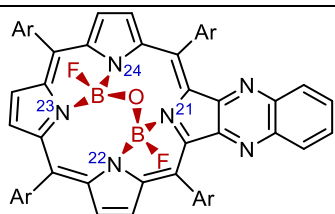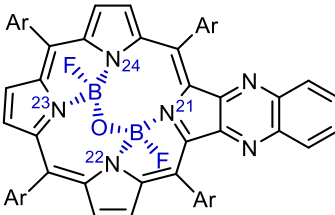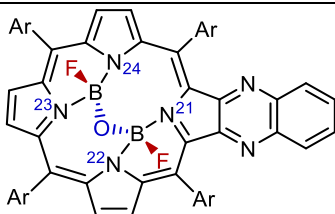
Figure 4.8   Naming convention for atoms involved in BAR processes, their application to **2a** – **5a**, the B transition-state structure linking **2a** and **3a**, and a defluorinated reaction intermediate, listing convenient abbreviations and full IUPAC substitutive systematic names, augmented with *parvo/amplo* stereodescriptors, that commence with "5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−b′]porphyrin-21,23-diido-" and finish as indicated.   The B−O−B plane is highlighted, and the *meso* 3,5-di-*tert*-butylphenyl groups and hydrogens are omitted for clarity.
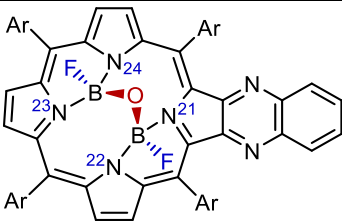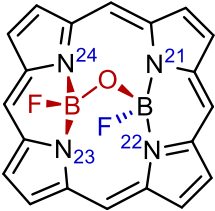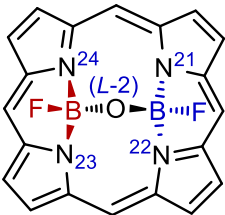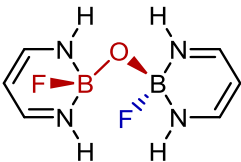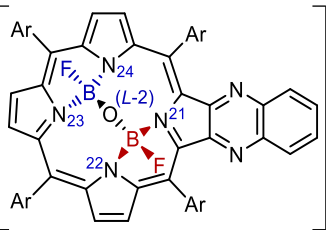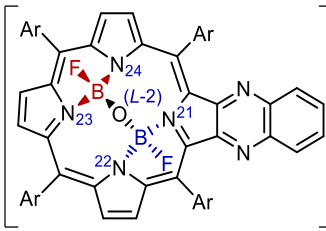
4.4.4.2 <u>Listing of all species and proposed systematic names</u>
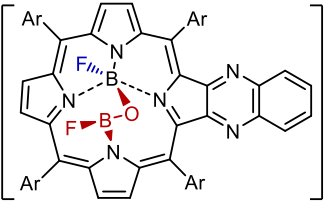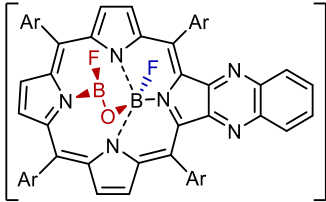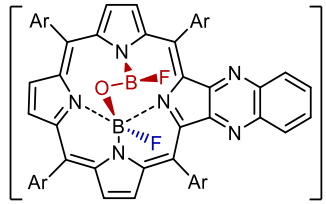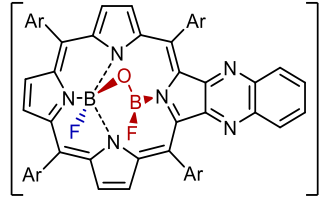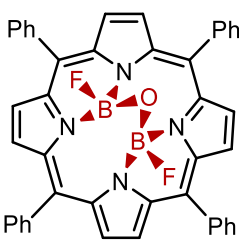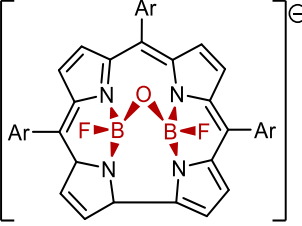
In Table 4.6, names are given to all neutral molecules and anionic corroles considered in this Chapter, their abbreviated names, and proposed full systematic substitutive IUPAC names modified with the *parvo* and *amplo* stereodescriptors. Additionally, alternative systematic additive IUPAC names are shown in blue for some species. For all 2D structures, all indicated stereochemistry is relative to the porphyrin ligand with atoms and bonds marked in red being above the quasi-plane of the porphyrin whilst those in blue are below it. Table 4.7 provides corresponding information for all mono-defluorinated cationic species considered. The geometric angles leading to the specified values of the *parvo* and *amplo* stereodescriptors in both tables are listed in Table 4.8.

Table 4.6   Names for all neutral porphyrins and anionic corroles considered. References listed as, for example IR-9.3.2.1, refer to specific sections in the "IUPAC Red Book"[3].

Structure	Names
 Ar = 3,5-di-*tert*-butylphenyl **1** H$_2$pqx	**1**, H$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)-22*H*,24*H*-quinoxalino[2,3−*b*′]porphyrin
 Ar = 3,5-di-*tert*-butylphenyl **2a** (*R$_a$*,*R$_p$*)-B$_2$OF$_2$pqx	**2a**, (*R$_a$*,*R$_p$*)-B$_2$OF$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1κ^2*N*^{21,22}:3κ^2*N*^{23,24}-{[1(*T*-4-*R$_{amplo}$*),2(*A*-2),3(*T*-4-*R$_{parvo}$*)]-1,3-difluoro-1λ^4,3λ^4-diboroxane} 1,2-difluoro-μ-oxo(μ-[5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxal-ino[2,3−*b*′]-21*H*,23*H*-porphinato(2-)-1κ^2*N*^{21,22}:2κ^2*N*^{23,24}])-[1(*T*-4-*R$_{amplo}$*),2(*T*-4-*R$_{parvo}$*)]-diboron
 Ar = 3,5-di-*tert*-butylphenyl **2b** (*S$_a$*,*S$_p$*)-B$_2$OF$_2$pqx	**2b**, (*S$_a$*,*S$_p$*)-B$_2$OF$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1κ^2*N*^{21,22}:3κ^2*N*^{23,24}-{[1(*T*-4-*S$_{amplo}$*),2(*A*-2),3(*T*-4-*S$_{parvo}$*)]-1,3-difluoro-1λ^4,3λ^4-diboroxane} 1,2-difluoro-μ-oxo(μ-[5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxal-ino[2,3−*b*′]-21*H*,23*H*-porphinato(2-)-1κ^2*N*^{21,22}:2κ^2*N*^{23,24}])-[1(*T*-4-*S$_{amplo}$*),2(*T*-4-*$_{parvo}$*)]--diboron

Structure	Names
Ar = 3,5-di-*tert*-butylphenyl **3a** (R_p,R_a)-B$_2$OF$_2$pqx	**3a**, (R_p,R_a)-B$_2$OF$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b'*]porphyrin-21,23-diido-1κ$^2N^{21,22}$:3κ$^2N^{23,24}$-{[1(*T*-4-R_{parvo}),2(*A*-2),3(*T*-4-R_{amplo})]-1,3-difluoro-1λ^4,3λ^4-diboroxane} 1,2-difluoro-μ-oxo(μ-[5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxal-ino[2,3−*b'*]-21*H*,23*H*-porphinato(2-)-1κ$^2N^{21,22}$:2κ$^2N^{23,24}$])-[1(*T*-4-R_{parvo}),2(*T*-4-R_{amplo})]-diboron
Ar = 3,5-di-*tert*-butylphenyl **3b** (S_p,S_a)-B$_2$OF$_2$pqx	**3b**, (S_p,S_a)-B$_2$OF$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b'*]porphyrin-21,23-diido-1κ$^2N^{21,22}$:3κ$^2N^{23,24}$-{[1(*T*-4-S_{parvo}),2(*A*-2),3(*T*-4-S_{amplo})]-1,3-difluoro-1λ^4,3λ^4-diboroxane} 1,2-difluoro-μ-oxo(μ-[5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxal-ino[2,3−*b'*]-21*H*,23*H*-porphinato(2-)-1κ$^2N^{21,22}$:2κ$^2N^{23,24}$])-[1(*T*-4-S_{parvo}),2(*T*-4-S_{amplo})]-diboron
Ar = 3,5-di-*tert*-butylphenyl **4a** (R_a,S_a)-B$_2$OF$_2$pqx	**4a**, (R_a,S_a)-B$_2$OF$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b'*]porphyrin-21,23-diido-1κ$^2N^{21,22}$:3κ$^2N^{23,24}$-{[1(*T*-4-R_{amplo}),2(*A*-2),3(*T*-4-S_{amplo})]-1,3-difluoro-1λ^4,3λ^4-diboroxane} 1,2-difluoro-μ-oxo(μ-[5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxal-ino[2,3−*b'*]-21*H*,23*H*-porphinato(2-)-1κ$^2N^{21,22}$:2κ$^2N^{23,24}$])-[1(*T*-4-R_{amplo}),2(*T*-4-S_{amplo})]-diboron
Ar = 3,5-di-*tert*-butylphenyl **4b** (S_a,R_a)-B$_2$OF$_2$pqx	**4b**, (S_a,R_a)-B$_2$OF$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b'*]porphyrin-21,23-diido-1κ$^2N^{21,22}$:3κ$^2N^{23,24}$-{[1(*T*-4-S_{amplo}),2(*A*-2),3(*T*-4-R_{amplo})]-1,3-difluoro-1λ^4,3λ^4-diboroxane} 1,2-difluoro-μ-oxo(μ-[5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxal-ino[2,3−*b'*]-21*H*,23*H*-porphinato(2-)-1κ$^2N^{21,22}$:2κ$^2N^{23,24}$])-[1(*T*-4-S_{amplo}),2(*T*-4-R_{amplo})]-diboron
Ar = 3,5-di-*tert*-butylphenyl **5a** (R_p,S_p)-B$_2$OF$_2$pqx	**5a**, (R_p,S_p)-B$_2$OF$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b'*]porphyrin-21,23-diido-1κ$^2N^{21,22}$:3κ$^2N^{23,24}$-{[1(*T*-4-R_{parvo}),2(*A*-2),3(*T*-4-S_{parvo})]-1,3-difluoro-1λ^4,3λ^4-diboroxane} 1,2-difluoro-μ-oxo(μ-[5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxal-ino[2,3−*b'*]-21*H*,23*H*-porphinato(2-)-1κ$^2N^{21,22}$:2κ$^2N^{23,24}$])-[1(*T*-4-R_{parvo}),2(*T*-4-S_{parvo})]-diboron

Structure	Names
 Ar = 3,5-di-*tert*-butylphenyl **5b** (S_p,R_p)-B$_2$OF$_2$pqx	**5b**, (S_p,R_p)-B$_2$OF$_2$pqx: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1$\kappa^2N^{21,22}$:3$\kappa^2N^{23,24}$-{[1(*T*-4-S_{parvo}),2(*A*-2),3(*T*-4-R_{parvo})]-1,3-difluoro-1λ^4,3λ^4-diboroxane} 1,2-difluoro-μ-oxo(μ-[5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxal-ino[2,3−*b*′]-21*H*,23*H*-porphinato(2-)-1$\kappa^2N^{21,22}$:2$\kappa^2N^{23,24}$])-[1(*T*-4-S_{parvo}),2(*T*-4-R_{parvo})]-diboron
 (*parvo,amplo*)-B$_2$OF$_2$p	(*parvo,amplo*)-B$_2$OF$_2$p: porphyrin-21,23-diido-1$\kappa^2N^{21,22}$:3$\kappa^2N^{23,24}$-{[1(*T*-4-*parvo*),2(*A*-2),3(*T*-4-*amplo*)]-1,3-difluoro-1λ^4,3λ^4-diboroxane} difluoro-μ-oxo[μ-(*21H,23H*-porphinato(2-)-1$\kappa^2N^{21,22}$:2$\kappa^2N^{23,24}$)]-[1(*T*-4-*parvo*),2(*T*-4-*amplo*)]-diboron
 TS$_{ak}$ 2(*L*-2)-B$_2$OF$_2$p‡	**TS$_{ak}$**, 2(*L*-2)-B$_2$OF$_2$p, BAR transition state porphyrin-21,23-diido-1$\kappa^2N^{21,22}$:3$\kappa^2N^{23,24}$-{[1(*T*-4),2(*L*-2),3(*T*-4)]-1,3-difluoro-1λ^4,3λ^4-diboroxane} ‡ difluoro-μ-oxo[μ-(21*H*,23*H*-porphinato(2-)-1$\kappa^2N^{21,22}$:2$\kappa^2N^{23,24}$)]-(2(*L*-2))-diboron‡
 (ipaBF)$_2$O	*transoid* (ipaBF)$_2$O: bis[2-fluoro-1,2-dihydro-1,2λ^4,2-diazaborinine]oxide
 TS$_{ak}^R$ Ar = 3,5-di-*tert*-butylphenyl	**TS$_{ak}^R$**, (*R,L*-2,*R*)-B$_2$OF$_2$pqx‡, 2a↔3a akamptisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1$\kappa^2N^{21,22}$:3$\kappa^2N^{23,24}$-{[1α(*T*-4-*R*),2(*L*-2),3β(*T*-4-*R*)]-1,3-difluoro-1λ^4,3λ^4-diboroxane}‡ See IR-9.3.2.1[3] for use of (*L*-2) polyhedral symbol
 TS$_{ak}^S$ Ar = 3,5-di-*tert*-butylphenyl	**TS$_{ak}^S$**, (*S,L*-2,*S*)-B$_2$OF$_2$pqx‡, 2b↔3b akamptisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1$\kappa^2N^{21,22}$:3$\kappa^2N^{23,24}$-{[1α(*T*-4-*S*),2(*L*-2),3β(*T*-4-*S*)]-1,3-difluoro-1λ^4,3λ^4-diboroxane}‡ See IR-9.3.2.1[3] for use of (*L*-2) polyhedral symbol

Structure	Names
 TS₃ₐ₂ₐ Ar = 3,5-di-*tert*-butylphenyl	**TS$_{2a3a}$**, (*C_p*,*amplo*)-B$_2$OF$_2$pqx‡, **3a↔2a** strepsisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1κ$^3N^{21,22,23}$:3κN^{24}-{[1(*TBPY*-5-*C_{parvo}*),2(*A*-2),3(*TP*-3-*amplo*)]-1,3-difluoro-1λ^5-diboroxane}‡ See IR-9.3.4.5^3 for use of (*C/A*) stereoconfiguration symbol
 TS₃ᵦ₃ₐ Ar = 3,5-di-*tert*-butylphenyl	**TS$_{3a3b}$**, (*parvo*,*amplo*)-B$_2$OF$_2$pqx‡, **3b↔3a** strepsisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1κ$^3N^{21,22,24}$:3κN^{23}-{[1(*TBPY*-5-*parvo*),2(*A*-2),3(*TP*-3-*amplo*)]-1,3-difluoro-1λ^5-diboroxane}‡
 TS₂ᵦ₃ᵦ Ar = 3,5-di-*tert*-butylphenyl	**TS$_{2b3b}$**, (*A_p*,*amplo*)-B$_2$OF$_2$pqx‡, **2b↔3b** strepsisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1κ$^3N^{21,22,23}$:3κN^{24}-{[1(*TBPY*-5-*A_{parvo}*),2(*A*-2),3(*TP*-3-*amplo*)]-1,3-difluoro-1λ^5-diboroxane}‡ See IR-9.3.4.5^3 for use of (*C/A*) stereoconfiguration symbol
 TS₂ₐ₂ᵦ Ar = 3,5-di-*tert*-butylphenyl	**TS$_{2a2b}$**, (*amplo*,*parvo*)-B$_2$OF$_2$pqx‡, **2a↔2b** strepsisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-1κN^{21}:3κ$^3N^{22,23,24}$-{[1(*TP*-3-*amplo*),2(*A*-2),3(*TBPY*-5-*parvo*)]-1,3-difluoro-3λ^5-diboroxane}‡
	(*amplo*,*amplo*)-B$_2$OF$_2$tpp: 5,10,15,20-tetraphenyl-porphyrin-21,23-diido-1κ$^2N^{21,22}$:3κ$^2N^{23,24}$-{[1(*T*-4-*amplo*),2(*A*-2),3(*T*-4-*amplo*)]-1,3-difluoro-1λ^4,3λ^4-diboroxane} difluoro-μ-oxo[μ-(5,10,15,20-tetraphenyl-21*H*,23*H*-porphinato(2-)-1κ$^2N^{21,22}$:2κ$^2N^{23,24}$)]-[1(*T*-4-*amplo*),2(*T*-4-*amplo*)]-diboron
 Ar = 4-methylphenyl (*amplo*,*amplo*)-B$_2$OF$_2$corrolate(1-)	(*amplo*,*amplo*)-B$_2$OF$_2$corrole(1-), conjugate base of CAS number 1013891-91-2 (Component: 1013965-96-2): [μ-[1,2,3,7,8,12,13,17,18,19-decadehydro-21,22-dihydro-5,10,15-tris(4-methylphenyl)corrinato(3-)-κN^{21},κN^{24}:κN^{22},κN^{23}]]difluoro-μ-oxo-[1(*T*-4-*amplo*),2(*T*-4-*amplo*)]-diborate(1-)
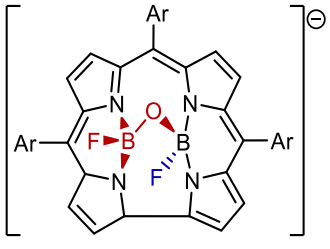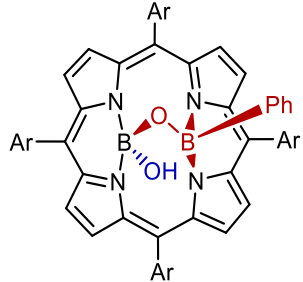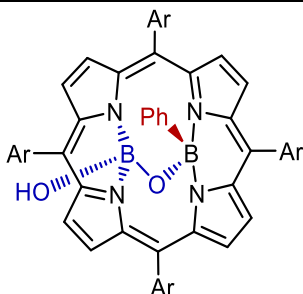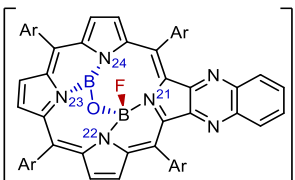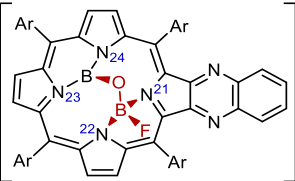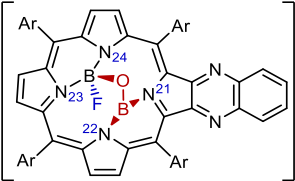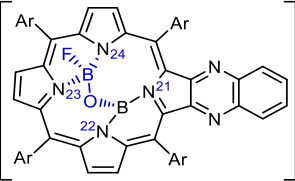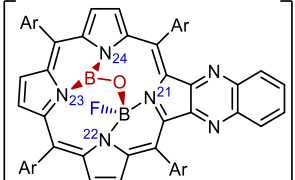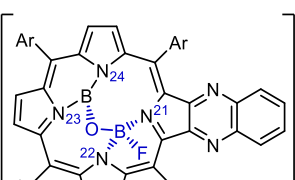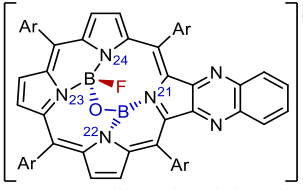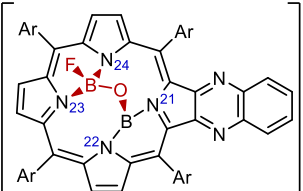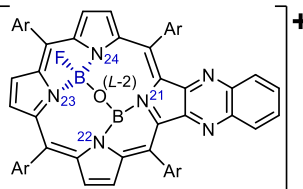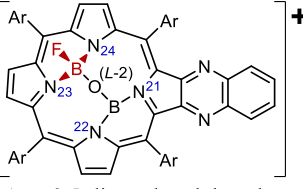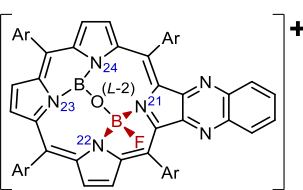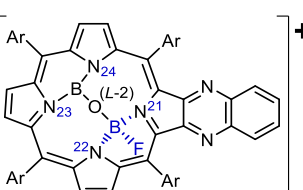
Structure	Names
 Ar = 4-methylphenyl (S_p,S_a)-B$_2$OF$_2$corrolate(1-)	(S_p,S_a)-B$_2$OF$_2$corrole(1-): [μ-[1,2,3,7,8,12,13,17,18,19-decadehydro-21,22-dihydro-5,10,15-triphenylcorrinato(3-)-κN^{21},κN^{24}:κN^{22},κN^{23}]]difluoro-μ-oxo-[1(T-4-S_{parvo}),2(T-4-S_{amplo})]-diborate(1-)
 Ar = 4-methylphenyl ($parvo$,$amplo$)-B$_2$O(Ph)(OH)-p	($parvo$,$amplo$)-B(OH)OB(Ph)ttp, CAS number 1022915-75-8: 5,10,15,20-tetrakis(4-methylphenyl)porphyrin-21,23-diido-1κ$^2N^{21,22}$:3κ$^2N^{23,24}$-{[1(T-4-$parvo$),2(A-2),3(T-4-$amplo$)]-1-hydroxy-3-phenyl-1λ^4,3λ^4-diboroxane} hydroxy-μ-oxophenyl[μ-[5,10,15,20-tetrakis(4-methylphenyl)-21H,23H-porphinato(2-)-κN^{21},κN^{22}:κN^{23},κN^{24}]]-[1(T-4-$parvo$),2(T-4-$amplo$)]-diboron.
 Ar = 4-methylphenyl ($amplo$,$parvo$)−B$_2$O(Ph)(OH)-p	($amplo$,$parvo$)-B(OH)OB(Ph)ttp: 5,10,15,20-tetrakis(4-methylphenyl)porphyrin-21,23-diido-1κ$^2N^{21,22}$:3κ$^2N^{23,24}$-{[1(T-4-$amplo$),2(A-2),3(T-4-$parvo$)]-3-hydroxy-1-phenyl-1λ^4,3λ^4-diboroxane] hydroxy-μ-oxophenyl[μ-[5,10,15,20-tetrakis(4-methylphenyl)-21H,23H-porphinato(2-)-κN^{21},κN^{22}:κN^{23},κN^{24}]]-[1(T-4-$amplo$),2(T-4-$parvo$)]-diboron.

107

Table 4.7 Names for all mono-defluorinated cations considered. References listed as, for example IR-9.3.2.1, refer to specific sections in the "IUPAC Red Book"[3].

Structure	Names
Ar = 3,5-di-*tert*-butylphenyl $(R_p,\_)$-B₂OFpqx⁺	$(R_p,\_)$-B₂OFpqx⁺: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b′*] porphyrin-21,23-diido-1κ²$N^{21,22}$:3κ²$N^{23,24}$-{[1(*T*-4-R_{parvo}),2(*A*-2),3(*TP*-3)]-1-fluoro-1λ⁴-diboroxanium} See IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl $(R_a,\_)$-B₂OFpqx⁺	$(R_a,\_)$-B₂OFpqx⁺: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b′*] porphyrin-21,23-diido-1κ²$N^{21,22}$:3κ²$N^{23,24}$-{[1(*T*-4-R_{amplo}),2(*A*-2),3(*TP*-3)]-1-fluoro-1λ⁴-diboroxanium} See IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl $(\_,R_p)$-B₂OFpqx⁺	$(\_,R_p)$-B₂OFpqx⁺: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b′*] porphyrin-21,23-diido-1κ²$N^{23,24}$:3κ²$N^{21,22}$-{[1(*TP*-3),2(*A*-2),3(*T*-4-R_{parvo}),]-1-fluoro-1λ⁴-diboroxanium} See IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl $(\_,R_a)$-B₂OFpqx⁺	$(\_,R_a)$-B₂OFpqx⁺: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b′*] porphyrin-21,23-diido-1κ²$N^{23,24}$:3κ²$N^{21,22}$-{[1(*TP*-3),2(*A*-2),3(*T*-4-R_{amplo}),]-1-fluoro-1λ⁴-diboroxanium} See IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl $(S_p,\_)$-B₂OFpqx⁺	$(S_p,\_)$-B₂OFpqx⁺: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b′*] porphyrin-21,23-diido-1κ²$N^{21,22}$:3κ²$N^{23,24}$-{[1(*T*-4-S_{parvo}),2(*A*-2),3(*TP*-3)]-1-fluoro-1λ⁴-diboroxanium} See IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl $(S_a,\_)$-B₂OFpqx⁺	$(S_a,\_)$-B₂OFpqx⁺: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b′*] porphyrin-21,23-diido-1κ²$N^{21,22}$:3κ²$N^{23,24}$-{[1(*T*-4-S_{amplo}),2(*A*-2),3(*TP*-3)]-1-fluoro-1λ⁴-diboroxanium} See IR-6.4.3[3] for nomenclature of substituted cations

Structure	Names
Ar = 3,5-di-*tert*-butylphenyl (_,S_p)-B$_2$OFpqx$^+$	(_,S_p)-B$_2$OFpqx$^+$: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′] porphyrin-21,23-diido-1$\kappa^2N^{23,24}$:3$\kappa^2N^{21,22}$-{[1(*TP*-3),2(*A*-2),3(*T*-4-S_{parvo}),]-1-fluoro-1λ^4-diboroxanium} See IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl (_,S_a)-B$_2$OF$_2$pqx$^+$	(_,S_a)-B$_2$OFpqx$^+$: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′] porphyrin-21,23-diido-1$\kappa^2N^{23,24}$:3$\kappa^2N^{21,22}$-{[1(*TP*-3),2(*A*-2),3(*T*-4-S_{amplo}),]-1-fluoro-1λ^4-diboroxanium} See IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl **TS**$_{ak1}^{R+}$ (_,*R*)-B$_2$OFpqx$^{+‡}$	**TS**$_{ak1}^{R+}$, (_,*R*)-B$_2$OFpqx+‡, cationic akamptisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′] porphyrin-21,23-diido-1$\kappa^2N^{23,24}$:3$\kappa^2N^{21,22}$-{[1(*TP*-3),2(*L*-2),3(*T*-4-*R*)-1-fluoro-1λ^4-diboroxanium}‡ See IR-9.3.2.1[3] for use of (*L*-2) polyhedral symbol and IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl **TS**$_{ak1}^{S+}$ (_,*S*)-B$_2$OFpqx$^{+‡}$	**TS**$_{ak1}^{S+}$, (_,*S*)-B$_2$OFpqx$^{+‡}$, cationic akamptisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′] porphyrin-21,23-diido-1$\kappa^2N^{23,24}$:3$\kappa^2N^{21,22}$-{[1(*TP*-3),2(*L*-2),3(*T*-4-*S*)-1-fluoro-1λ^4-diboroxanium}‡ See IR-9.3.2.1[3] for use of (*L*-2) polyhedral symbol and IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl **TS**$_{ak2}^{R+}$ (_,*R*)-B$_2$OFpqx$^{+‡}$	**TS**$_{ak2}^{R+}$, (*R*,_)-B$_2$OFpqx+‡, cationic akamptisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′] porphyrin-21,23-diido-1$\kappa^2N^{21,22}$:3$\kappa^2N^{23,24}$-{[1(*T*-4-*R*),2(*L*-2),3(*TP*-3)]-1-fluoro-1λ^4-diboroxanium}‡ See IR-9.3.2.1[3] for use of (*L*-2) polyhedral symbol and IR-6.4.3[3] for nomenclature of substituted cations
Ar = 3,5-di-*tert*-butylphenyl **TS**$_{ak2}^{S+}$ (_,*R*)-B$_2$OFpqx$^{+‡}$	**TS**$_{ak2}^{S+}$, (*S*,_)-B$_2$OFpqx+‡, cationic akamptisomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′] porphyrin-21,23-diido-1$\kappa^2N^{21,22}$:3$\kappa^2N^{23,24}$-{[1(*T*-4-*S*),2(*L*-2),3(*TP*-3)]-1-fluoro-1λ^4-diboroxanium}‡ See IR-9.3.2.1[3] for use of (*L*-2) polyhedral symbol and IR-6.4.3[3] for nomenclature of substituted cations
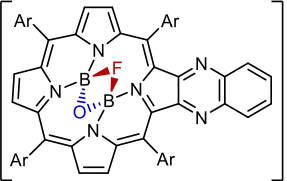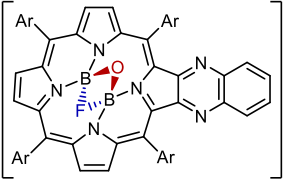
Structure	Names
 Ar = 3,5-di-*tert*-butylphenyl **TS$_{en}^{RS+}$** (*R,S*)-B$_2$OFpqx$^{+\ddagger}$	**TS$_{en}^{RS+}$**, (*R,S*)-B$_2$OFpqx$^{+\ddagger}$, cationic enantiomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-2κ$^2N^{21,22}$:4κ$^2N^{23,24}$-{[1(*A*-2),2(*T*-4-*R*),3(*A*-2),4(*T*-4-*S*)]-1λ^2,3,2λ^4,4λ^4-fluoraoxadiboretanium}‡
 Ar = 3,5-di-*tert*-butylphenyl **TS$_{en}^{SR+}$** (*S,R*)-B$_2$OFpqx$^{+\ddagger}$	**TS$_{en}^{SR+}$**, (*S,R*)-B$_2$OFpqx$^{+\ddagger}$, cationic enantiomerisation transition state: 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b*′]porphyrin-21,23-diido-2κ$^2N^{21,22}$:4κ$^2N^{23,24}$-{[1(*A*-2),2(*T*-4-*S*),3(*A*-2),4(*T*-4-*R*)]-1λ^2,3,2λ^4,4λ^4-fluoraoxadiboretanium}‡

Table 4.8 Listing of all species internal angles in degrees, angle ratios and *parvo/amplo* assignments.

Compound	θ_1	θ_2	θ_{BOB}	Reference akamptisomer	θ_1'	θ_2'	θ_1/θ_1'	θ_2/θ_2'	Stereo-descriptors
2a, (R_a,R_p)-B$_2$OF$_2$pqx	60	5	118	**3a** (R_p,R_a)-B$_2$OF$_2$pqx	5	60	11.6	0.1	*amplo,parvo*
2b, (S_a,S_p)-B$_2$OF$_2$pqx	60	5	118	**3b**, (S_p,S_a)-B$_2$OF$_2$pqx	5	60	11.6	0.1	*amplo,parvo*
3a, (R_p,R_a)-B$_2$OF$_2$pqx	5	60	118	**2a**, (R_a,R_p)-B$_2$OF$_2$pqx	60	5	0.1	11.6	*parvo,amplo*
3b, (S_p,S_a)-B$_2$OF$_2$pqx	5	60	118	**2b**, (S_a,S_p)-B$_2$OF$_2$pqx	60	5	0.1	11.6	*parvo,amplo*
4a, (R_a,S_a)-B$_2$OF$_2$pqx	65	65	143	**5a**, (R_p,S_p)-B$_2$OF$_2$pqx	39	39	1.7	1.7	*amplo,amplo*
4b, (S_a,R_a)-B$_2$OF$_2$pqx	65	65	143	**5b**, (S_p,R_p)-B$_2$OF$_2$pqx	39	39	1.7	1.7	*amplo,amplo*
5a, (R_p,S_p)-B$_2$OF$_2$pqx	39	39	121	**4a**, (R_a,S_a)-B$_2$OF$_2$pqx	65	65	0.6	0.6	*parvo,parvo*
5b, (S_p,R_p)-B$_2$OF$_2$pqx	39	39	121	**4b**, (S_a,R_a)-B$_2$OF$_2$pqx	65	65	0.6	0.6	*parvo,parvo*
6, (*parvo,amplo*)-B$_2$OF$_2$p	4	61	118	**6**, (*amplo,parvo*)-B$_2$OF$_2$p	61	4	0.1	14.1	*parvo,amplo*
TS$_{ak}^6$, $2(L$-$2)$-B$_2$OF$_2$p	39	39	180	**TS$_{ak}^6$**, $2(L$-$2)$-B$_2$OF$_2$p	39	39	1	1	none,none
TS$_{ak}^R$, **2a**↔**3a** BAR TS	40	38	180	**TS$_{ak}^R$**, **2a**↔**3a** BAR TS	40	38	1	1	none,none
TS$_{ak}^S$, **2b**↔**3b** BAR TS	40	38	180	**TS$_{ak}^S$**, **2b**↔**3b** BAR TS	40	38	1	1	none,none
TS$_{2a3a}$, **2a**↔**3a** streps. TS	4	57	118	**TS$_{3b2b}$**, **3b**↔**2b** streps. TS	57	4	0.1	15.2	*parvo,amplo*
TS$_{3a3b}$, **3a**↔**3b** streps. TS	4	56	117	**TS$_{2b2a}$**, **2b**↔**2a** streps. TS	56	4	0.1	15.2	*parvo,amplo*
TS$_{3b2b}$, **3b**↔**2b** streps. TS	57	4	118	**TS$_{2a3a}$**, **2a**↔**3a** streps. TS	4	57	15.2	0.1	*amplo,parvo*[a]
TS$_{2b2a}$, **2b**↔**2a** streps. TS	56	4	117	**TS$_{3a3b}$**, **3a**↔**3b** streps. TS	4	56	15.2	0.1	*amplo,parvo*
(*parvo,amplo*)-B$_2$O(Ph)(OH)-p	13	62	121	(*amplo,parvo*)-B2O(Ph)(OH)-p	62	8	0.2	8.1	*amplo,parvo*
(*amplo,parvo*)-B$_2$O(Ph)(OH)-p	62	8	119	(*parvo,amplo*)-B2O(Ph)(OH)-p	13	62	0.1	4.8	*parvo,amplo*
$(R_p,\_)$-B$_2$OFpqx$^+$	7	29	104	$(R_a,\_)$-B$_2$OFpqx$^+$	64	32	0.1	0.9	*parvo*,none
$(R_a,\_)$-B$_2$OFpqx$^+$	64	32	129	$(R_p,\_)$-B$_2$OFpqx$^+$	7	29	9.1	1.1	*amplo*,none
$(\_,R_p)$-B$_2$OFpqx$^+$	29	7	104	$(\_,R_a)$-B$_2$OFpqx$^+$	32	65	0.9	0.1	none,*parvo*
$(\_,R_a)$-B$_2$OFpqx$^+$	32	65	129	$(\_,R_p)$-B$_2$OFpqx$^+$	29	7	1.1	9.7	none,*amplo*
$(S_p,\_)$-B$_2$OFpqx$^+$	7	29	104	$(S_a,\_)$-B$_2$OFpqx$^+$	64	32	0.1	0.9	*parvo*,none
$(S_a,\_)$-B$_2$OFpqx$^+$	64	32	129	$(S_p,\_)$-B$_2$OFpqx$^+$	7	29	9.1	1.1	*amplo*,none
$(\_,S_p)$-B$_2$OFpqx$^+$	29	7	104	$(\_,S_a)$-B$_2$OFpqx$^+$	32	65	0.9	0.1	none,*parvo*
$(\_,S_a)$-B$_2$OFpqx$^+$	32	65	129	$(\_,S_p)$-B$_2$OFpqx$^+$	29	7	1.1	9.7	none,*amplo*
TS$_{ak1}^{R+}$, $(\_,R)$-B$_2$OFpqx$^{+\ddagger}$	5	44	155	**TS$_{ak1}^{R+}$**, $(\_,R)$-B$_2$OFpqx$^{+\ddagger}$	5	44	1	1	none,none
TS$_{ak1}^{S+}$, $(\_,S)$-B$_2$OFpqx$^{+\ddagger}$	5	44	155	**TS$_{ak1}^{S+}$**, $(\_,S)$-B$_2$OFpqx$^{+\ddagger}$	5	44	1	1	none,none
TS$_{ak2}^{R+}$, $(R,\_)$-B$_2$OFpqx$^{+\ddagger}$	44	5	155	**TS$_{ak2}^{R+}$**, $(R,\_)$-B$_2$OFpqx$^{+\ddagger}$	44	5	1	1	none,none
TS$_{ak2}^{S+}$, $(S,\_)$-B$_2$OFpqx$^{+\ddagger}$	44	5	155	**TS$_{ak2}^{S+}$**, $(S,\_)$-B$_2$OFpqx$^{+\ddagger}$	44	5	1	1	none,none
TS$_{en}^{RS+}$	11	10	93	**TS$_{en}^{SR+}$**	10	11	1	1	none,none[b]
TS$_{en}^{SR+}$	10	11	93	**TS$_{en}^{RS+}$**	11	10	1	1	none,none[b]

a: alternate ring numbering giving priority to *parvo* leads to **TS$_{3b2b}$** being named $(1A_p,3amplo)$-B$_2$OF$_2$pqx and not $(1amplo,3Ap)$-B$_2$OF$_2$pqx, *b*: BAR of a F$^-$ transfer transition state requires the conceptual BAR of both the F and O atoms.

Primary conclusions concerning terminology and nomenclature:

New terminology is essential for describing the BAR-related stereoisomerism and has been introduced. Additionally, new nomenclatural stereodescriptors are sometimes required to describe the BAR-related geometric details of the bridgehead atoms L^1 and L^2 comprising an embedded ML_2 stereocentre. The stereodescriptors *parvo* and *amplo* are introduced and rigorously defined. They are compatible with existing IUPAC stereodescriptors. The IUPAC recommended wedge-hash notation fails for embedded BAR-related stereoisomers with non-standard usage introduced.

4.5 Detailed DFT understanding of reaction mechanisms

To show the relationships between the conceived compounds **2** – **5**, (DFT) calculations at the B3LYP/6-31+G* level were performed, seeking stable isomers corresponding to **2** – **5** and transition states for BAR and other possible unimolecular processes, as well as for probable bimolecular processes, that could interconvert them. Full details are given in Methods Section 4.7.1.3, with optimised coordinates provided in E_File_8 (see page xxv).

Previously, Table 4.1 presented a results overview pertaining to isomerisation reactions possible in simple solution under ambient conditions. Here, this is extended by the consideration of more processes and by the examination of reactions pertaining to the conditions used in the synthesis, with results presented in Table 4.9. Considered in detail are the following possible reaction mechanisms:

Unimolecular processes:

      i)   B−O−B BAR,

      ii)   B−O−B hula-twist-like mechanism,

      iii)   B(F)OB(F)-group bond-walk-rotation (strepsisomerisation).

Bimolecular processes:

      iv)   Uncatalysed F^- loss,

      v) $HNEt_3^+$ catalysed F^- loss,

      vi) $BF_3 \cdot NEt_3$ catalysed F^- loss,

      vii)   uncatalysed F^- S_N2,

      viii) F^- loss with counterfacial addition.

Table 4.9 Observed and B3LYP/6-31G+(d)-calculated reaction activation free energies in kJ mol^{-1}, for various types of isomerisation mechanisms in (a) simple CH_2Cl_2 solutions[a] and (b) under the conditions of synthesis [b]. Entries (a) are reproduced from Table 4.1 and included here to facilitate direct comparison with entries (b).

Environment	Process	Akamptisomerisation		Enantiomerisation		*cisoid* relaxation		
		2a⇌3a 3b⇌2b	3a⇌2a 2b⇌3b	2a⇌2b	3a⇌3b	4a→2a 4a→3b 4b→2b 4b→3a	4a→2a 4b→3b	4a→2b 4b→3b
	Observed	*104 ± 2[c]*	*104 ± 2[c]*	N.O.[e]	N.O.[e]	–	–	–
	B−O−B BAR	108	108	N.A.[f]	N.A.[f]	–	–	–
(a) CH_2Cl_2 solution [a]	B−O−B hula-twist-like	>108	>108	N.A.[f]	N.A.[f]	–	–	–
	B(F)OB(F) bond-walk-rotation [d]	225	225	251	247	–	–	–
	F⁻ loss [a]	284	285	220	220	–	–	–
(b) Synthesis conditions [b]	HNEt$_3^+$ catalysed F⁻ loss	180	180	116	119	–	–	–
	BF$_3$·NEt$_3$ catalysed F⁻ loss	180	180	106	109	–	–	–
	Uncatalysed F⁻ S$_N$2	–	–	–	–	82	–	–
	F⁻ loss, counterfacial addition	–	–	–	–	–	143	149

a: see Sections 4.7.1.3.1, 4.7.1.3.2, 4.7.1.3.3, and 4.7.1.3.4 for details, including pathways and considered alternatives. *b*: see Section 4.7.1.3.4. *c*: 96% confidence interval.
d: Strepsisomerisation – see Section 4.7.1.3.3. *e*: N.O. – Not observed. *f*: N.A. – Not applicable.

4.5.1 BAR mechanism

BAR R_{st}^c1 mechanisms are, by definition, characterised by *A*-2 polytopes interchanging *via* an intermediate *L*-2 polytope (see Chapter 3, Section 3.5). There exists a wide range of barriers to linearisation for ML$_2$ systems as shown in Table 4.10 for bicoordinate oxygen and sulfur species with the nature of the L-atoms playing as significant role as the nature of M. Whilst noting the rich diversity of equilibrium bond angles and barriers to linearisation shown in Table 4.10, a full discussion is beyond the scope of this Thesis.

To provide insight into the nature of the energy barrier associated with BAR, the model system ipaB(F)−O−B(F)ipa (see Figure 4.9) is considered where ipa = 3-iminoprop-1-en-1-amine. It is introduced as it embodies the essential electronic and structural character present in B(F)−O−B(F)-porphyrins but without the topological and steric constraints of the full macrocycle. Further, the compact size of the chelating ipa ligands avoid intramolecular close-contact interactions and hence this model compound provides an excellent reference

system for the DFT calculated barrier to linearisation without the steric constraints imposed by a porphyrin macrocycle.

Table 4.10 BAR barriers for ML$_2$ centres in some simple compounds.

Species	Bond angle	Equilibrium angle	$\Delta G^{\ddagger}$ / kJ mol^{-1} ^{a}	$\Delta G^{\ddagger}$ / kJ mol^{-1} ^{d}
H−O−H ^{b}	H−O−H	104.7°	127	132.51
Li−O−Li ^{b}	Li−O−Li	[180°]	[0]	
H$_3$N:B(H$_2$)−O−B(H$_2$):NH$_3$ ^{c}	B−O−B	120.3°	38	
ipaB(F)−O−B(F)ipa ^{cf}	B−O−B	129.8°	23	
Me−O−Me ^{b}	C−O−C	112.5°	149	154.66
F−O−F ^{be}	F−O−F	104.0°	326	
Cl−O−Cl ^{b}	Cl−O−Cl	112.8°	262	
H$_3$Si−O−SiH$_3$ ^{b}	Si−O−Si	137.9°	6	
F$_3$Si−O−SiF$_3$ ^{b}	Si−O−Si	[180°]	[0]	
H−S−H ^{b}	H−S−H	93.1°	284	289.67
Li−S−Li ^{b}	Li−S−Li	143.3°	11	
H$_3$N:B(H$_2$)−S−B(H$_2$):NH$_3$ ^{c}	B−S−B	99.9°	116	
ipaB(F)−S−B(F)ipa ^{cf}	B−S−B	106.3°	114	
Me−S−Me ^{b}	C−S−C	99.5°	293	310.16
H$_3$Si−S−SiH$_3$ ^{b}	Si−S−Si	101.1°	112	

a: relative electronic energy of species with bond angle = 180°, *b*: B3LYP-D3/6-311++G(2d,p) water solvation, *c*: B3LYP-D3/6-31+G(d) chloroform solvation, *d*: at the W2-F12 level of theory.[151] *e*: linear case constrained to $D_{\infty h}$ point group. *f*: the bidentate ligand ipa = 3-iminoprop-1-en-1-amine with the symmetries of complexes constrained to the C_s and C_{2h} point groups for the bent and linear geometries, respectively. These complexes conform to the structure with the F atoms *trans* to each other.

transoid
ipaB(F)OB(F)ipa

Potential-energy local-minimum structure under the constraint that the *Cs* symmetry point group is maintained. There are no close intramolecular contacts between each side of the complex.
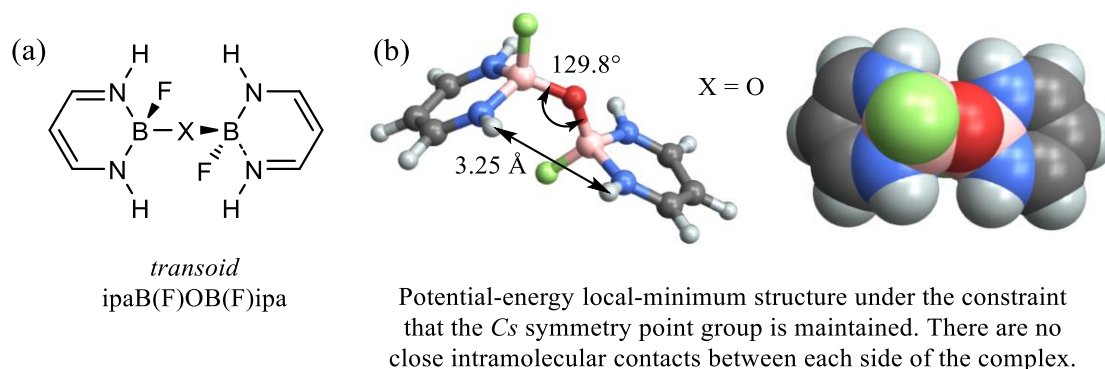
Figure 4.9 The topologically-unconstrained model compound *transoid* ipaB(F)OB(F)ipa complex, where ipa = 3-iminoprop-1-en-1-amine. (a) The 2D representation. (b) DFT calculated model of the model compound where symmetry is constrained to the C_s point group. The ball and stick figure indicate the B–O–B bond angle of 129.8° and the internuclear distance of the imino H atoms at 3.25 Å. The space-filling model clearly shows no close intramolecular contacts between each side of the complex. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

As shown in Table 4.10 (blue highlighted entry), the DFT calculated B–O–B bond angle for ipaB(F)OB(F)ipa is 129.8° and the DFT calculated barrier to linearisation is only 23 kJ mol^{-1}. In comparison, in **2** and **3** the DFT calculated B–O–B bond angle is 117.8° and

the corresponding DFT calculated barrier 108 kJ mol[-1]. This five-fold increase in the energy barrier and large reduction in the B–O–B bond angle are readily interpreted as a consequence of the compression of the coordinated B–O–B unit imposed by the porphyrin macrocycle with a concomitant tetragonal macrocycle distortion.[143, 152] These compressional forces are communicated *via* the N–B bonds of the complex. Whilst these forces are in balance between the macrocycle and ligated B(F)OB(F) unit, the strain energy is not necessarily equipartitioned.

The effects of the strain are shown in Table 4.11. The strain energy of the macrocycle alone is readily calculated with reference to a free-base porphyrin analogue. Full technical details are found in Methods Section 4.7.1.2.

Table 4.11 DFT calculated macrocycle strain energies ΔE for model (Ar = H) species **2a**, **3a** and $\mathbf{TS}_{ak}^{R}$, strain energy differences $\Delta\Delta E$, and tetragonal elongations Δd.

Species	Macrocycle strain energy^a ΔE / kJ mol[-1]	$\Delta\Delta E^{\,b}$ / kJ mol[-1]	Δd / Å ^{c}
2a	43	64	0.25
3a	44	63	0.25
$\mathbf{TS}_{ak}^{R}$	107	-	0.50

a: B3LYP-D3/6-311G(d,p). *b*: $\Delta\Delta E$ values are $\Delta E(\mathbf{TS}_{ak}^{R})$ - ΔE(**2a** or **3a**). *c*: Tetragonal elongation for the (N21-N24) internuclear distance relative to a free-base porphyrin reference structure distance of 3.35Å. See Methods Section 4.7.1.2 for details.

The large macrocycle strain energies ΔE, strain energy differences $\Delta\Delta E$, and tetragonal elongations Δd all indicate that the macrocycle is exhibiting significant strain in order to accommodate the compressed B–O–B unit. Considering the magnitude of these strain energies, it is reasonable to conclude that the facile formation of and stability of the B$_2$OF$_2$pqx stereoisomers must be due to the high strength of the B–N and B–O bonds.

An important practical consequence of the akamptisomerisation barrier being of the order of 100 kJ mol[-1] is that, whilst diastereomeric racemates **2** and **3** were readily separable using standard flash-silica column chromatography, their specific pattern of diastereomeric interconversion was also observable under ambient conditions during the chiral stationary-phase HPLC resolution of stereoisomers. If the barrier to stereoisomerisation was higher or lower by only 12 kJ mol[-1] then the serendipitous observation of this unusual dynamical behaviour is unlikely to have occurred and lead to the philosophical discourse that comprises this Thesis.

The BAR R_{st}^c1 transition structure linking **2a** and **3a** on the model system with Ar = H was optimised at the B3LYP-D3/6-31(+)G(d) level. The transition state, called $\boldsymbol{TS_{ak}^R}$, is shown in detail in Figure 4.10. An animation of the BAR R_{st}^c1 reaction **2a** ⇌ **3a** (*meso*-unsubstituted) is provided as **E_File_4** with the animation of the corresponding vibrational normal mode for *meso*-unsubstituted **2a** shown in **E_File_5** (vibrational normal mode 12, frequency 204.36 cm^{-1}, IR intensity 0.7641 km mol^{-1})

TS$_{ak}^R$

$(R,L$-2,$R)$-B$_2$OF$_2$pqx‡
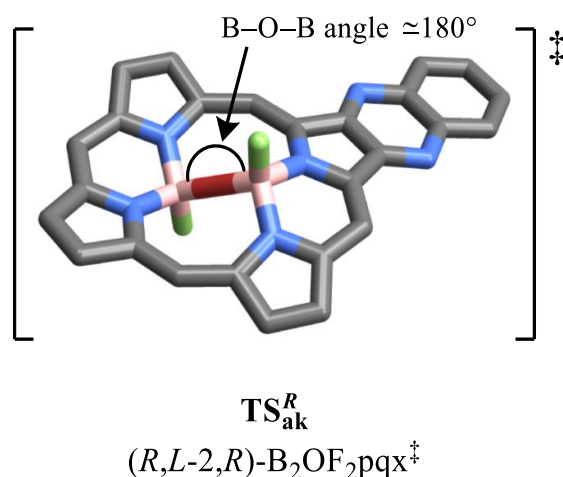
Figure 4.10 DFT calculated transition structure of the model system (unsubstituted porphyrin *meso* positions) for the BAR R_{st}^c1 reaction **2a** ⇌ **3a** (akamptisomerisation). Hydrogens omitted for clarity. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

Primary conclusion concerning the BAR mechanism:

The BAR mechanism is compatible with the observations.

4.5.2 Hula-twist-like mechanism

The linear nature of the B−O−B bond angle present in the DFT modelled BAR transition structures such as **TS$_{ak}^R$** (See Section 4.5.1, above) is a particular instance of the more general case where the oxygen atom moves through the porphyrin quasi-plane but adopts B−O−B angles other than 180°. This "swinging" motion is referred to as the "hula-twist-like mechanism" (see Sections 2.4.11 and 0).

The energetics and resulting *constrained* potential energy surface (PES) of this general scenario was modelled using the simpler unsubstituted high-symmetry *transoid* B(F)OB(F)-porphyrin **6** according to the autakamptisomerisation scheme shown in Figure 4.11.

6   where B−O−B is nonlinear
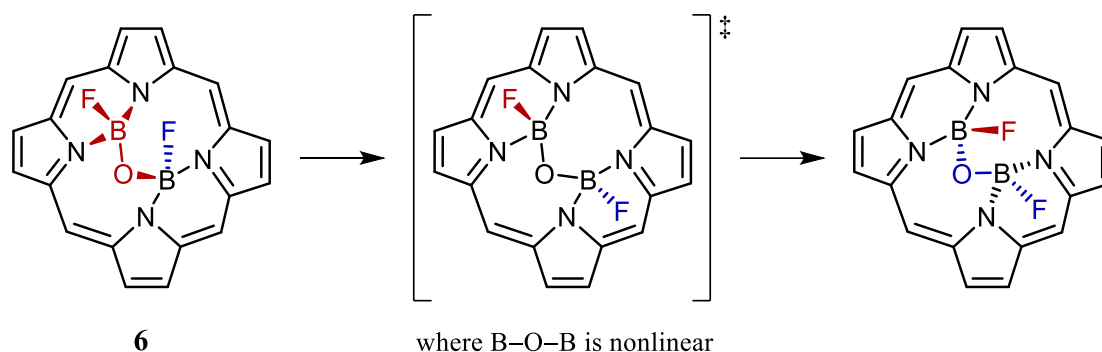
Figure 4.11   Scheme showing the unsubstituted *transoid* B(F)OB(F)-porphyrin **6** undergoing an R_{st}^c1 mechanism involving a nonlinear B−O−B bond angle. Such a process is called the hula-twist-like mechanism.

The resulting DFT PES is shown in Figure 4.12a under the geometrical constraint that the oxygen atom lies on the C_2 axis as indicated in Figure 4.12b.

Three semi-empirical methods were compared to the DFT results and are shown in Figure 4.12c. The older AM1 and PM3 methods both predict a potential energy double well indicative of a hula-twist-like mechanism. The newer and more accurate PM6 method reproduces a single potential energy well that closely tracks the DFT values for B−O−B angles in the range of ∼140° to ∼220°. A single potential energy well indicates that a linear B−O−B geometry corresponds to the akamptisomerisation transition structure.

A hula-twist-like mechanism not involving appreciable bond-angle flexion character would be expected to involve a corresponding transition structure with the B−O−B angle comparable to that found the ground state structure. **E_File_6** is an animation of such a mechanism for meso-unsubstituted **2a** ⇌ **3a** reaction with the B–O–B angle fixed at 130° throughout. For the DFT model described here, this would correspond to a B−O−B bond angle of around 118°. In Figure 4.12a, the vertical green and red dashed lines indicate B−O−B bond angles representative of ground state structures. The energy penalty for these values clearly indicate that a hula-twist-like mechanism is energetically uncompetitive with the linear B−O−B bond angle indicative of the BAR R_{st}^c1.

This analysis, as applied to the *transoid* B_2OF_2pqx stereoisomeric system shows that this system exhibits *both* geometric and mechanistic relationships arising purely by BAR.
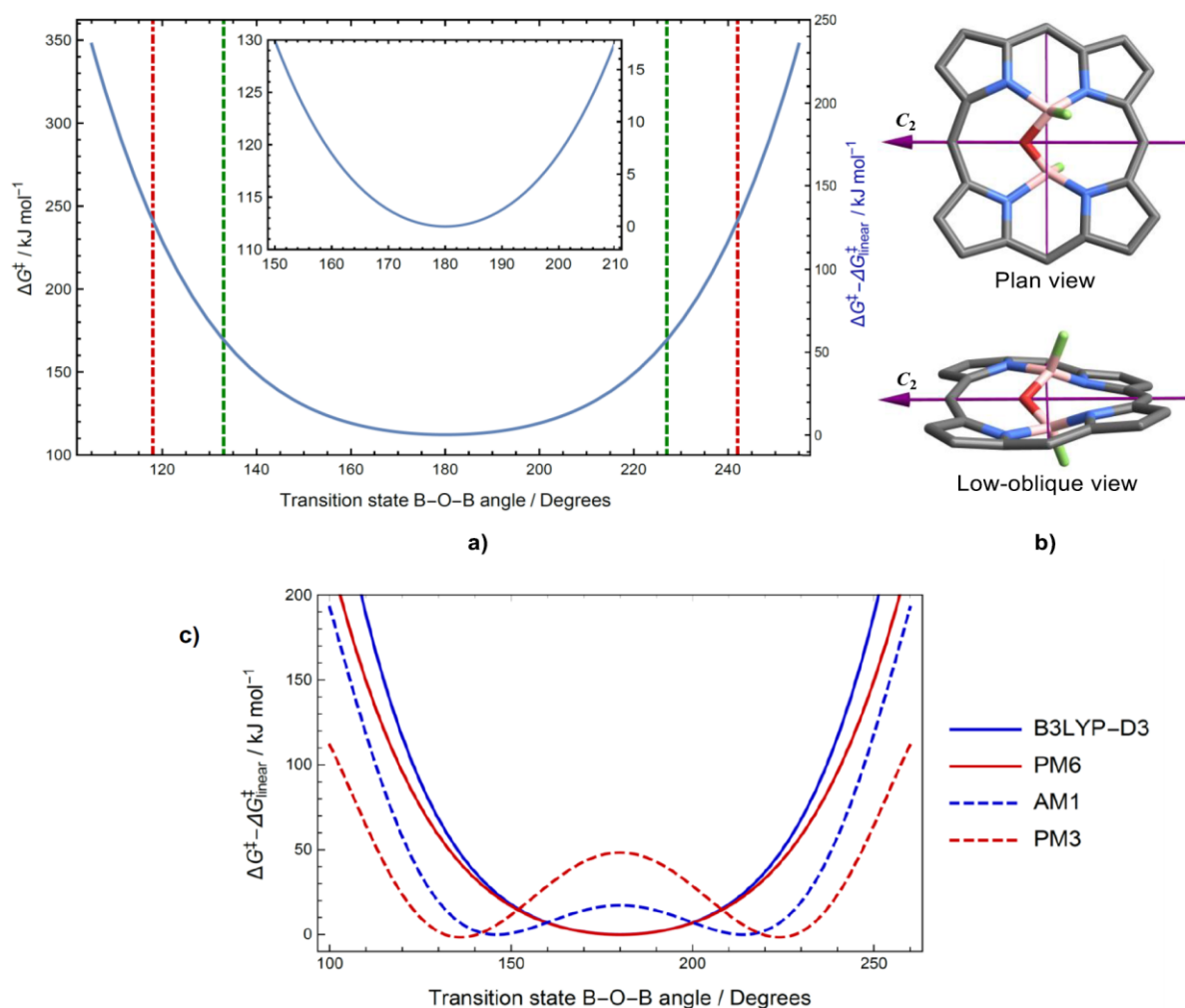
Figure 4.12 The hula-twist-like transition state energy dependencies on B−O−B angle. (a) B3LYP-D3/6-31G(d) potential-energy surface for akamptisomerisation hula-twist-like transition structure (TS) of **6**, featuring the transition structure shown in (b). The left axis shows the TS energy relative to the ground state. The right axis shows the TS energy relative to that for the linear B−O−B transition structure. The red dot-dashed vertical lines correspond to the B−O−B angles of 118° and 242° found in the *ground state* (potential energy local minimum) of **6**. The green dashed vertical lines correspond to the less constrained B−O−B angles of 133° and 227°. b) Plan and low-oblique views of a transition structure. The oxygen and two *meso* carbons, C-10 and C-20, lie on the C_2 axis represented by the purple arrow. The second line joining *meso* carbons C-5 and C-15, intersects the C_2 axis at right angles. Hydrogen atoms omitted for clarity. c) Comparison of potential well as calculated by different methods. The older AM1 and PM3 semiempirical methods (dashed lines) predict a double well indicative of a double torsional mechanism, whereas the more accurate B3LYP-D3/6-31G(d) (solid blue) and newer PM6 semiempirical (solid red) methods predict very similar single wells with a linear BAR transition state. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

A broader understanding of the relationship between the BAR mechanism and the hula-twist-like mechanisms is provided in Figure 4.13. First, Figure 4.13a depicts some aspects of the conformational isomerism of ethylmethyl ether, labelled to make each hydrogen atom distinguishable. A pair of akamptisomeric conformations are identified and the reactions between them considered. First, the BAR mechanism is applied to Akamptisomer 1, converting it to Akamptisomer 2 *via* a (high energy) BAR transition state. Note that this preserves the configurational relationships of ligands (the two carbon atoms

and attendant substituents) to the oxygen atom. Then, hula-twist-like (low energy) coupled torsions around the C–O bonds are applied, producing a structure that is rotationally equivalent to Akamptisomer 1. In this example, the coupled torsions comprising the hula-twist-like mechanism can be uncoupled with each allowed to occur sequentially to the same affect. Conformational isomerisation by BAR is not observable in this molecule as BAR is energetically uncompetitive compared to either coupled (hula-twist-like) or sequential separate torsion-based reaction mechanisms. This situation is extremely widely applicable to the chemistry of ethers.

Next, Figure 4.13b examines the same processes but with the key atoms encapsulated ("embedded" – see Section 2.4.9) inside a porphyrin macrocycle. A phenyl derivative is used to stress that it is not possible to transfer this substituent through the centre of the macrocyclic ring. First, the figure shows the BAR mechanism in direct analogy to that portrayed in Figure 4.13a for ethylmethyl ether. The oxygen atom is transferred through the macrocyclic ring to the other side by the BAR process, with the B–O–B bond angle becoming linear at the transition-state structure. The situation is very different for the hula-twist-like mechanism, however, as this mechanism demands that the phenyl ring by transported through the centre of the macrocyclic ring. In this way, confinement of the reacting group by the macrocycle inhibits hula-twist-like reactions much more than it does BAR reactions.

Nevertheless, coupled torsions may still be involved in the reaction mechanism if, as illustrated in Figure 4.13c, the B–O–B bond angle is constrained to be non-linear, large bond-angle changes around the boron atoms are permitted, and the resultant twisting-bending motion simply forces the oxygen through the macrocyclic plane. When the B–O–B bond angle becomes linear, couped-torsional motions can no longer be defined and this mechanism collapses to BAR. The interpretation of Figure 4.12 is that the DFT (and semi-empirical PM6) calculations predict that indeed this occurs, making any attempt at performing a coupled-torsion-based reaction mechanism combined with B–O–B bond angle changes collapse to BAR rather than providing a viable alternative.
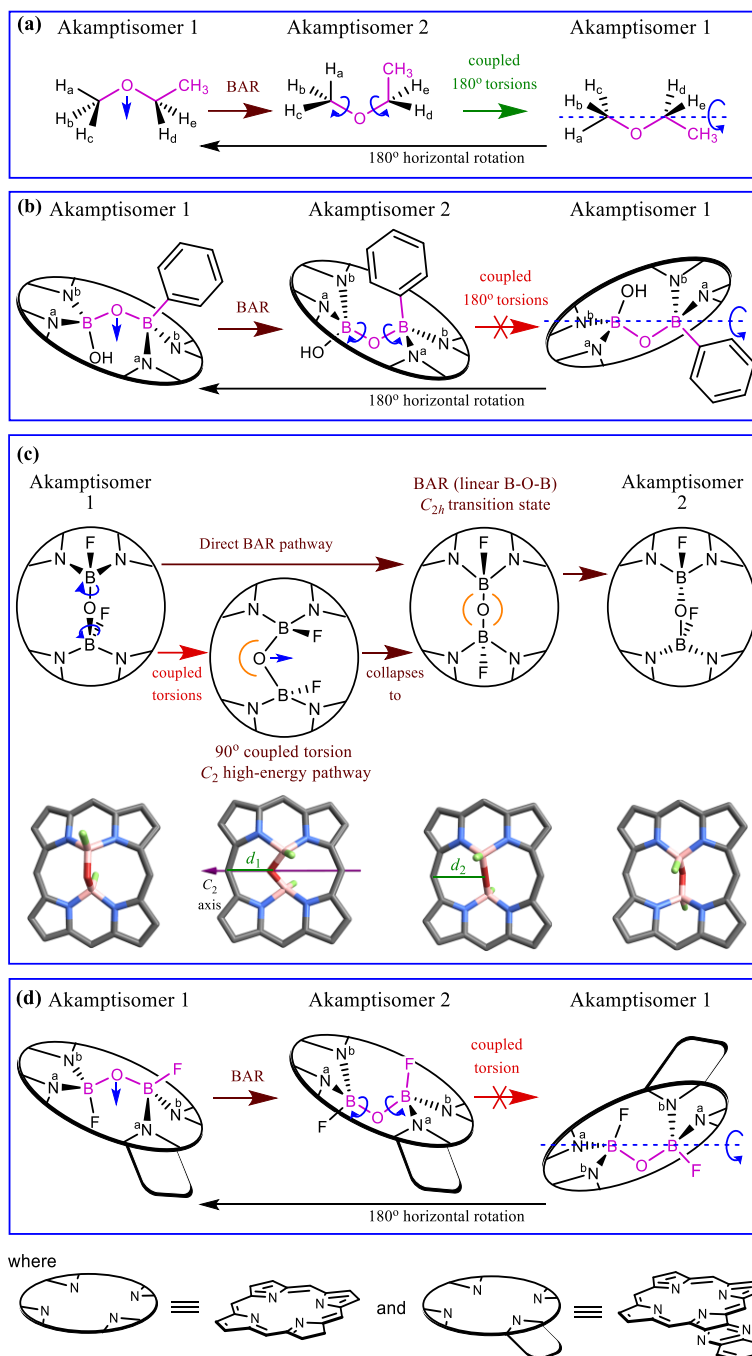
Figure 4.13   BAR processes within single bonds.   (a) BAR can conceivably occur in many molecules but coupled torsional motions conceivably undo the reaction at minimal energy cost, with two 180° torsional motions followed by a global 180° rotation demonstrating the effect most simply.   (b) Side view of B−O−B showing that embedding the atoms of interest within a macrocycle, couples and sterically hinders such torsional motions, producing distinct akamptisomers.   (c) Top view of macrocycle featuring an alternative double torsion involving only 90° motions could also undo BAR, but owing to compression from the macrocycle, double-torsion pathways involve strong steric repulsions (shown B3LYP/6-31G* optimised O−C non-bonded distance $d_1 = 2.27$ Å) are all of higher energy than the BAR transition state (C_{2h} optimised O−C non-bonded distance $d_2 = 2.94$ Å) and so do not undo the process.   (d) Asymmetric substitution of the macrocycle allows BAR to produce a distinct akamptisomer that also cannot revert through coupled torsional motions.   Views (a), (b), and (d) show the fundamental processes with respect to the magenta-coloured B−O−B (or C−O−C) atoms placed at characteristic positions and hence involve large motions of the surrounding atoms, whereas in (c) the view is with respect to the macrocycle, showing the O atom moving through it. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

Finally, Figure 4.13d depicts the situation for the stereoisomers **2** and **3**. The fusion of the quinoxaline ring onto the porphyrin makes **2** and **3** distinguishable. However, the distinguishing interactions involve atoms at the porphyrin centre interacting with the substituents out in the periphery and are therefore very weak. In this way, distinguishable compounds are produced with very different 3D shapes but very similar energies. BAR can interchange the akamptisomers, but torsional motions are inhibited by the connections of the reacting group to the macrocycle. This suppression of degrees of freedom of the flexing group is detailed in Chapter 3 in Figure 3.13.

Primary conclusion concerning the hula-twist-like mechanism:

A hula-twist-like mechanism involving torsional motions around the B–O bonds coupled with significant B–O–B bond angle changes, whilst compatible with the observations, is found to be energetically uncompetitive by DFT calculations.

4.5.3 Strepsisomerism

A unimolecular "bond-walk" mechanism termed strepsisomerisation (see Section 2.4.26) was considered as a potential stereoisomerisation process. During strepsisomerisation, the chelated B(F)OB(F) group rotates in a coherent manner inside the porphyrin cavity with each 90° rotation of the B(F)OB(F) group yielding a different stereoisomer. A full 360° rotation completes a strepsisomerisation cycle returning to the original stereoisomer.

For the 2,3-porphyrin substitution pattern studied here, the resulting strepsisomerisation cycle gives rise to two alternating rounds of akamptisomerisation (though not by the BAR mechanism) and enantiomerisation, *interconverting all* products **2a**, **2b**, **3a**, and **3b**. This is shown in Figure 4.14, with results collected in Table 4.9.
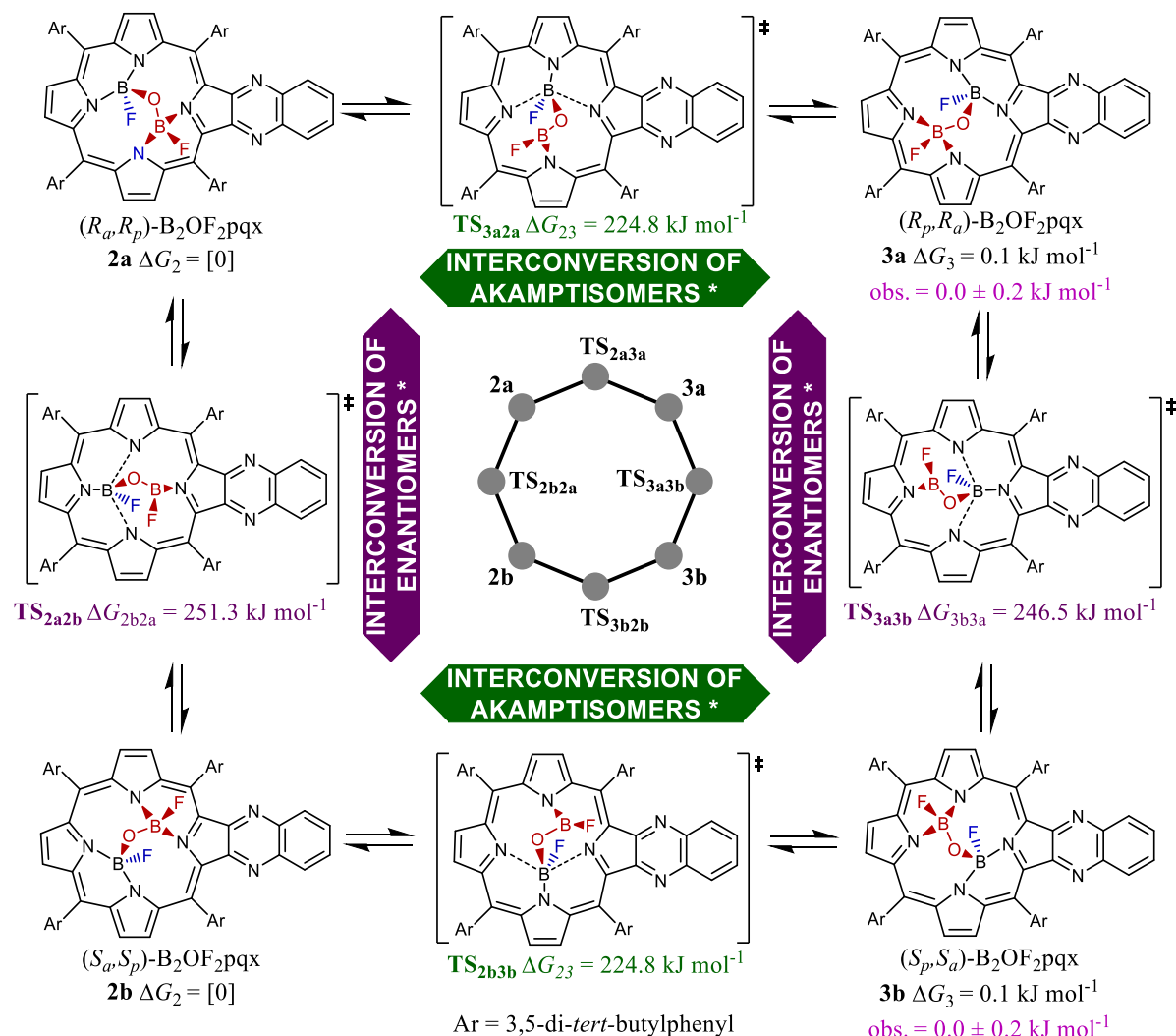
Figure 4.14 Concerted unimolecular "bond-walk" stereoisomerisation pathways affected through the coherent rotation of the B(F)OB(F) group inside the cavity of the macrocycle. All isomers interconvert through successive 90° rotations with intervening transition structures indicated. Ar = 3,5-di-*tert*-butylphenyl for the real systems. Calculated free energies for the Ar = H analogues are at the B3LYP-D3/6-31G+(d) level including implicit dichloromethane solvation. All indicated stereochemistry is relative to the porphyrin ligand. Atoms and bonds marked in red are above the quasi-plane of the porphyrin whilst those in blue are below it. * The interconversion of akamptisomers is not through BAR nor is the interconversion of enantiomers through the dual inversion of the stereocentres; in each case these isomerisations are a feature specifically arising from the combined symmetries of both the ligand and the *transoid*-B(F)OB(F) group. Inset in the centre of the scheme is the corresponding labelled reaction graph.

The strepsisomerisation transition-state structure **TS$_{2a3a}$** connecting stereoisomers **2a** and **3a** (with Ar = H) is shown in Figure 4.15 and is representative of all four strepsisomerisation transition-state structures.
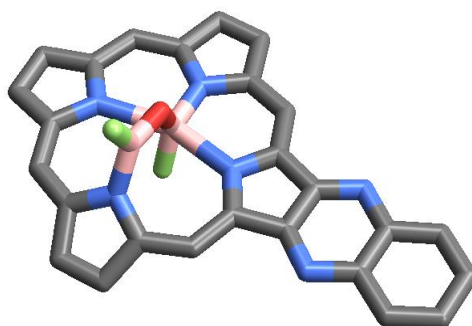
TS$_{3a2a}$

(C_p,*amplo*)-B$_2$OF$_2$pqx‡

Figure 4.15 3D structure the strepsisomerisation "bond-walk" mechanism transition structure for the reaction **2a** ⇌ **3a** as featured in Figure 4.14. Peripheral hydrogen atoms omitted for clarity. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

The PES of the full strepsisomerisation cycle is shown in Figure 4.16. The structure of this surface is strongly indicative of an asynchronous process where a weaker N–B(*amplo*) bond breaks first as the energy rises changing the coordination of this boron atom from *T*-4 to *TP*-3. On completion of this stage the B(*parvo*) atom moves changing coordination from *T*-4 to *TBPY*-5.
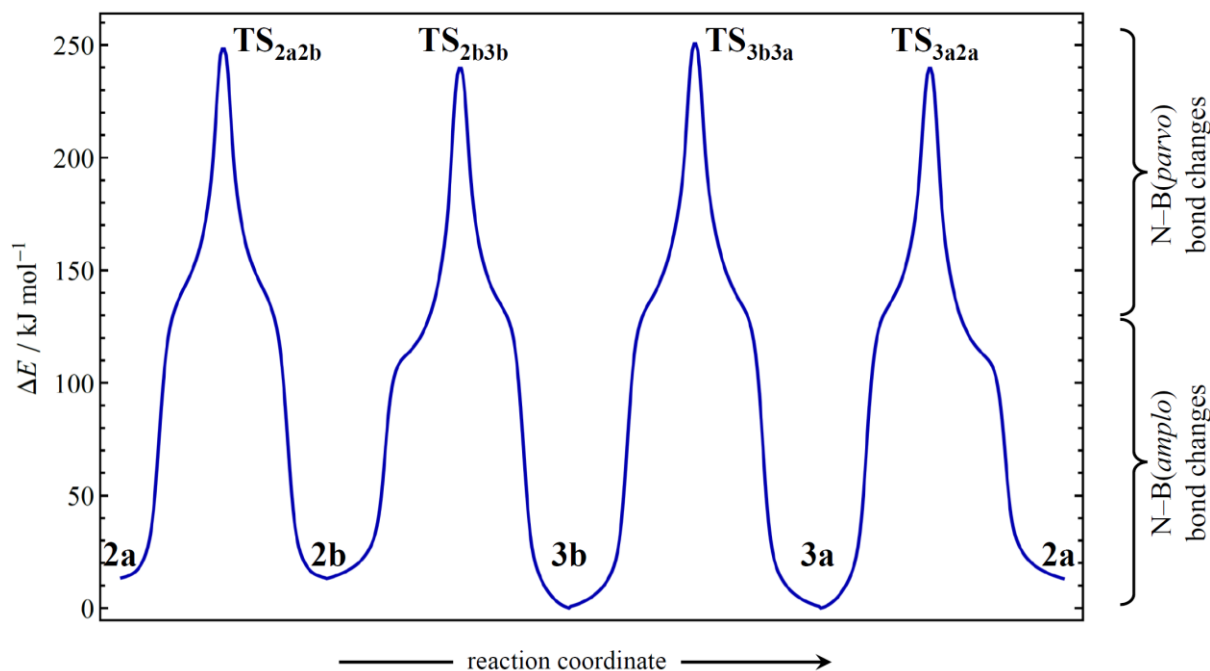
Figure 4.16 Potential energy surface scan of the strepsisomerisation cycle corresponding to Figure 4.14 for the model (Ar = H) system calculated at the PM7 level with energies uniformly scaled to match the DFT energy range. The reaction coordinate corresponds to the rotation of the B(F)OB(F) group within the macrocycle cavity. The shape of the surface indicates that each local minimum to transition-state-structure proceeds in a strongly asynchronous manner with the weaker N–B(*amplo*) bonds breaking first as the energy rises. An animation of the strepsisomerisation process is illustrated in **E_File_7**.

Figure 4.17 shows structures at three stages along the reaction path **2a** to **TS₂ₐ₂ᵦ** corresponding to the first eighth of the PES in Figure 4.16. The intermediate structure shows that the *amplo* boron has become *TP*-3 whilst the *parvo* boron remains *T*-4. Proposed systematic names are given for each structure reflecting the differences in the coordination of the boron atoms. An animation of the strepsisomerisation process is illustrated in **E_File_7** (see page xxv).

2a   N22–B(*amplo*) bond breakage   **TS₂ₐ₂ᵦ**

…1κ²$N^{21,22}$:3κ²$N^{23,24}$-{[1(*T*-4-R_{amplo}),2(*A*-2), …1κN^{21}:3κ²$N^{23,24}$-{[1(*TP*-3-*amplo*),2(*A*-2), …1κN^{21}:3κ³$N^{22,23,24}$-{[1(*TP*-3-*amplo*),2(*A*-2),
3(*T*-4-R_{parvo})]-1,3-difluoro-1λ⁴,3λ⁴-diboroxane} 3(*T*-4-R_{parvo})]-1,3-difluoro-3λ⁴-diboroxane} 3(*TBPY*-5-*parvo*)]-1,3-difluoro-3λ⁵-diboroxane}‡
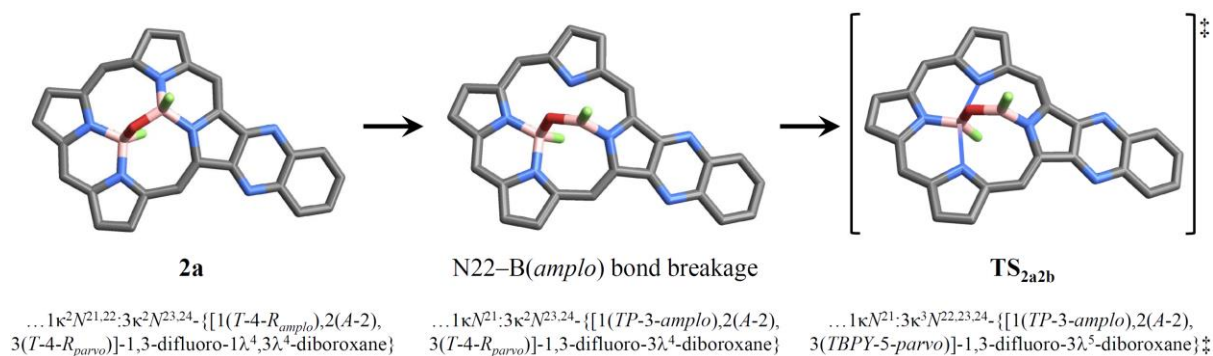
Figure 4.17 Modelled structures along the path from **2a** to **TS₂ₐ₂ᵦ** with an intermediate (centre) showing the changes in coordination of the boron atoms. The proposed systematic name of each structure is assigned and reflects the changes in coordination number with each name commencing 5,10,15,20-tetrakis(3,5-di-*tert*-butylphenyl)quinoxalino[2,3−*b′*]porphyrin-21,23-diido-… Hydrogens omitted for clarity. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

Primary conclusion concerning strepsisomerisation-mediated stereoisomerisation:

DFT modelling of the strepsisomerisation process shows its activation energies to be very high. Further, the strepsisomerisation mechanism interconverts all stereoisomers **2a**, **2b**, **3a**, and **3b** which is not observed in simple solutions. Modelling indicates that strepsisomerisation in the *transoid* B₂OF₂pqx system proceeds in a highly asynchronous manner.

4.5.4 *Bimolecular processes*

Whilst B−F bonds are particularly strong, dissociation of B₂OF₂pqx in solution to [B₂OFpqx]⁺ + F⁻ was anticipated to be energetically accessible due to the highly delocalised nature of the charge in the resulting cationic porphyrin species [B₂OFpqx]⁺, with the nature of the resulting anion playing an equally critical role. Loss of two fluoride anions to produce dicationic species [B₂Opqx]²⁺ was deemed energetically inaccessible under the conditions of synthesis and in simple solutions.

The special importance of these reactions arises from the ability of the cationic porphyrin to undergo BAR R_{st}^c1 mediated akamptisomerisation processes analogous to those of the neutral species. Unlike the neutral species B_2OF_2pqx, cationic B_2OFpqx^+ species can also undergo an intramolecular nucleophilic substitution (S_Ni, see Section A.52) which can also be described as 1,3-fluoride shift from one boron to the other. Such a process facilitates enantiomerisation and, hence, is a key step in racemisation. Example transition structures for each of these are shown in Figure 4.19.

Overall, the effect of these fluoride loss reactions is to enable a wide range of pathways interconverting all the isomers upon subsequent fluoride recombination. Interrelationships between **2a**, **2b**, **3a**, **3b, 4a** and **4b** are sketched in Figure 4.18, with key data collected in Table 4.9.
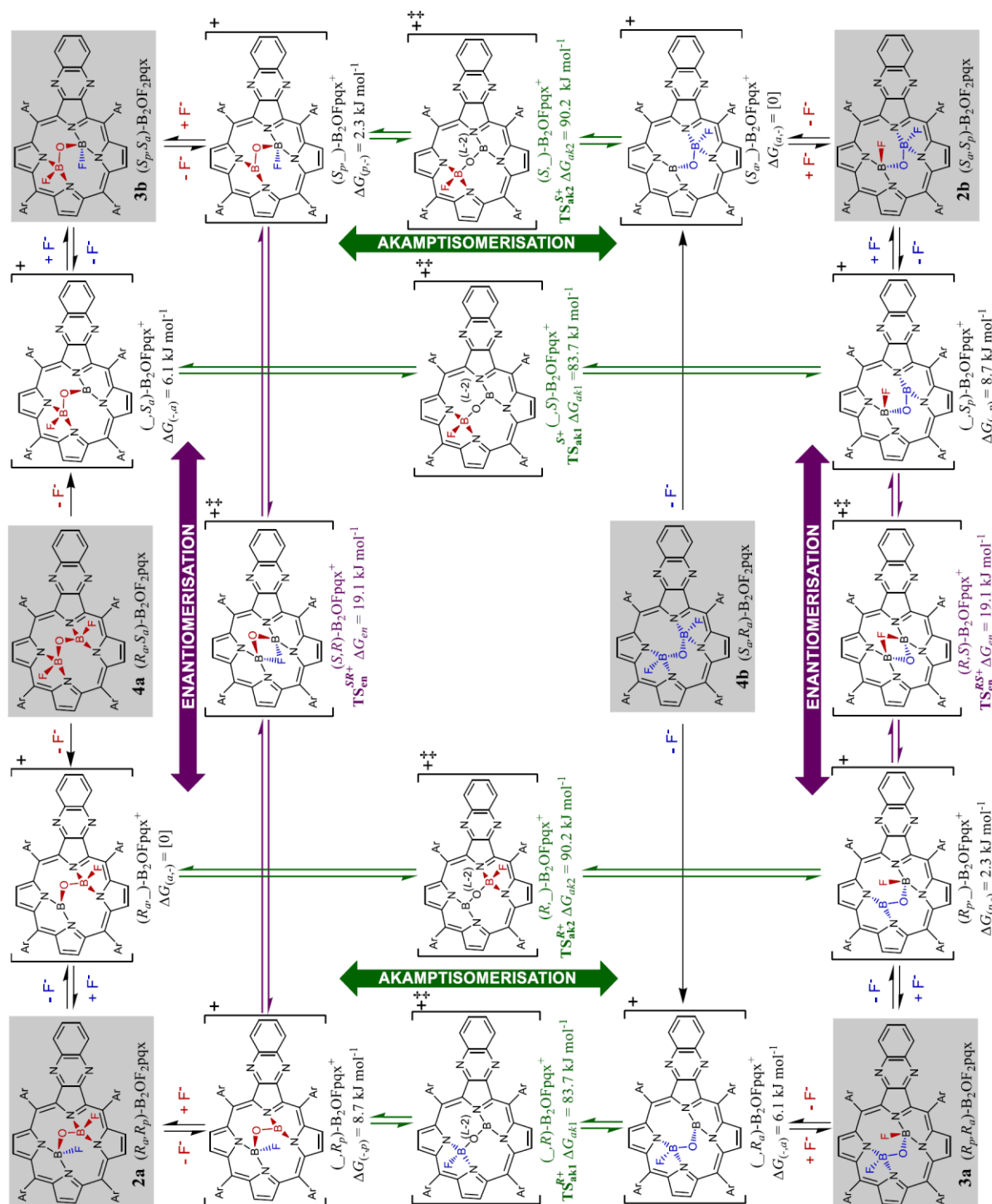
Figure 4.18 Bimolecular isomerisation pathways involving initial loss of fluoride ion. Ar = 3,5-di-*tert*-butylphenyl for the real systems. Free energies for the cationic species are calculated at B3LYP-D3/6-31+G(d) for the Ar = H analogues using an implicit dichloromethane solvation model. All indicated stereochemistry is relative to the porphyrin ligand. Atoms and bonds marked in red are above the quasi-plane of the porphyrin whilst those in blue are below it.

Figure 4.19 shows the DFT modelled structures for some of the transformations (local minima and transition-state structures) appearing in Figure 4.18 involving cationic species with an emphasis on an "enantiomerisation" step that is key to racemisation reactions, and a

BAR reaction. In Figure 4.19, the central local-minimum structure labelled as $(R_p,\_)$-B_2OFpqx^+ can react in the left direction to $(\_,S_p)$-B_2OFpqx^+ *via* a transition-state structure involving a 1,3 fluoride transfer. If $(R_p,\_)$-B_2OFpqx^+ reacts in the direction to the right, a BAR mechanism occurs giving $(R_a,\_)$-B_2OFpqx^+.
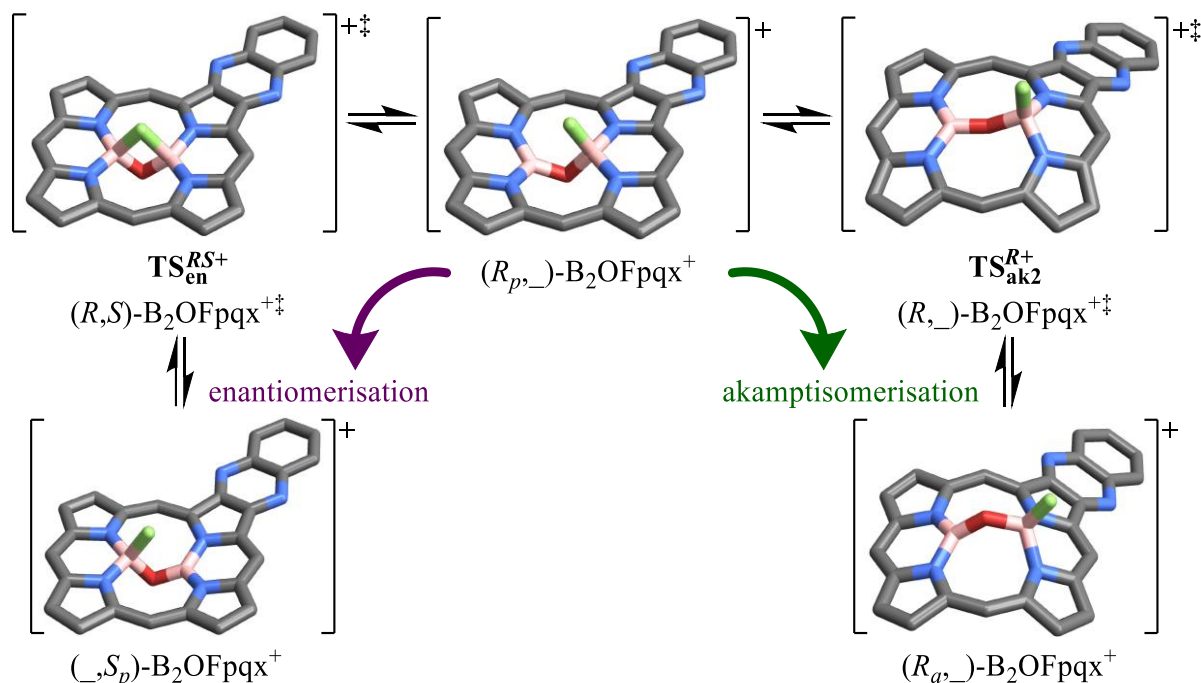
Figure 4.19 Example local minima and transition structures for some interconversions of cationic species shown in Figure 4.18. From the centre structure representing the $(R_p,\_)$-B_2OFpqx^+ species, the pathway to the left, as indicated by the purple arrow, results in the generation of a tetrahedral boron atoms with the (*S*) configuration – a key step in the enantiomerisation of *transoid* B_2OF_2pqx species. The pathway to the right, as indicated by the green arrow, results in akamptisomerisation of the cationic species. All DFT models were on *meso*-unsubstituted species. Hydrogens omitted for clarity. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

In Figure 4.20 is the labelled full reaction graph showing all the interconversion that appear in Figure 4.18 plus BAR R_{st}^c1 processes between all neutral species, the barrierless relaxation of **5 → 4,** and the linking of **5** to cationic enantiomerisation transition-state structures. This reaction graph, in conjunction with the results summarised in Table 4.9, demonstrate that, under the conditions of synthesis, interconversion of all B_2OF_2pqx stereoisomers are thermally accessible and allow for a thermodynamic distribution of products.

$2a = (R_a, R_p)$-B_2OF_2pqx
$3a = (R_p, R_a)$-B_2OF_2pqx

2a

$(R_a,\_)$-B_2OFpqx^+ TS_{ak}^R $(\_,R_p)$-B_2OFpqx^+

TS_{ak2}^{R+} TS_{ak1}^{R+}

$(R_p,\_)$-B_2OFpqx^+ **3a** $(\_,R_a)$-B_2OFpqx^+

4a   **5a** TS_{en}^{RS+} **4b**   **5b** TS_{en}^{SR+}

$(\_,S_p)$-B_2OFpqx^+ **3b** $(S_a,\_)$-B_2OFpqx^+

TS_{ak1}^{S+} TS_{ak2}^{S+}

$(\_,S_a)$-B_2OFpqx^+ TS_{ak}^S $(S_p,\_)$-B_2OFpqx^+

2b

● neutral species
● cationic species

$2b = (S_a, S_p)$-B_2OF_2pqx
$3b = (S_p, S_a)$-B_2OF_2pqx

$4a = (R_a, S_a)$-B_2OF_2pqx
$4b = (S_a, R_a)$-B_2OF_2pqx
$5a = (R_p, S_p)$-B_2OF_2pqx
$5b = (S_p, R_p)$-B_2OF_2pqx
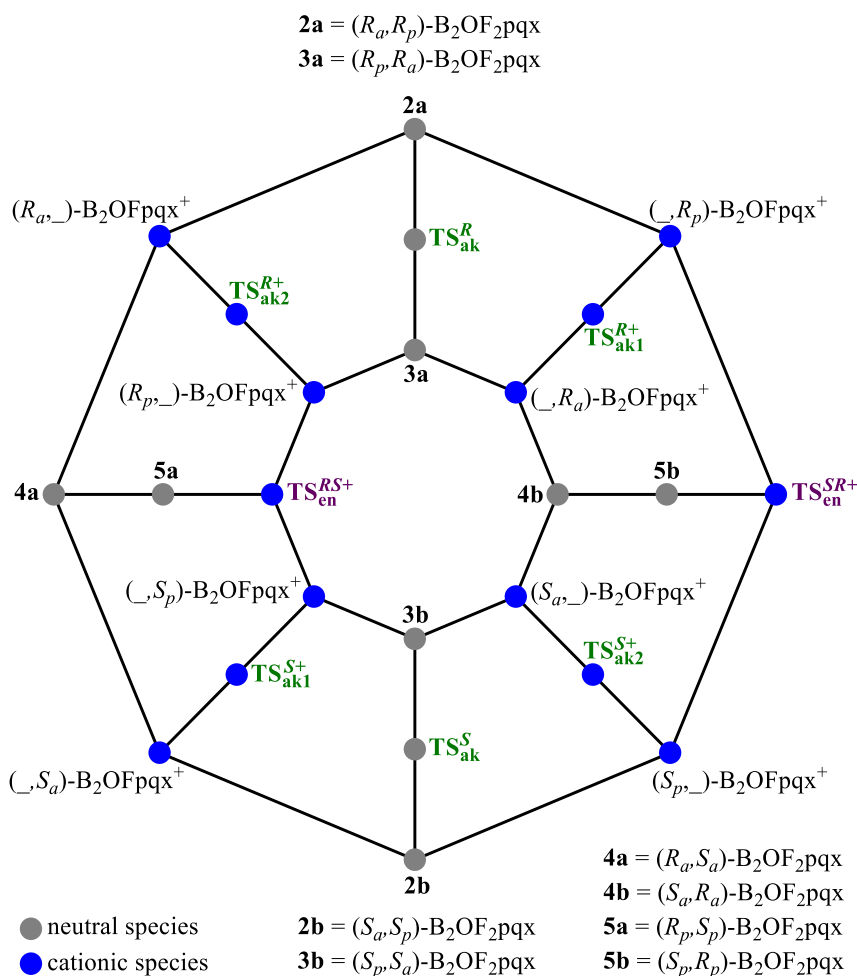
Figure 4.20 Labelled full reaction graph showing all the interconversions involving cationic local minima and transition-structure species (blue graph vertices) and neutral local minima and transition-structure species (grey graph vertices) as corresponding to those shown in Figure 4.18. Additionally, BAR $R_{st}^c 1$ processes between the neutral species, the barrierless relaxation of **5 ⇸ 4,** and the linking of **5** to cationic enantiomerisation transition-state structures are included.

A final bimolecular mechanism of note is the S_N2 reaction where fluoride attacks a boron atom *opposite* its substituent F-atom leading to Walden inversion of configuration (see Section A.51). For *transoid* species as exemplified by **2a**, neither of the constituent boron atoms is particularly exposed for attack by a nucleophile as shown in Figure 4.21a. In contrast and as shown in Figure 4.21b, one face of an *extended*-cisoid compound, as exemplified by **4a**, exposes both boron atoms opposite their substituent fluorine atoms. The conclusion is that for extended-*cisoid* compounds, there is a far larger likelihood of nucleophilic attack. Further, the products of such an S_N2 reaction are the thermodynamically favoured *transoid* species. Through this reasoning, S_N2 can be viewed as a viable mechanism for the one-way conversion of extended-*cisoid* B_2OF_2pqx species **4** to *transoid* stereoisomers **2** and **3**. The calculated reaction barrier for an uncatalysed S_N2 reaction is only 82 kJ mol^{-1}. Being uncatalysed, this provides an upper bound for S_N2 type reactions on reactant **4**.

minimal exposed boron surface opposite existing F-substituent

relatively large area of exposed boron surface opposite existing F-substituents

(a)   *transoid*

2a = (R_a,R_p)-B$_2$OF$_2$pqx

(b)   extended-*cisoid*

4a = (R_a,S_a)-B$_2$OF$_2$pqq
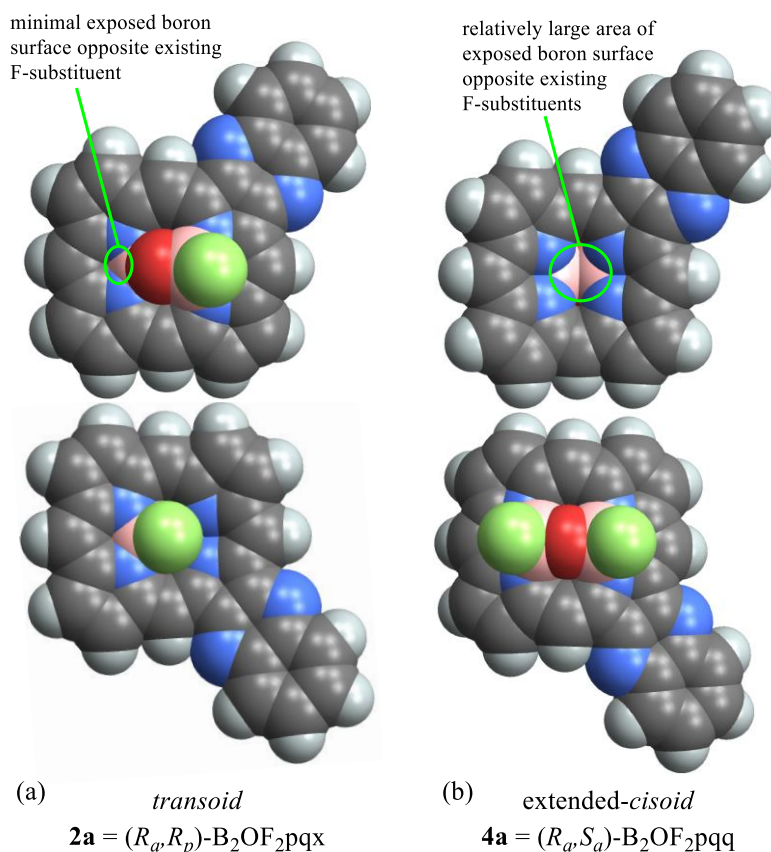
Figure 4.21 Space-filling models (at each element's the van der Waals radius) of the two faces of B$_2$OF$_2$pqx stereoisomers. (a) **2a**, representing a *transoid* configuration, does not allow for ready access of a nucleophile (fluoride) to either boron atom from the opposite side of their respective existing F-substituent. Both boron atoms are effectively sterically shielded from nucleophilic attack. (b) **4a**, representing an extended-*cisoid* configuration, does partially expose both boron atoms to attack from the side opposite both existing F-substituents. Only this extended-*cisoid* configuration was modelled for nucleophilic attack by F⁻. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

Primary conclusions concerning bimolecular processes:

DFT modelling of the bimolecular processes indicates that the energetics of a B$_2$OF$_2$pqx stereoisomer to lose a fluoride ion is a rate-determining factor. In simple solutions, the energy barrier to loss of fluoride is prohibitively high. Under the condition of synthesis, however, loss of fluoride is more facile *via* at least two catalysed mechanisms with all the resulting B$_2$OFpqx⁺ species capable of undergoing thermal isomerisation facilitating a thermodynamically controlled overall reaction and yielding only **2a**, **2b**, **3a**, and **3b**. Additionally, under the conditions of synthesis uncatalysed S$_N$2 reactions on extended-*cisoid* B$_2$OF$_2$pqx species **4** readily yield *transoid* B$_2$OF$_2$pqx species **2** and **3**.

4.5.5   Conclusions from the DFT perspective of stereoisomerisation

An examination of feasible unimolecular and bimolecular stereoisomerisation processes in these *transoid* B_2OF_2pqx systems demonstrate that *only* a BAR $R_{st}^c 1$ mechanism reproduces the pattern of stereoisomerisation observed in simple solutions with all other processes conceivably leading to the interconversion of all *transoid* and *cisoid* structures. The DFT calculated barrier for the BAR $R_{st}^c 1$ mechanism on a model system (Ar = H, see Section 4.5.1) is in excellent agreement with the observed experimental[7] value suggesting the choice of DFT modelling is appropriate. Applying this same DFT modelling approach to the other proposed stereoisomerisation mechanisms (see Section 4.5.2, Section 4.5.3, and Section 4.5.4) in simple solutions produces reaction barriers infeasibly high to contribute to stereoisomerisation at near-ambient temperatures providing quantitative support for the observed pattern.

Under the conditions of the synthesis, all processes considered here could conceivably contribute to interconversion of all stereoisomers facilitating thermodynamic control. This provides a plausible explanation for the distribution of isolated products.

4.6   Conclusions

X-ray crystallographic data was not able to be obtained to assign structures to the four chirally resolved HPLC B_2OF_2pqx stereoisomer fractions. An alternative approach where consideration of various theoretical and reported configurations of B(F)OB(F)-macrocycle complexes, led to the conclusion that there were eight distinct B_2OF_2pqx stereoisomeric possibilities. These comprised four chiral *transoid* complexes **2a**, **2b**, **3a**, and **3b** along with four chiral *cisoid* complexes **4a**, **4b**, **5a** and **5b**. Of these eight, DFT modelling indicated that only the *transoid* configurations need be considered. This was consistent with structures of previously reported B(F)OB(F)-porphyrin complexes and consistent with the observed 1:1:1:1 proportions of each B_2OF_2pqx isomer.

^{1}H, ^{13}C, and ^{19}F NMR data were interpreted using DFT assignments to definitively assign (**Fr1** + **Fr2**) to racemate **2** and (**Fr3** + **Fr4**) to racemate **3**. Further, CD and MCD measurements on enantiopure fractions, interpreted against DFT calculated chiroptical properties gave a robust assignment of **Fr1** to **2b**, **Fr2** to **2a**, **Fr3** to **3a**, and **Fr4** to **3b.**

The unprecedented nature of the demonstrated stereoisomerism arising from purely BAR-related geometric differences demanded new terminology. These BAR-related stereoisomers are described in geometry-only terms as "bond-angle reflectomers". In a similar manner to how the term *atropisomer* is defined by IUPAC, the term "akamptisomer" is introduced, this being the BAR analogue of the bond-torsion based atropisomer.

The IUPAC recommended wedge-hash notation for stereo-structures fails for embedded BAR-related stereoisomers, with non-standard usage therefore introduced. Additionally, existing IUPAC stereodescriptors were found to be inadequate for concisely and unambiguously describing these novel geometric relationships. The new "rigorously defined" BAR-specific stereodescriptors *parvo* and *amplo* are introduced and shown to be compatible with existing stereodescriptors.

A detailed DFT investigation into reaction mechanisms was undertaken. This included conceivable unimolecular and bimolecular processes that could occur in both simple solutions and under the conditions of the B_2OF_2pqx synthesis. The BAR mechanism was found to be compatible with the observed specific thermal diastereomeric interconversion under ambient conditions in simple solutions. The competing hula-twist-like mechanism was shown to be energetically uncompetitive to a BAR R_{st}^c1 mechnism.

The DFT modelling showed that under the conditions of synthesis, multiple reaction pathways were thermally accessible allowing full interconversion of stereoisomers including the *cisoid* stereoisomers. This provides a reasonable explanation for the observed thermodynamic control of synthesis products.

4.7   Methods

All geometry optimisations, vibrational normal mode analyses, NMR, electronic absorption and electronic-transition circular dichroism calculations were carried out using Gaussian-16[153]. All structures were characterised using vibrational frequency analyses with the harmonic-oscillator approximation used to estimate free energies.

B3LYP[154-155] was chosen as a "general purpose" functional that, with the inclusion of empirical dispersion corrections, is capable of reproducing acceptable geometries.[156] The combination of this functional and the relatively small partially augmented 6-31(+)G(d,p) basis set, along with other theoretical assumptions such as the polarised continuum solvation

model and the harmonic-oscillator approximation for free energies, would not typically be expected to yield free energies with high "chemical" accuracy,[156] however, for the processes and species described in this chapter, this typically leads to situations where there is a cancellation of any systematic errors and hence reasonable chemical accuracy is obtained.

At no point in this Chapter is the need for highly precise geometries or free energies of critical importance.

4.7.1.1 DFT Geometries

Structural models for **2a**, **3a**, and **4a**, [transition structures, intermediates, full system (Ar = 3,5-di-*tert*-butylphenyl) and model systems (Ar = H)] were optimised at the B3LYP-D3/6-31G(d) level,[154, 157-158] with diffuse functions on the boron, oxygen and fluorine atoms and the inner four porphyrin nitrogen atoms added. In this method, empirical dispersion is included using Grimme's original D3 function.[158] Solvent effects were included using the polarisable continuum model[159] utilising the standard parameters for chloroform and dichloromethane and the solute-solvent dispersion interaction[160] included. Integrals were calculated using an "ultrafine" grid and tight convergence criteria for the SCF. This method and level of theory is referred to throughout this Chapter as B3LYP-D3/6-31(+)G(d). Coordinates for **b** enantiomers were generated from the related **a** species by internal reflection.

4.7.1.2 DFT intrinsic barriers to BAR in simple molecules

Barriers to linearisation were calculated by first geometry optimising the coordinates of each A-2 species at the level of theory indicated in Table 4.10 with free energies corrected for using the harmonic-oscillator approximation giving $G^{A\text{-}2}$. The L^1–M–L^2 angles were then changed to 180° and the structures re-optimised under the constraint of maintaining a L-2 geometry at M and conserving any torsional relationships at L with the resulting free energies $G^{L\text{-}2}$. The barriers are then simply $\Delta G^{\ddagger} = G^{L\text{-}2} - G^{A\text{-}2}$.

The strain energy DFT calculations of the porphyrin unit for complexes **2a**, **3a**, and $\textbf{TS}_{ak}^{R}$ that comprise Table 4.11 were calculated by first optimising model compounds (Ar = H) of each system at the B3LYP-D3/6-311G(d,p) level of theory. A reference free-base porphyrin system was generated from **1** to reproduce the same bonding topology exhibited by

B₂OF₂pqx stereoisomers. This corresponds to structure **1ref** as shown in Figure 4.22 which is a second order transition-state structure (see Chapter 6).

The optimised B₂OF₂pqx structures were then modified to remove the B₂OF₂ unit and replace it with two hydrogen atoms with placements similar to those in **1ref**. A geometry optimisation to a second order transition-state structure was then conducted with all atoms of the molecule frozen except for the added hydrogen atoms. Given these structures do not correspond to true critical points on the potential energy surface, only the uncorrected electronic energies were subsequently used. The energy differences are then simply $\Delta E = E$(strained structure) – E(**1ref**) and the tetragonal elongation distances $\Delta d = d$(strained structure) – d(**1ref**)
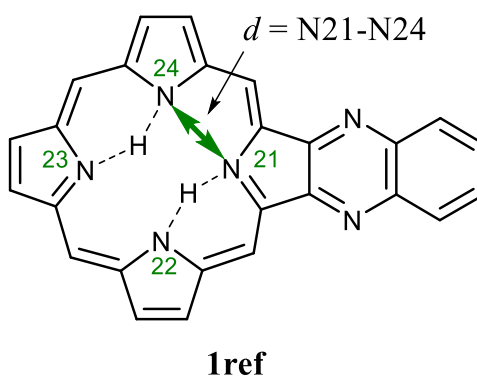
1ref

Figure 4.22 Reference molecule **1ref** derived from **1** to have the same bonding topology as the B₂OF₂pqx stereoisomers. This corresponds to a second order transition-state structure. The distance d indicated, is equal to the internuclear distance between atoms N21 and N24.

4.7.1.3 <u>DFT Reaction Pathways, Transition States and Reaction Rates</u>

4.7.1.3.1 DFT B−O−B bond-angle reflexion (akamptisomerisation) transition structure energies

The BAR transition state linking **2a** and **3a** on the model system with Ar = H was optimised at the B3LYP-D3/6-31(+)G(d) level. The transition state, called $\mathbf{TS_{ak}^{R}}$, is shown in detail in Figure 4.10.

4.7.1.3.2 DFT B−O−B coupled torsions transition structure energies

The transition state energy dependence upon the B−O−B angle for Hula-twist-like mechanism was calculated as follows and using an unsubstituted and high symmetry *transoid* B₂OF₂-porphyrin as a workable model.

In this general case, for a fixed B−O−B angle, the resulting *constrained* transition state exists when the oxygen atom moves through the quasi-plane of the macrocycle. In geometric terms, the following constraints must be met:

1. The *meso* carbons C-10 and C-20 and the oxygen atom must be *collinear* and lie on the C_2 axis as shown in Figure 4.23,

2. the vector formed by joining C-5 and C-15 must intersect the C_2 axis at right angles as shown in Figure 4.23, and

3.
4. the oxygen atom is located at this intersection point when, and only when; the B−O−B bond is linear.

For all non-linear B−O−B angles the model exhibits the C_2 point group. In the linear B−O−B case, the model exhibits the C_{2h} point group.
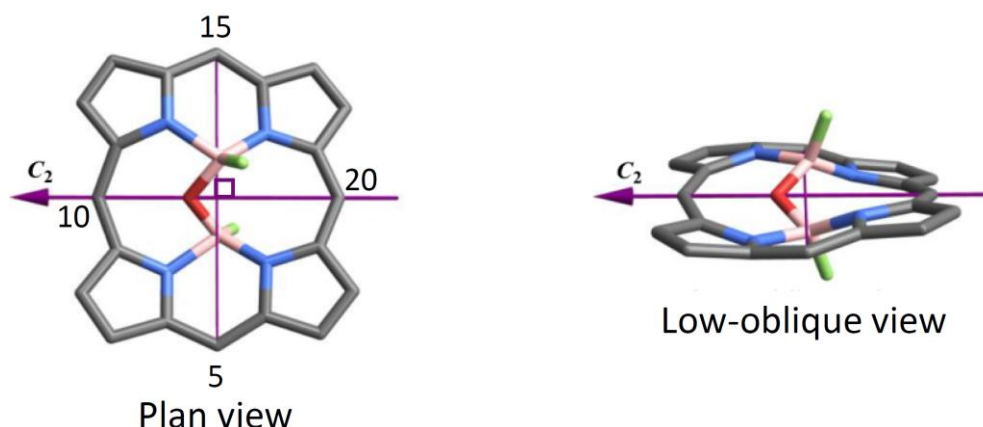
Figure 4.23 Geometric constraints for performing hula-twist-like constrained potential energy surface scans. Plan and low-oblique views of a transition state structure. The oxygen and two *meso* carbons, C-10 and C-20, lie on the C_2 axis represented by the purple arrow. The second line joining *meso* carbons C-5 and C-15, intersects the C_2 axis at right angles. Hydrogen atoms omitted for clarity. Atomic colours: white- H, peach- B, grey- C, blue- N, red- O, green- F.

Within Gaussian16[153], this was implemented by defining a dummy atom (X) to lie at the intersection of the orthogonal [C-10 – C-20] and [C-5 – C-15] vectors and to freeze the linear bonds [C-10 – O – C-20] and [C-5 – X – C-15] and the right angles [C-10 – X – C-20] and [C-5 – X – C-15]. The angle [B – O – B] was then scanned from 175° to 100° at the B3LYP-D3/6-31(+)G(d), AM1, PM3, and PM6 levels. For each of the different methods, the linear B−O−B case was calculated separately. The results are shown in Figure 4.12a and Figure 4.12c.

Single-point energy calculations at the B3LYP-D3/6-31(+)G(d) optimised geometries at 180° and 147° degrees give MP2/cc-pVDZ, MP2/cc-pVTZ, CCSD/cc-pVDZ, and

CCSD(T)/cc-pVDZ instability energies for the double torsion pathway of 13.7, 15.9, 13.7, and 12.3 kJ mol^{-1}, respectively, in qualitative agreement with the B3LYP-D3/6-31G(d) optimised value of 22.7 kJ mol^{-1}.

4.7.1.3.3 DFT in-plane B(F)OB(F) bond-walk rotation (strepsisomerisation) transition structure energies

Optimised geometries and free energies for solution phase structures of the *meso*-unsubstituted analogues of **2a**, **2b**, **3a**, and **3b**, as well as their interconnecting transition states obtained by rotating the B(F)OB(F) group inside the porphyrin macrocycle via a "bond walk" mechanism (*strepsisomerisation*, see Section 2.4.26), were calculated at the B3LYP-D3/6-31+G(d) level with the diffuse functions anticipated as critical to accurately model the bond-breaking/making process. Solvent effects were included using the polarisable continuum model[159] using the standard parameters for dichloromethane and the solute-solvent dispersion interaction[160] included. The results are given in Section 4.5.4.

4.7.1.3.4 DFT analysis of bimolecular processes involving fluoride dissociation and rearrangements of cationic species

The free energies for fluoride ion dissociation from the *meso*-unsubstituted analogues of **2** - **4** under experimental conditions (conditions of synthesis or just simple solutions) were calculated at the B3LYP-D3/6-31+G(d) level with the diffuse functions critical to accurately model the anionic species. Solvent effects were included using the polarisable continuum model[159] using the standard parameters for dichloromethane and CHCl$_3$ (representing CDCl$_3$) and the solute-solvent dispersion interaction[160] included.

For the following reactions, the un-ionised starting compounds are generalised as "P−F" where "P" represents the isomers of (B$_2$OFpqx)$^+$. Reaction rates are calculated for the dissociation-promoted pseudo-reaction:

$$\text{P-F}_{(solv)} \rightarrow \text{P}^{+\ddagger}_{(solv)} \qquad \text{with} \qquad e^{-\Delta G^{\ddagger}/kT} = K = \frac{[P^{+\ddagger}]}{[P–F]} \qquad (1)$$

where P$^{+\ddagger}$ represents the six different transitions states named in colour in Figure 4.18. Using the relationships for the activation of P$^+$ to a transition state

$$\text{P}^+ \xrightarrow{K_c} \text{P}^{+\ddagger} \qquad \text{with} \qquad e^{-\Delta G_c/kT} = K_c = \frac{[\text{P}^{+\ddagger}]}{[\text{P}^+]} \qquad (2)$$
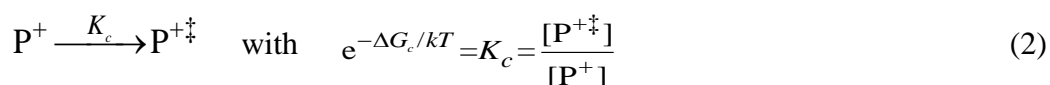
(values of ΔG_c shown in Figure 4.18) is used to solve each concentration-dependent reaction rate. As the Gaussian package calculates free energies with respect to a gaseous standard

state, the molar volume being 22.4 L, whilst in solution the standard concentration is 1 M, a correction to $\Delta G^{\ddagger}$ of $RT \ln (22.4 \text{ L} / 1 \text{ L}) = 8.0$ kJ mol^{-1} must be added. The method for different reactions is then:

<u>Uncatalysed simple dissociation:</u>

$$P{-}F_{(solv)} \rightarrow P^+{}_{(solv)} + F^-{}_{(solv)}$$
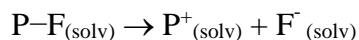
The barriers are simply calculated as $\Delta G_{products} - \Delta G_{reactants} + \Delta G_c$

<u>HNEt$_3{}^+$ catalysed dissociation:</u>

$$P{-}F_{(solv)} + H{-}NEt_3{}^+{}_{(solv)} \rightarrow P^+{}_{(solv)} + [F{-}H{-}NEt_3{}_{(solv)}]$$
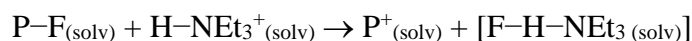
As HNEt$_3{}^+$ is a by-product of synthesis then $[HNEt_3{}^+] = 2[P{-}F]$. With $[F{-}H{-}NEt_3{}^+] = [P^+]$, the rate for this reaction becomes

$$K_F = [P^+]^2 / 2[P{-}F]^2$$

Combining this with (2) gives

$$K_F = ([P^{+\ddagger}]/ K_c)^2 / 2[P{-}F]^2$$

and combining with (1) and simplifying gives

$$K = K_c(2K_F)^{\frac{1}{2}} = e^{-\Delta G^{\ddagger}/RT}$$

so that

$$\Delta G^{\ddagger} = \Delta G_c + \Delta G_F - \tfrac{1}{2}RT \ln 2$$

<u>BF$_3${-}NEt$_3$ catalysed dissociation:</u>

$$P{-}F_{(solv)} + BF_3{-}NEt_{3(solv)} \rightarrow P^+{}_{(solv)} + BF_4{}^-{}_{(solv)} + NEt_{3(solv)}$$
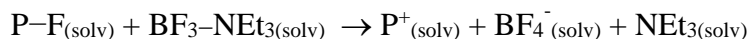
with NEt$_3$ modelled as NMe$_3$,

$$e^{-\Delta G_F/kT} = K_F = \frac{[Me_3N][BF_4^-][P^+]}{[Me_3N{-}BF_3][P{-}F]} = \frac{[Me_3N]}{[Me_3N{-}BF_3][P{-}F]}[P^+]^2$$

From the reaction conditions, $[Me_3N{-}BF_3] = 0.16$ M and can be treated as a constant, $[Me_3N] = 0.011$ M as this was added in slight excess and is taken as a constant, $[P{-}F] = 0.0042$ M is also taken as a constant (this assumes that only a small fraction of it is dissociated). Combining with (1) and (2) gives

$$K_F = \frac{[Me_3N]}{[Me_3N-BF_3][P-F]}\left(\frac{[P^{+\ddagger}]}{K_c}\right)^2 = \frac{[Me_3N][P-F]}{[Me_3N-BF_3]}\frac{1}{K_c^2}\left(\frac{[P^{+\ddagger}]}{[P-F]}\right)^2 = \frac{[Me_3N][P-F]}{[Me_3N-BF_3]}\left(\frac{K}{K_c}\right)^2$$

$$K = K_c\left(K_F\frac{[Me_3N-BF_3]}{[Me_3N][P-F]}\right)^{1/2} \quad \text{so that}$$

$$\Delta G^{\ddagger} = \Delta G_c + \frac{1}{2}\Delta G_F - \frac{kT}{2}\ln\frac{[Me_3N-BF_3]}{[Me_3N][P-F]} = \Delta G_c + \frac{1}{2}\Delta G_F + 10.0 \text{ kJ mol}^{-1}.$$

Uncatalysed F⁻ S$_N$2 reactions on **4a** and **4b**:

$$P-F_{(solv)} + F^-{}_{(solv)} \rightarrow [F-P-F]^-{}_{(solv)} \rightarrow \textit{iso}\text{-}P-F_{(solv)} + F^-{}_{(solv)}$$
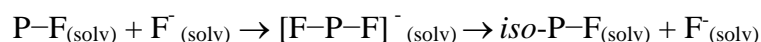
The *cisoid* B(F)OB(F) motif in **4a** and **4b** exposes the two boron atoms to the attack from the opposite face of the macrocycle to give either **2a**, **2b**, **3a**, and **3b**. The reaction was modelled by optimising the stepwise approach and withdrawal of F⁻ and fitting an exponential function to the ΔG data points to estimate the energy of each P−F$_{(solv)}$ + F⁻$_{(solv)}$ at infinite separation and the maximum energy of the bound complex [F−P−F]⁻$_{(solv)}$.

Uncatalysed F⁻ loss with counter-facial addition reactions on **4a** and **4b**:

$$P-F_{(solv)} \rightarrow P^+{}_{(solv)} + F^-{}_{(solv)} \rightarrow \textit{iso}\text{-}P-F_{(solv)}$$
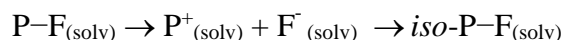
This uncatalysed reaction converting **4a** and **4b** to **2a**, **2b**, **3a**, and **3b** provides an estimated *upper limit* for the barrier to this depletion mechanism.

4.7.1.4 <u>DFT Potential Energy Surface scans</u>

Optimised geometries and enthalpies for the gas-phase structures of the *meso-*unsubstituted analogues of **2a** and **3a** were calculated at the B3LYP-D3/6-31G(d) level under the constraint of fixed B−O−B bond angles ranging from 95° to 180° in 27 equal angle increments to generate the two halves of the potential energy surface along the B−O−B BAR reaction coordinate shown in Figure 4.5. A similar scan was performed for the interconversion of **4a** and **5a** with the B–O–B angle of unbound **5a** structure arbitrarily assigned as 220° where the gradient of the PES in this region was lowest. The resulting

surfaces are shown in Figure 4.5. How the molecular structures evolve from **2a** and **3a** is shown in **E_File_4**.

A PES scan of the strepsisomerisation cycle for the purposes of visualisation was performed using the PM7 semi-empirical method.[161] The method used the intrinsic reaction coordinate (IRC) method[162-163] starting with the four strepsisomerisation transition-state structures of the model (Ar = H) system in the gas phase and calculating steps in both forward and reverse directions to the respective local minima. Analytical force constants were calculated at every step with step sizes of 0.15 a_0 and with up to 72 steps. The eight segments of the PES were assembled in order and scaled to match the DFT energy range (E_{max}[DFT] = 1.15 E_{max}[PM7]).

4.7.1.5 DFT ^{1}H and ^{19}F NMR spectra

NMR calculations using the optimised **2a** and **3a** geometries for the full systems (Ar = 3,5-di-*tert*-butylphenyl) (see Section 4.7.1.1) were carried out under the Gauge-Independent Atomic Orbital method[164] using the B3LYP functional[154] with the 6-311++G(2d,p) basis set applied to all atoms of the extended macrocycle and the 6-31+G(d,p) basis set to all *meso*-aryl atoms. Empirical dispersion was included using Grimme's original D3 function.[158] Solvent effects were included using the polarisable continuum model[159] with the standard parameters for chloroform and the solute-solvent dispersion interaction included.

Isotropic chemical shifts for ^{1}H and ^{13}C are reported relative to the calculated ^{1}H and ^{13}C shifts of tetramethylsilane (TMS) and those for fluorine relative to the calculated ^{19}F shift of trifluorochloromethane (CF$_3$Cl). The TMS and CF$_3$Cl geometries were optimised at the B3LYP-D3/6-311++G(2d,p) level. Integrals were calculated using an "ultrafine" grid and tight convergence criteria for the SCF. The TMS and CF$_3$Cl NMR calculations were conducted at the same level of theory as for the porphyrins using identical settings.

4.7.1.6 DFT Absorption and Circular Dichroism spectra

Electronic absorption and circular dichroism spectra were modelled using the optimised geometries for **2a** and **2b** only as their enantiomers give identical results except for the opposite sign of the calculated rotatory strengths. At least 8 electronic transitions were determined, the first two occurring in the Q-band region while the remainder are the Soret and nearby bands.

Two different asymptotically corrected density functionals, CAM-B3LYP[165-166] and ωB97XD[167] were used as well as the configuration-interaction singles[168] (CIS) method, and three different basis sets were used. The ωB97XD results include corrections for intramolecular dispersion interactions, while for the largest basis set used, 6-31++G(2d,p), the D3 dispersion correction[158] was added to CAM-B3LYP calculations.

Solvent corrections were also sometimes applied using the polarisable continuum model[159] with the "noneq=read/write" commands to properly treat charge-transfer bands. The implicit solvent was set to either $CHCl_3$ or else to 2-MeTHF parametrised as THF with the dielectric constant changed to $\varepsilon = 6.97$. Calculations were also performed both for the full molecule with Ar = 3,5-di-*tert*-butylphenyl substituents and for simplified molecules with Ar = H

4.7.1.6.1 Room temperature spectra

Room temperature UV-Vis and simultaneous absorption (ABS) and circular dichroism (CD) spectra of purified chiral stationary phase HPLC fractions **Fr1**, **Fr2**, **Fr3** and **Fr4** in solvents of different polarity have been reported.[7] Figure 4.24 and Figure 4.25 show spectra for **Fr1** and **Fr4**, respectively, observed using a range of sample concentrations in the 5-20 μM region. The spectra of **Fr1** and **Fr2** were not significantly different from each other except for a sign reversal of the CD, indicating that these fractions are enantiomers, and hence spectra are shown only for **Fr1**. Similarly, **Fr3** and **Fr4** are also enantiomers. The spectra show little variation with solvent and concentration, indicating that the intrinsic CD of isolated molecules is being measured rather that the CD of oligomeric aggregates. Porphyrinoid compounds, including chlorophylls, often form oligomers in solutions at the relatively high concentrations often used to measure weak spectral components. These oligomers typically exhibit CD associated with exciton coupling within the oligomers that is much stronger than monomeric (molecular) CD. The presence of the B(F)OB(F) group coordinated in the middle of the porphyrin would prevent π-stacking and hence inhibit oligomerisation, the likely cause of the absence of the expected spectral signatures at the highest concentrations used.

In Figure 4.26, the observed room temperature spectra for all isomers in 2-MeTHF are compared. The CD spectra in the Q-band region show distinctly different patterns for each fraction varying in the sign and relative amplitudes of the two Q-band component peaks. This distinction is used to assign structures to the isolates in Section 4.3.6. Figure 4.26 shows

that the Soret-band region of the spectrum which also displays unique patterns for each fraction and hence could (in principle) facilitate assignment. However, using the available computational techniques it was not possible to authoritatively predict these patterns.
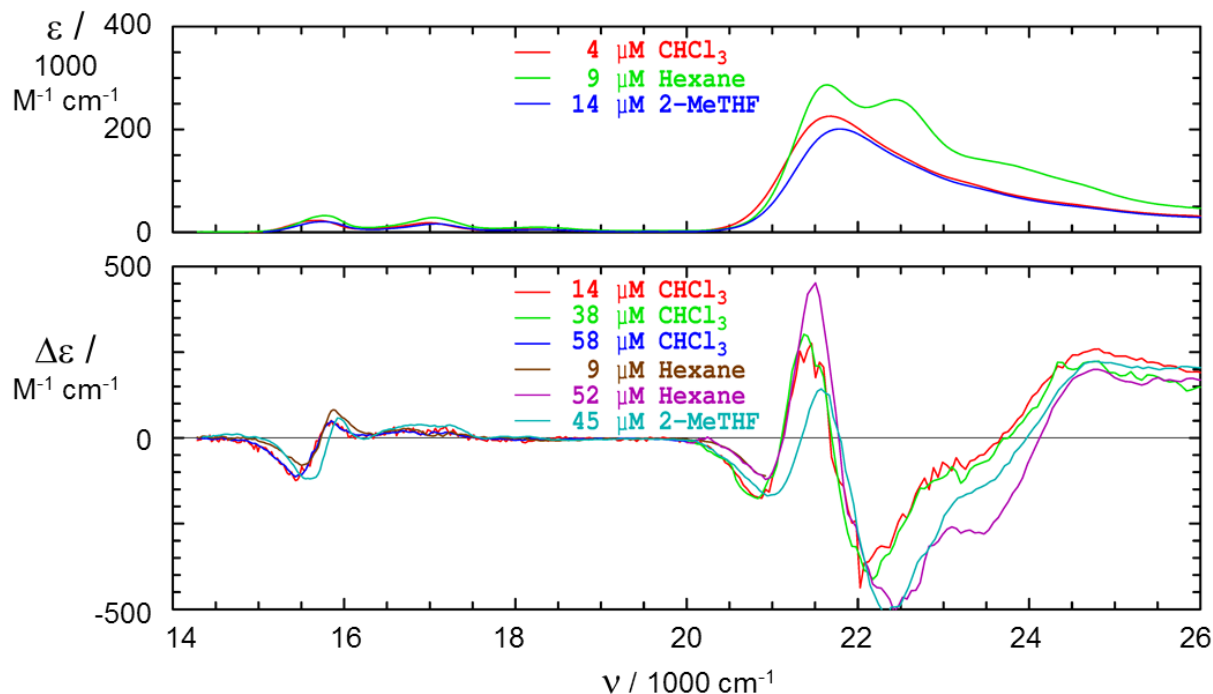
Figure 4.24 Reproduced[7] comparative 298 K experimental CD spectra $\Delta\varepsilon$ and ABS spectra ε of **Fr1** in CHCl$_3$, hexane and 2-MeTHF obtained using various porphyrin concentrations in different solvents.
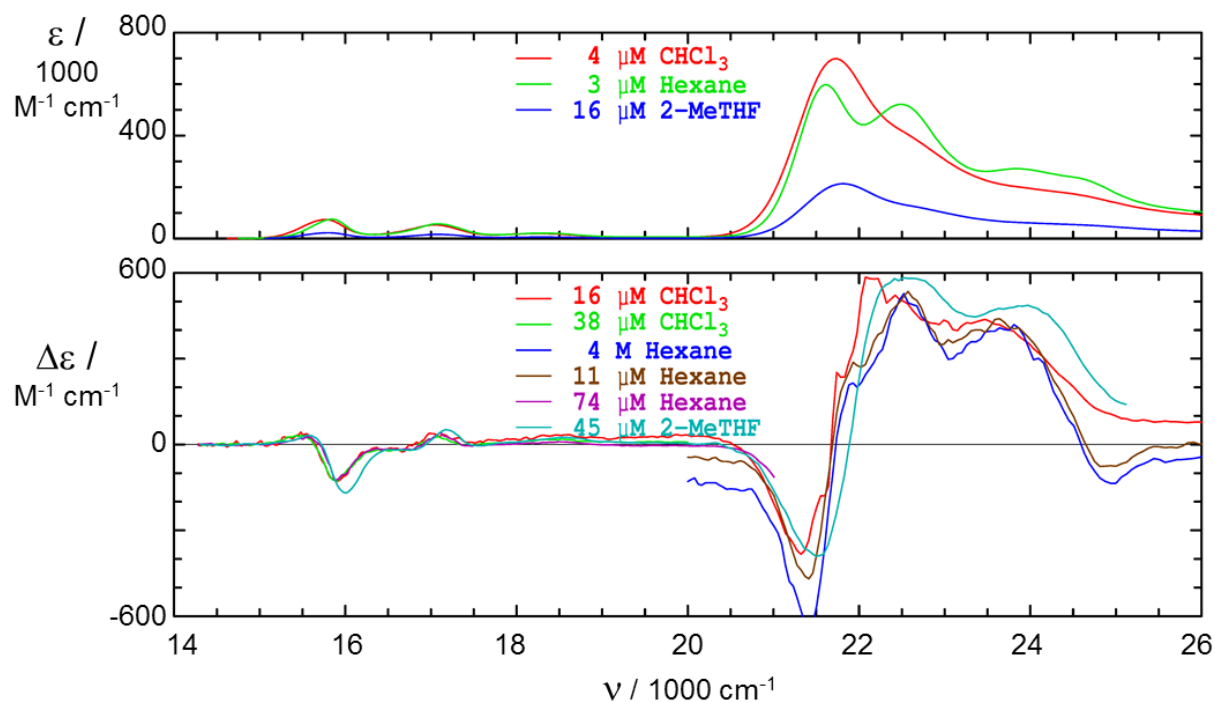
Figure 4.25 Reproduced[7] comparative 298 K experimental CD spectra $\Delta\varepsilon$ and ABS spectra ε of **Fr4** in CHCl$_3$, hexane and 2-MeTHF obtained using various porphyrin concentrations in different solvents.
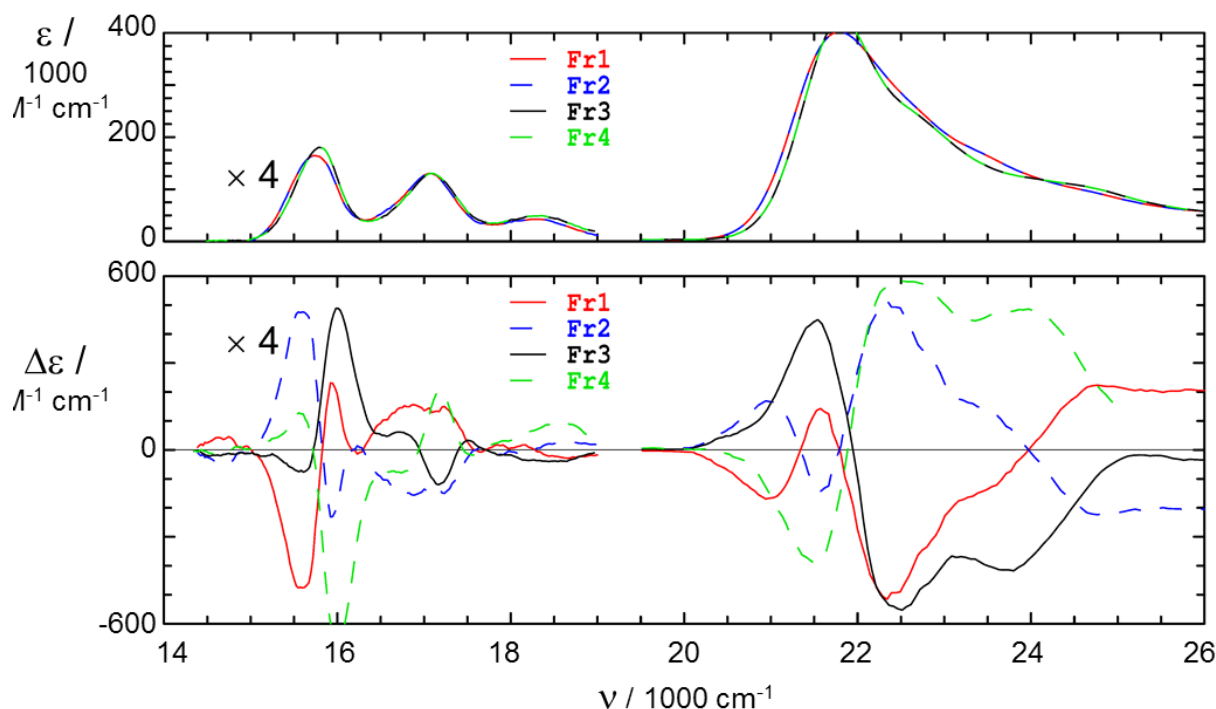
Figure 4.26 Reproduced[7] comparative 298 K experimental CD spectra $\Delta\varepsilon$ and ABS spectra ε of **Fr1** - **Fr4** in 2-MeTHF obtained using various porphyrin concentrations (14-45 μM).

4.7.1.6.1 Low temperature spectra

Low-temperature ABS, CD, and magnetic circular dichroism (MCD) spectra have been reported.[7] Figure 4.27 reproduces the observed low-temperature ABS, CD, and MCD spectra. The CD signal is very weak and hence shows considerable noise under the experimental conditions, *i.e.*, conditions optimised to minimise any possibility of oligomerisation and to provide high reliability ABS and MCD data. The primary purpose of these low-temperature measurements is to be able to separate the observed ABS spectra into its orthogonally polarised components. Two independent and orthogonally polarised electronic transitions, usually named Q_x and Q_y, are expected in the Q-band region, whilst the related bands B_x and B_y are expected in the Soret-band region. However, the Soret bands will mix with charge-transfer states arising from combinations of the porphyrinic N bands with transitions that move charge between the macrocycle ring and the fused quinoxaline group.[127]
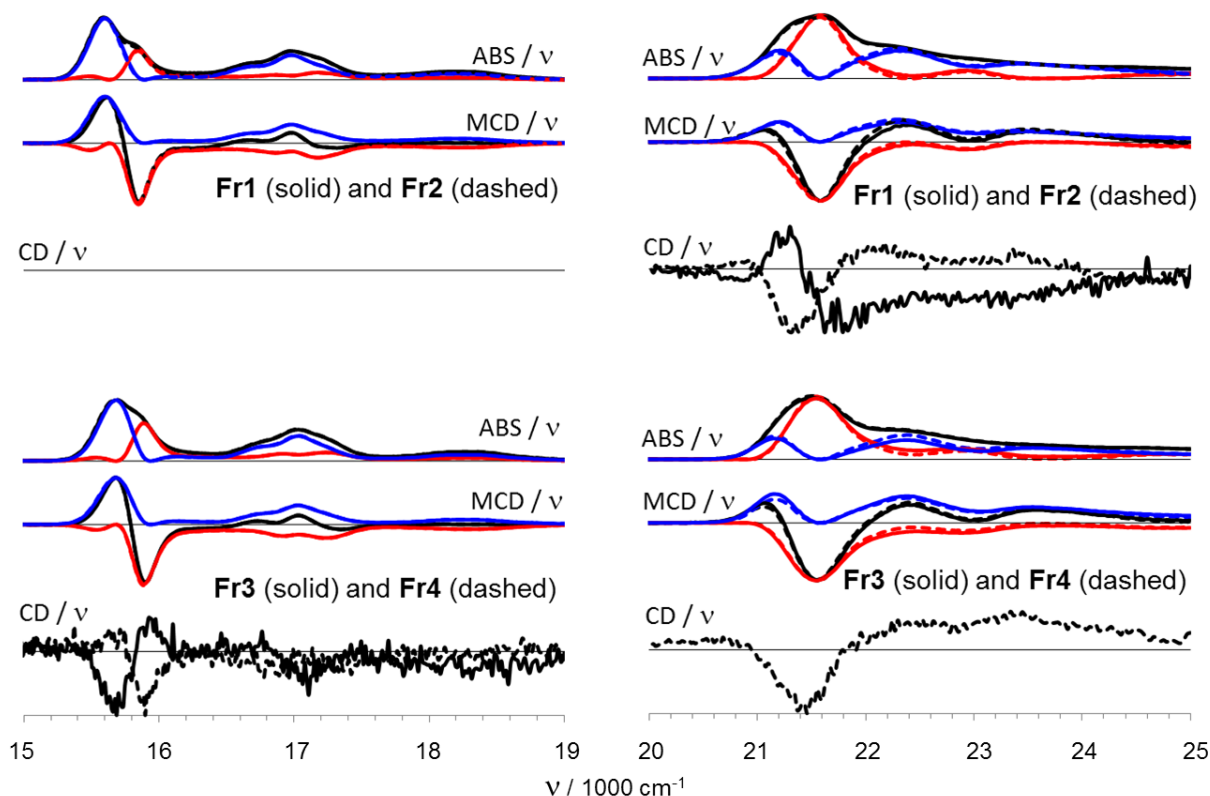
Figure 4.27 Reproduced[7] comparative of relative experimental ABS, MCD, and CD low temperature spectra taken in 2-MeTHF at 55-68 K in the Q-band (60-80 μM) (left) and Soret band (4 – 40 μM) (right) regions for **Fr1** - **Fr4**. Analytical data inversion methods are used to convert the ABS and MCD raw spectra (black) into two orthogonal spectral components (red and blue) on the assumption of maximal intensity in the weaker band (the band coloured red). Concentrations used: 37 μM for **Fr1**, 43 μM for **Fr2**, 27 μM for **Fr3**, and 28 μM for **Fr4**.; maximum ABS extinction always < 1.2.

In MCD experiments, component states with one polarisation appear with +*ve* sign whilst those of the orthogonal polarisation appear with a –*ve* sign. When combined with ABS data, this observation allows the ABS to be deconvolved into two its *x* and *y* polarisation components. This was achieved not by band fitting as is the usual practice, but instead using an analytical data inversion scheme recently introduced.[169] In principle, such an analytical data inversion does not appear to be feasible, as it requires two more parameters to be deduced than there are experimentally observed quantities. However, it has been shown[169] that one equation of constraint is already known so that only one unspecified parameter remains. Further, this remaining variable can be expressed in a bounded form to which deconvoluted band shapes are insensitive.[169] Spectral components deduced using the minimum-allowable value for this unknown variable are shown in the figure, coloured red for the weaker component and blue for the stronger one. The assumption used attributes the maximum possible intensity to the red components, with alternate values for the arbitrary parameter serving to downplay the absorption strength of the red bands without significantly

altering the shapes of the red and blue components. As a result, the band centres of the extracted components are reliably determined and may be compared to computational predictions. The deduced properties of the absorption bands are listed in Table 4.12 whilst Table 4.13 gives the ratios of the MCD to ABS susceptibilities B/D. The ratios of $-B_y/B_x$ for each band are also given, these are unity for a 4-fold symmetric porphyrin.

For each sample, the Q-band origin region of 15400-16200 cm^{-1} decomposes into an intense lower-frequency band (blue) and a weak high-frequency shoulder (red) shifted by 250 cm^{-1} for **Fr1** and **Fr2** and by 220 cm^{-1} for **Fr3** and **Fr4**. The analytical inversion analysis does not allow for comparisons between molecules, however, and it is uncertain as to whether the intense band coloured blue for **Fr1** and **Fr2** corresponds to the band coloured blue for **Fr3** and **Fr4**.

In the Soret-band region, the observed ABS spectra for all fractions are deconvolved into four spectral components: one intense component of one polarisation (coloured blue) and three components of the alternate polarisation (coloured red). As the analyses of the Soret and Q-band regions are essentially independent of each other, there is no proscribed relationship between the state polarisations coloured the same way for each band.

Table 4.12 Deconvoluted properties of observed absorption spectral bands.[a]

Band	Pol.	Fr1 ν_{00}	Fr1 $\overline{\nu}$	Fr1 f	Fr2 ν_{00}	Fr2 $\overline{\nu}$	Fr2 f	Average ν_{00}	Average $\overline{\nu}$	Average f
S1	x	15600	16382	0.125	15600	16371	0.129	15600	16377	0.127
S2	y	15850	16467	0.05	15850	16470	0.048	15850	16469	0.049
S2-S1		250	85		250	99		250	92	
S3	x	21210		0.25	21210		0.25	21210		0.250
S4	y	21600	21906	0.93	21600	21906	0.93	21600	21906	0.930
S5	x	22320		0.51	22320		0.62	22320		0.565
S6	x	~23600		0.5	~23600		0.5	~23600		0.500

Band	Pol.	Fr3 ν_{00}	Fr3 $\overline{\nu}$	Fr3 f	Fr4 ν_{00}	Fr4 $\overline{\nu}$	Fr4 f	Average ν_{00}	Average $\overline{\nu}$	Average f
S1	x	15670	16445	0.11	15670	16450	0.111	15670	16448	0.111
S2	y	15890	16539	0.057	15890	16535	0.056	15890	16537	0.057
S2-S1		220	94		220	85		220	90	
S3	x	21160		0.18	21160		0.19	21160		0.185
S4	y	21550	22050	1.05	21550	21950	1.003	21550	22000	1.027
S5	x	22370		0.38	22370		0.43	22370		0.405
S6	x	~23000		0.38	~23000		0.43	~23000		0.405

a: Polarisation directions x and y are arbitrary and not correlated between Q and Soret bands. ν_{00} is the band origin, $\overline{\nu}$ the average absorption energy, and f the deconvolution-assumption-dependent oscillator strength.

Table 4.13 MCD susceptibility ratios B/D, in Tesla^{-1}, and also $-B_y/B_x$.

. Fraction	Q band B_x/D_x	Q band B_y/D_y	Q band $-B_y/B_x$	Soret Band B_x/D_x	Soret Band B_y/D_y	Soret Band $-B_y/B_x$
Fr1	0.0287	–0.0755	1.03	0.00277	–0.00358	1.05
Fr2	0.0245	–0.0680	1.02	0.00214	–0.00280	0.94
Fr3	0.0056	–0.0114	1.04	0.00075	–0.00052	0.95
Fr4	0.0056	–0.0117	1.04	0.00061	–0.00054	0.92

The available (noisy) low-temperature CD spectra are also reproduced in Figure 4.27. Comparison of relative ABS, MCD, and CD spectra taken in 2-MeTHF at 55-68 K in the Q-band (60 – 80 μM) (left) and Soret band (4 – 40 μM) (right) regions for **Fr1** – **Fr4**. Analytical data inversion methods are used to convert the ABS and MCD raw spectra (black) into two orthogonal spectral components (red and blue) on the assumption of maximal intensity in the weaker band (the band coloured red). The peaks in the CD spectra correlate with the maxima of the deconvolved spectral components, indicating again that they relate to molecular features rather than to properties of aggregates. The qualitative shapes of the CD correspond to those observed at room temperatures (Figure 4.24 – Figure 4.26), indicating that these much-better-resolved features can be attributed to the peaks in the deconvolved low-temperature ABS/MCD spectra, for the purposes of spectral assignment.

5 The ML$_2$ stereogenic unit: the diversity of structural possibilities, failings of current stereochemical practice, and the broader implications for stereoisomerism

5.1 Summary

Recognising that an embedded ML$_2$ centre must be treated as a single stereogenic unit, oxygen-bridged oxabicyclo[*m.n.*1] systems are identified as a simple but large class of compounds that are akamptisomeric. A search of the Chemical Abstracts Service (CAS) database for such compounds was undertaken to investigate the performance of current IUPAC recommendations to akamptisomeric systems.

The search results based on various combinations of *m* and *n* show that, for small bicyclo systems, the current IUPAC recommendations are fit for purpose. For a range of intermediate to larger bicyclo systems, the IUPAC recommendations prove inadequate with special attention to akamptisomerism-related increased scope for stereoisomers being necessary.

Some of these compounds, examined in detail, are representative of a great diversity of both synthetic and natural products with applications ranging from materials science to pharmacological agents.

As the IUPAC recommendations provide the foundation on which much chemistry related infrastructure is built, the deficiencies exposed here have immediate implications for automatic naming algorithms, digital representation and manipulation of structures, and patents and pharmaceutical-registration requirements, to name but a few.

A deeper and more significant finding is that this deficiency exposes a fundamental weakness in the overall approach to stereoisomerism to date. Historically, the subject has adapted to the discovery of novel phenomena by introducing novel, but *ad hoc*, amendments. Whilst this approach provided immediate practical solutions, it is not capable of predicting what undiscovered possibilities remain. What has been lacking is an all-encompassing conceptual framework allowing for a systematic description for what is possible. It is proposed here that the Polytope Formalism of stereoisomerism can meet that need.

5.2   Introduction

Chapter 3 of this Thesis concluded with the mathematical basis for how a novel bond-angle reflexion (BAR) based class of stereoisomerism, called akamptisomerism, could manifest in embedded ML_2 systems. Chapter 4 described the experimental demonstration of akamptisomerism in B(F)OB(F)-quinoxalinoporphyrins (B_2OF_2pqx) where both the geometrical relationships (stereoisomerism) and the associated concerted unimolecular stereoisomerisation ($R_{st}^c 1$) were exclusively BAR based. Attempts to describe, name, and depict the relationships inherent to these BAR-related porphyrin stereoisomers revealed that the current IUPAC recommendations for terminology,[4] nomenclature,[3, 170] and 2D depiction,[91] respectively, fail to do so.

Originally, akamptisomers were presented as an exotic form of *conformational* isomerism (Chapter 4), but this view now needs to be broadened as most compounds investigated can be considered instead as *configurational* isomers.

Whilst the geometric relationships inherent to akamptisomers can, in principle, occur at any bicoordinate centre, in Chapter 4 it was shown that embedding the ML_2 stereogenic unit to give an external frame of reference was critical to allowing these geometric relationships to manifest. The essential molecular features needed to exhibit akamptisomerism include treating the entire ML_2 grouping as a *single stereogenic unit* and having this stereogenic unit connect to a larger molecular structure to provide the embedding. Figure 5.1a shows this where ML_2 is now represented as X(A)–Y–Z(B) with A and B representing any number of possible substituents to X and Z, respectively. X and Z have coordination numbers equal to or greater than two, and Y is of *A*-2 coordination. Importantly, A and B are not part of the embedding. For embedded molecular systems, at least one topological connection (black arc) must connect X and Z independent of Y. Whilst this single topological connection is a *minimal requirement*, at least one additional topological connection (green arc) serves to provide a practical embedding as a more "general case".

(a)

(b)

ML$_2$ = X(A)−Y−Z(B)

oxabicyclo[*m.n*.1]
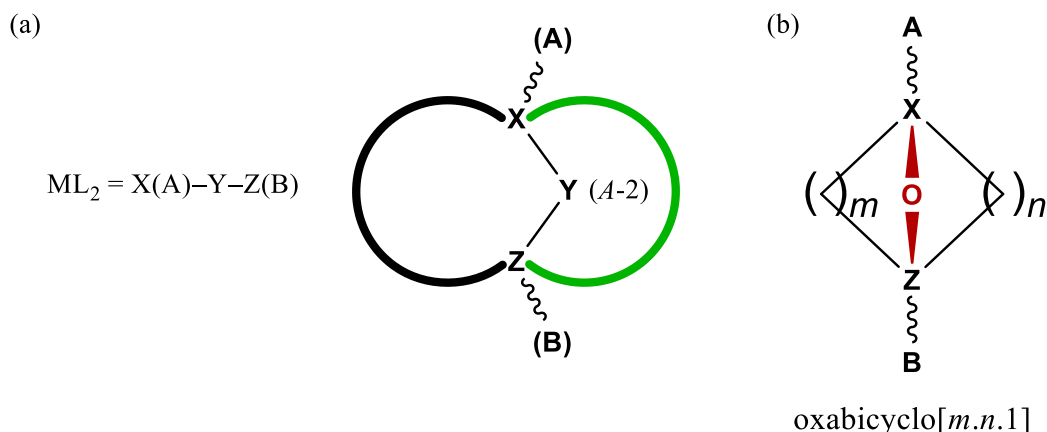
Figure 5.1 General depiction of an embedded BAR-based stereogenic unit. (a) Assigning the ML$_2$ unit as X(A)–Y–Z(B), where A and B are possible substituents to X and Z, respectively. X and Z have coordination numbers equal to two or greater, and Y is of *A*-2 coordination. A and B are not part of the embedding. For molecular systems, at least one topological connection (black arc) must connect X and Z independent of Y. At least one additional topological connection (green arc) serves to provide a practical embedding. (b) General formula for oxygen-bridged oxabicyclo[*m.n*.1] compounds embodying the form shown in (a).

A broad class of compounds that conforms to these general features is the oxygen-bridged oxabicyclo[*m.n*.1] class as shown in Figure 5.1b The hypothesis, examined herein, is that the IUPAC recommendations for nomenclature[3, 170] and 2D depiction[91] – the modern chemistry software embodying these and the basis for chemical intuition that it provides – must be deficient regarding its handling of properties associated with akamptisomerism. Utilising the Chemical Abstracts Service[171] (CAS) database and narrow search criteria for oxygen-bridged oxabicyclo[*m.n*.1] compounds, 443,807 known compounds are considered that fulfil the classification as akamptisomers. Of these compounds that were examined, 17,807 were found to be deficiently represented in some way.

As the records that comprise the CAS database utilise the IUPAC recommendations (though generally treated as "rules") this leads to deficiencies in structure representation. Some of the compound records found in the database search results, examined in detail, are representative of a great diversity of both synthetic and natural products with applications ranging from materials science to pharmacological agents.

5.3   Results

5.3.1 *Recognising compounds as akamptisomers*

In Chapter 4, a non-standard wedge-hash usage was required to faithfully depict the full scope of stereo-relationships exhibited by the B$_2$OF$_2$pqx compounds. The question is:

how general is this requirement for non-standard 2D depiction of akamptisomeric structures? A systematic examination of the 2D depiction and nomenclatural needs is undertaken with solutions given that concisely and unambiguously describe all stereo-relationships involving any embedded ML_2 stereogenic unit.

In Chapter 4, porphyrin akamptisomers **1** and the related forms **2** and **3a** (Figure 5.2) were identified[148] in porphyrazines; indeed, **1** – **3** all stem from related works on macrocycles.[145, 148, 172-173] In Figure 5.2, distinguishable isomers with geometries related by at least one akamptisomeric centre are named as **a** and **b**, *e.g.*, the porphyrins **1a** and **1b**. A pair of akamptisomers that are not chemically distinguishable, for example owing to high symmetry, are labelled simply by a number, *e.g.*, **2**; these are called *autakamptisomers* (see Section 2.4.4).
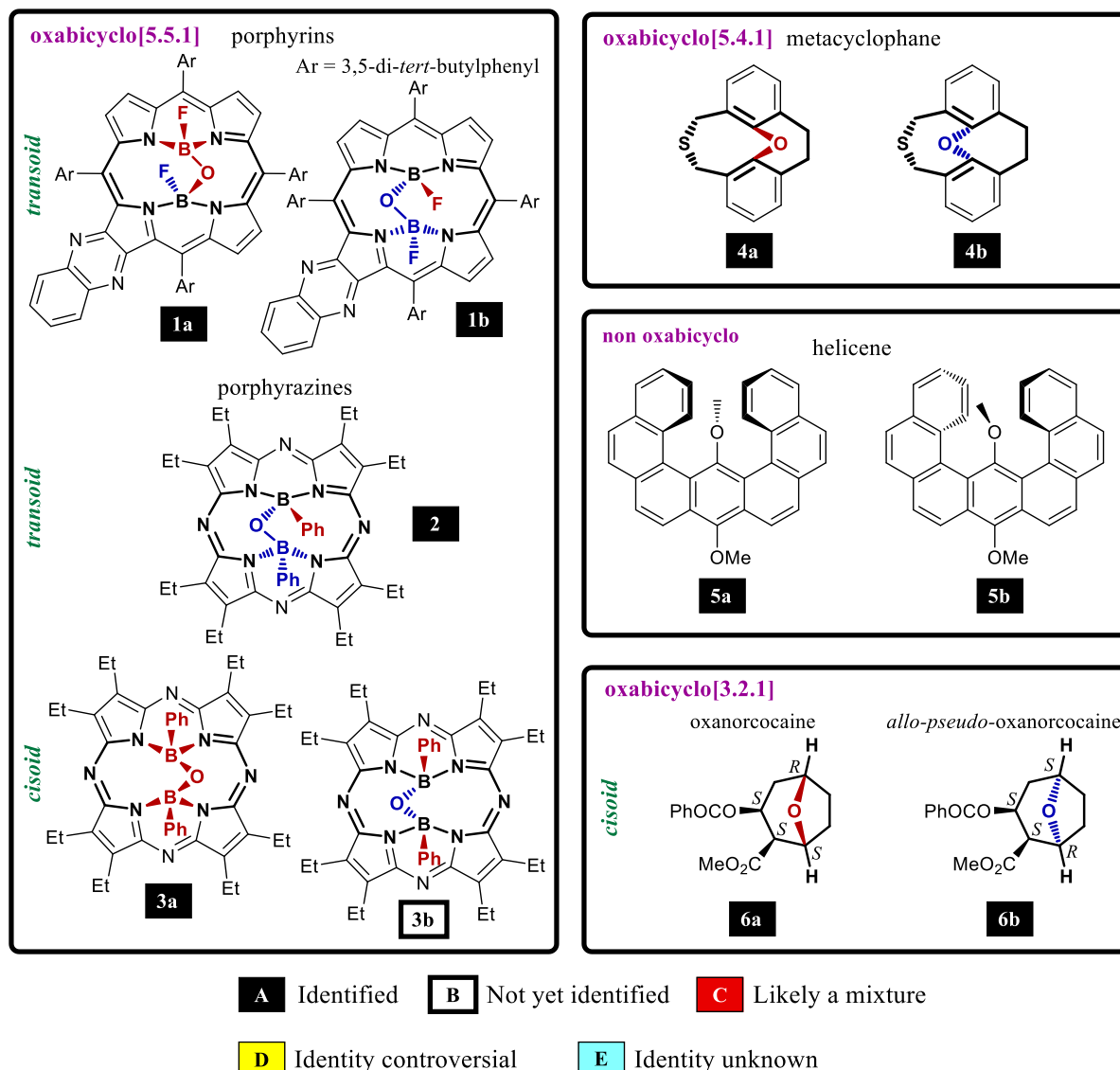
Figure 5.2 Conformational and configurational bicyclo[*m.n.*1] isomers with ML$_2$ stereogenic units. Oxygen-bridged bicyclo[*m.n.*1] systems featuring akamptisomeric centres are emphasised in bold and in colour, with atoms above a quasi-plane shown in red, with those below shown in blue. Labelled are *transoid* and *cisoid* configurations, as appropriate, of bridgehead substituents at akamptisomeric centres. The molecule numbers are shaded to indicate experimental identification, with categories of (A) identified, (B) not yet identified, (C) the observed compound is most likely some mixture of the shown possibilities, (D) ongoing discussion is suggested as to the identity of observed isomer(s), and (E) the identity of the observed compound remains unspecified; for molecules with *cisoid/transoid* variants, up to 8 isomers are conceptually possible, with all distinguishable unlisted isomers considered to be unlikely possibilities.

Figure 5.2 *continued.* Conformational and configurational bicyclo[*m.n.*1] isomers with ML$_2$ stereogenic units. Oxygen-bridged bicyclo[*m.n.*1] systems featuring akamptisomeric centres are emphasised in bold and in colour, with atoms above a quasi-plane shown in red, with those below shown in blue. Labelled are *transoid* and *cisoid* configurations, as appropriate, of bridgehead substituents at akamptisomeric centres. The molecule numbers are shaded to indicate experimental identification, with categories of (A) identified, (B) not yet identified, (C) the observed compound is most likely some mixture of the shown possibilities, (D) ongoing discussion is suggested as to the identity of observed isomer(s), and (E) the identity of the observed compound remains unspecified; for molecules with *cisoid*/*transoid* variants, up to 8 isomers are conceptually possible, with all distinguishable unlisted isomers considered to be unlikely possibilities.

Figure 5.2 *continued.* Conformational and configurational bicyclo[*m.n.*1] isomers with ML$_2$ stereogenic units. Oxygen-bridged bicyclo[*m.n.*1] systems featuring akamptisomeric centres are emphasised in bold and in colour, with atoms above a quasi-plane shown in red, with those below shown in blue. Labelled are *transoid* and *cisoid* configurations, as appropriate, of bridgehead substituents at akamptisomeric centres. The molecule numbers are shaded to indicate experimental identification, with categories of (A) identified, (B) not yet identified, (C) the observed compound is most likely some mixture of the shown possibilities, (D) ongoing discussion is suggested as to the identity of observed isomer(s), and (E) the identity of the observed compound remains unspecified; for molecules with *cisoid/transoid* variants, up to 8 isomers are conceptually possible, with all distinguishable unlisted isomers considered to be unlikely possibilities.

Previous to this work, a search for isomerism distinguished by "inversion" (actually a reflection – see Section 2.4.5) at oxygen centres was initiated in 1933 by von Bruchhausen *et al.*,[174] leading to the 1970 NMR-only identification of the oxabicyclo[5.4.1]metacyclophane isomeric pair **4** by Gordon and Gallagher,[175] compounds that can be taken as models for certain marine natural products and oligomeric lignans.[176-178] These are oxygen-bridged oxabicyclo[*m.n.*1] species, but alternatives such as **5** have also been considered;[179] however,

in no case were isolable products produced in these earlier works as the isomers rapidly interconverted at room temperature. Those studies were based on the notion that akamptisomers would be purely *conformational* isomers that, like atropisomers and pyramidal reflectomers, required special constraints to affect isolability. The particular oxygen-bridged oxabicyclo[5.4.1] scaffold in the akamptisomers **4a** and **4b** did not provide sufficient constraint to make them isolable, nor the sterically constrained helicene isomers **5a** and **5b**, whereas the macrocycle rings in **1a**, **1b**, **2** and **3a** from recent studies (Chapter 4, *etc.*[7, 148]) did prove sufficient. Note that, whilst these porphyrinoid macrocyclic rings embody internal oxygen-bridged oxabicyclo[5.5.1] systems (as emphasised in Figure 5.2), they have additional, critical, connections.

Oxygen-bridged oxabicyclo compounds are generally considered from the alternative perspective as *configurational* isomers, *i.e.*, stereoisomers classically requiring bond breakage and reformation to interconvert. This is exemplified by the isomeric pair oxanorcocaine[180-181] **6a** and *allo-pseudo*-oxanorcocaine[181-182] **6b**. These molecules have been described[181] as being "endowed with interesting cocaine-like activity" and feature an akamptisomeric C(H)–O–C(H) centre. They embody the small, all-convex, and highly constraining oxabicyclo[3.2.1] ring system, with any physically plausible akamptisomerisation between these isomers requiring bond breakage and reassembly. Further, akamptisomerisation of these all-convex, small-ring isomers necessarily involves Walden inversion (see Section A.51) of the two bridgehead stereocentres in addition to reflection at the bridging O atom; they serve to stress that the akamptisomerism phenomenon can be present in molecules *in addition* to other stereoisomeric properties.

Hence akamptisomers can show wide-ranging properties: some conformational isomers like **4a**, **4b**, **5a** and **5b** have been shown to undergo rapid concerted unimolecular akamptisomerisation at room temperature, some isolable molecules like **1a** and **1b** can undergo laboratory-controllable akamptisomerisation, and some like **6a** and **6b** are classical configurational isomers. Additionally, some of the structures shown like **3b** are currently purely conceptualised stereoisomers that may turn out to be energetically infeasible to produce (Chapter 4).

To overview the diversity of akamptisomerism, focus is now placed upon compounds with *concave* ring systems of intermediate size that exhibit properties intermediate between configurational and conformational isomerism: [5.3.1], [4.4.1], and [4.3.1] rings systems.

A substructure search of the CAS database (in "SciFinder classic") was conducted on differing sized oxygen-bridged oxabicyclo[*m.n*.1] structures according to the following criteria:

1. The bicyclo peripheral (non-bridging) atoms can be any element,
2. all bonds between the peripheral atoms are of arbitrary bond order,
3. the bicyclic system has the "lock ring" function applied to it. Whilst this criterion excludes the more "interesting" compounds **1 – 5** and **7 – 19**, given the limitations of the SciFinder search functionality, it provides a practical and expedient way to obtain a large number of search results that are strictly bicyclo in nature.

These search results were then refined by the additional constraints:

4. metal-free,
5. single component,
6. isotope free.

A typical output is shown in Figure 5.3. Full results are found in **E_Files_7** (see page xxv).

Many basic akamptisomeric oxygen-bridged oxabicyclo[4.4.1] compounds such as 1,6-epoxy[10]annulene[183-184] **7** and their derivatives have been synthesised,[183-188] with some such as **9** becoming the topic of patents.[189] Also of significance are the diastereoisomers **8**[190] containing two akamptisomeric centres that are identified as potential model compounds[191] for graphene oxide. Also considered are the partially saturated compounds **10 – 19** possessing widely diverse biological properties; unlike the aromatic compounds **7 – 10**, these are not particularly acid labile.[192] As detailed in Table 5.1, the example compounds **1 – 19** are structurally highly diverse and include various natural products, synthetic intermediates, *seco*-steroids, specific and generic anti-cancer agents, kinase inhibitors, IKKβ inhibitors, dopaminergic agents, estrogenic agents, antibacterial agents, dipolar switches, and basic chemical/biochemical and graphene oxide models. Further details of these compounds are listed in Table 5.2

(a)

(b)

oxabicyclo[4.4.1]

Figure 5.3   Typical output from CAS database substructure search to the given criteria.   (a) Search results output.   (b) Search criterion input for oxabicyclo[4.4.1].   The bold bonds (single and dashed) indicate the "lock rings" functionality.

.

Table 5.1 Representation of the stereochemistry at the akamptisomeric centre. Results for some example molecules are shown, considering what is known concerning the 3D structure from raw experimental data, how this is represented in 2D in the original literature and in the CAS database, software interconversion of isomer name and 2D structure, and immediate implications for patent specifications.

#	molecules	application	3D structure	2D literature	2D CAS	2D→name→2D conversion[a]	patent implications
1[j]	porphyrin	dipole switch	known	accurate	unspecified	fail	already addressed[193]
2[148]	porphyrazines	dipole switch	known	accurate	unspecified	fail	already addressed[193]
3[148]	porphyrazines	dipole switch	known	accurate	unspecified	fail	new possibilities[f]
4[175]	metacyclophane	model system	known	discussed	unspecified	fail[b]	
5[179]	helicene	model system	known	accurate	unspecified	fail[b]	
6[180-182]	cocaine derivatives	dopaminergic	known	accurate	accurate	succeed	none
7[183-184]	epoxy-10-annuene	model system	known	accurate	inconsistent[c]	succeed	
8[190]	diepoxy-14-annulene	graphene oxide model	known	accurate	inconsistent[c]	fail[b]	
9[189]	an epoxy[4.4.1]annulene	estrogenic agent	unspecified	accurate	overspecified[d]	fail[b]	new possibilities[f]
10[194]	*Toona* lathyrane	ovarian cancer	unspecified	accurate	overspecified[d]	fail[b]	new possibilities[f]
11[195]	*cis* badgerin	natural product	putative[e]	accurate	inconsistent[c]	partial	
12[196]	*trans* badgerin	natural product	putative[e]	accurate	inconsistent[c]	partial	
13[197]	austroliolide	antibacterial	unspecified	accurate	accurate	partial	
14[198-199]	*Santolina* germacrane	colon cancer	known	accurate	accurate	partial	
15[200]	jereisterol A	*seco*-steroid model	known	accurate	inconsistent[c]	fail	new concepts[g]
16[176]	substituted bicyclo-ether	polyether toxin synthetic intermediate	unspecified	accurate	overspecified[d]	partial	
17[201]	alterbrassicene A	IKKβ inhibitor	known	accurate	accurate	partial	
18[202-204]	staurosporine	kinase inhibitor	known	accurate	accurate	partial	new possibilities[h]
19[205-207]	cyclostreptin[i] acid	cytotoxic (as lactone)	known	accurate	accurate	partial	

a: Using ChemDraw,[208] MarvinSketch[209] and ChemSketch[210] to perform an automated [structure]→[name]→[structure] two-step procedure, see Table 5.2 for details. In summary: succeed- both akamptisomers correctly cycled, fail- neither akamptisomer successful cycled, partial- works for one akamptisomer only. Note that **4**, **7-9**, **15** fail as the software does not include existing IUPAC α/β stereodescriptors, other failures are purely associated with non-recognition of akamptisomeric variations. *b*: failure may not be considered relevant as, for these molecules, the isomerism can be considered as conformational, but the reasons for this failure apply also to other molecules listed that are stable isolable structures. *c*: 2D graphical representation of the stereochemistry is unspecified, but the associated compound name is accurate. *d*: provides more information than is reported. *e*: fully specified but controversy is revealed as to the actual stereochemical identity. *f*: possibilities exist for new patents for active akamptisomers. *g*: in August 2019, a search in CAS for "steroid" refined by "patent" yielded 30396 results; incorporating the *seco*-steroid motif provides new structural scope featuring akamptisomers, with CAS returning only 89 analogous results for "*seco*-steroid" refined by "patent". *h*: original patent excludes all transoid possibilities; in August 2019, a search in CAS for staurosporin (ID 62996-74-1) refined by "patent" yielded 1048 records. *i*: also known as FR182877; *j*: Chapter 4.

Table 5.2 How literature, chemical software and databases depict stereoisomerism features associated with akamptisomerism.

#	structure a / b	→ auto-name *e.g.*, a	→ auto-structure a / b	cleaned ChemDraw structures	stereochemistry handling by CD/MS/CS software[a] a b	stereochemical specification and representation literature spec.[b] rep.	CAS spec. rep.	required stereodescriptors
1	Ar = 3,5-di-*tert*-butylphenyl simplified to:	13b,14a-difluoro-13b,14a-dihydro-14-oxa-4a¹,7a¹λ⁴,13a,14bλ⁴-tetraaza-13bλ⁴,14a λ⁴-dibora-1,13-(metheno)cycloocta[8,1,2-*cd*:4,5,6-*e′f*]di-*s*-indacene-2,3-diamine	Ar = 3,5-di-*tert*-butylphenyl simplified to:		F/F/F stereo. lost F/F/F stereo. lost	complete N, M complete	complete (accepted *parvo*/*amplo* descriptors) specifications inadequate representations	*parvo*/*amplo* and *R*/*S* for each of the 4-coordinate bridgehead atoms

#	structure → a / b	auto-name *e.g.*, a →	auto-structure a / b	cleaned ChemDraw structures	stereochemistry handling by CD/MS/CS software[a] a b	stereochemical specification and representation — literature spec.[b]	rep.	CAS spec.	rep.	required stereodescriptors
2		2,3,5,6,8,9,11,12-octaethyl-13b-fluoro-14a-phenyl-13b,14a-dihydro-14-oxa-4,4a¹ λ⁴,7,7a¹,10,13a λ⁴,14b,15-octaaza-13b λ⁴,14a λ⁴-dibora-1,13-(metheno)cycloocta[8,1,2-*cd*:4,5,6-*e'f'*]di-*s*-indacene			F/F/F stereo. lost	complete (as *transoid*") N, M	complete	none	inadequate representation	*parvo/amplo* for each of the 4-coordinate bridgehead atoms
3		2,3,5,6,8,9,11,12-octaethyl-13b-fluoro-14a-phenyl-13b,14a-dihydro-14-oxa-4,4a¹ λ⁴,7,7a¹,10,13a λ⁴,14b,15-octaaza-13b λ⁴,14a λ⁴-dibora-1,13-(metheno)cycloocta[8,1,2-*cd*:4,5,6-*e'f'*]di-*s*-indacene			F/F/F F/F/F stereo. lost stereo. lost	**3a** complete (as *cisoid*") N, M **3b** not seen	**3a** complete	**3a** none	**3a** inadequate representation	*parvo/amplo* for each of the 4-coordinate bridgehead atoms
4		5*H*,7*H*-1,11-ethanodibenzo[*b,g*][1,5]oxathiocine			F/F/F F/F/F stereo. lost stereo. lost	unknown but fluxionality shown by NMR	no stereochem. indicated	none	no stereochem. indicated	α/β for bridging O atom and the S atom

#	structure → auto-name → auto-structure a / b *e.g.*, a a / b		cleaned ChemDraw structures	stereochemistry handling by CD/MS/CS software[a]		stereochemical specification and representation				required stereodescriptors
						literature		CAS		
				a	b	spec.[b]	rep.	spec.	rep.	
5	9,18-dimethoxybenzo [*a*]naphtho[2,1-*m*] tetraphene			F/F/F stereo. lost	F/F/F stereo. lost	complete X	complete	none	no stereochem. indicated	*meso* for 5a, *M/P* for 5b
6	methyl (1*S*,2*R*,3*S*,5*R*)-3- (benzoyloxy)-8- oxabicyclo[3.2.1] octane-2-carboxylate			S/S/S	S/S/S	complete X, N	complete	complete	complete	*R/S* for each of the 4-coordinate bridgehead atoms
7	11-oxabicyclo[4.4.1] undeca-1,3,5,7,9- pentaene			S/S/S * fails when symmetry broken	S/S/S * fails when symmetry broken	complete X, N	complete	complete	complete	none for this highly symmetric system. α/β for bridging O atom if symmetry broken

#	structure a / b → auto-name *e.g.*, a → auto-structure a / b	cleaned ChemDraw structures	stereochemistry handling by CD/MS/CS software[a] a / b	stereochemical specification and representation — literature spec.[b] / rep.	CAS spec. / rep.	required stereodescriptors
8	15,16-dioxatricyclo [8.4.1.1^{3,8}] hexadeca-1,3,5,7,9,11,13-heptaene		F/F/F stereo. lost / F/F/F stereo. lost	8a complete X, N / 8a complete; 8b complete (inferred) / 8b complete	8a complete ("α,α or *syn*) / 8a inadequate representation; 8b complete (as *anti*) / 8b inadequate representation	α/β for bridging O atoms
9	(1*S*,2*R*)-1-ethyl-8-methoxy-2-methyl-1,2,3,4-tetrahydro-5,10-epoxybenzo[10]annulene-2-carboxylic acid		F/F/F stereo. lost / F/F/F stereo. lost	unspecified / not represented	none / no stereochem. indicated	α/β for bridging O atom
10	(1a*S*,7*R*,8a*S*,9a*R*,11*R*,11a*R*)-3,9,9,11-tetramethyl-6-methylene-6,7,8,8a,9,9a,10,11-octahydro-4*H*-7,11a-epoxycyclopenta[1,11]cyclopropa[5,6]cycloundeca[1,2-*b*]oxiren-4-one		F/F/F stereo. inverted / F/F/F stereo. lost	10a none / inadequate representation	10a none / inadequate representation	*parvo*/*amplo* and *R/S* for each of the 4-coordinate bridgehead atoms

159

#	structure a / b →	auto-name *e.g.*, a →	auto-structure a / b	cleaned ChemDraw structures	stereochemistry handling by CD/MS/CS software[a] a	b	stereochemical specification and representation — literature spec.[b]	rep.	CAS spec.	rep.	required stereodescriptors
11		(3a*S*,4*S*,5*R*,6*S*,9*R*,11a*S*)-4,6-dihydroxy-6-methyl-3,10-dimethylenedecahydro-5,9-epoxycyclodeca[*b*]furan-2(3*H*)-one			S/S/S	F/F/F stereo. lost	none (putative) NMR data does not rule out **11a ⇌ 11b** fluxionality	configurations indicated but inadequate representation	stereochem. different from primary reference	inadequate and inconsistent representation	*parvo/amplo* and *R/S* for each of the 4-coordinate bridgehead atoms
12		(3a*S*,4*R*,5*R*,6*S*,11a*S*)-4,6-dihydroxy-6-methyl-3,10-dimethylenedecahydro-5,9-epoxycyclodeca[*b*]furan-2(3*H*)-one			F/F/F stereo. lost	S/S/S	none (putative) NMR data does not rule out **12a ⇌ 12b** fluxionality	configurations indicated but inadequate representation	complete *R*/S**	inadequate and inconsistent representation	*parvo/amplo* and *R/S* for each of the 4-coordinate bridgehead atoms

#	structure a / b	auto-name *e.g.*, a	auto-structure a / b	cleaned ChemDraw structures	stereochemistry handling by CD/MS/CS software[a] a	stereochemistry handling by CD/MS/CS software[a] b	stereochemical specification and representation literature spec.[b]	stereochemical specification and representation literature rep.	stereochemical specification and representation CAS spec.	stereochemical specification and representation CAS rep.	required stereodescriptors
13		(3a*S*,4*R*,5*S*,11a*S*)-4-hydroxy-3,6,10-trimethylenedecahydro-5,9-epoxycyclodeca[*b*]furan-2(3*H*)-one			F/F/F stereo. lost	S/S/S	none NMR data does not rule out **13a ⇌ 13b** fluxionality	configurations indicated but inadequate representation	complete *R*/S**	inadequate representation	*parvo/amplo* and *R/S* for each of the 4-coordinate bridgehead atoms
14		(1*R*,2*S*,3*R*,6*R*,7*R*,8*S*)-8-isopropyl-1-methyl-5-methylene-11-oxabicyclo[4.4.1]undecane-2,3,7-triol			S/S/S	F/F/F stereo. lost	**14a** complete X **14b** not seen	**14a** complete	**14a** complete	**14a** complete	*parvo/amplo* and *R/S* for each of the 4-coordinate bridgehead atoms
15		(1*S*,8*S*,10a*S*,13a*R*)-1-nonyl-8-methoxy-10a,13a-dimethyl-2,3,3a,6,6a,7,8,9,10,10a,13,13a-dodecahydro-1*H*-4,11-epoxybenzo[*a*]cyclopenta[*f*][10]annulene			F/F/F stereo. lost	F/F/F stereo. lost	**15a** complete X **15b** not identified	**15a** complete	**15a** complete **15b** no record	**15a** complete	α/β for bridging O atom

161

#	structure a / b → auto-name *e.g.*, a → auto-structure a / b			cleaned ChemDraw structures	stereochemistry handling by CD/MS/CS software[a]		stereochemical specification and representation				required stereodescriptors
							literature		CAS		
					a	b	spec.[b]	rep.	spec.	rep.	
16		(1R,2R,3S)-11-oxabicyclo[4.4.1] undec-6-ene-2,3-diol			S/S/S	F/F/F stereo. inverted	none - no evidence for 2 species.	**16a** incomplete	**16a** complete *R*/S** **16b** no record	**16a** complete	*parvo/amplo* and *R/S* for the single 4-coordinate bridgehead atom
17		(2aR,3S,5R,7aS,8R,9S)-3,5-dihydroxy-1-isopropyl-5-(methoxymethyl)-2a,7a,8-trimethyl-3,5,6,7,7a,8,9,10-octahydro-3,9-epoxycyclobuta[a] cyclopenta[e][9] annulen-2(2aH)-one			S/S/S	F/F/F stereo. lost	**17a** complete N, M **17b** not identified	**17a** complete	**17a** complete **17b** no record	**17a** complete	*parvo/amplo* and *R/S* for the single 4-coordinate bridgehead atom

#	structure → auto-name → auto-structure a / b		cleaned ChemDraw structures	stereochemistry handling by CD/MS/CS software[a]		stereochemical specification and representation				required stereodescriptors
						literature		CAS		
				a	b	spec.[b]	rep.	spec.	rep.	
18		(5*S*,6*R*,7*R*,9*R*)-6-methoxy-5-methyl-7-(methylamino)-6,7,8,9,15,16-hexahydro-5*H*,14*H*-17-oxa-4b,9a,15-triaza-5,9-methanodibenzo[*b*,*h*]cyclonona[*jkl*]cyclopenta[*e*]-*as*-indacen-14-one		S/S/S	F/F/F stereo. lost	**18a** complete X **18b** not seen	**18a** complete	**18a** complete, **18b** no record	**18a** complete	*parvo/amplo* and *R/S* for each of the 4-coordinate bridgehead atoms
19		(1*S*,2*S*,3*R*,3a*R*,5a*R*,6a*R*,7*S*,10*S*,11*R*,13a*S*,13b*R*,13c*R*)-1,3,10-trihydroxy-2,5,7,11-tetramethyl-2,3,3a,5a,6,6a,7,8,9,10,11,13a,13b,13c-tetradecahydro-1*H*-7,12-epoxycyclonona[*a*]-*as*-indacene-13-carboxylic acid		S/S/S	F/F/F stereo. lost	**19a** complete (as lactone) X **19b** not seen	**19a** complete	**19a** complete 19b no record	**19a** complete	*parvo/amplo* and *R/S* for each of the 4-coordinate bridgehead atoms

a. CD: ChemDraw, MS: MarvinSketch and CS: ACD/ChemSketch. ChemDraw and MarvinSketch support [structure]→[name]→[structure] whilst ChemSketch only supports a [structure]→[InChI]→[structure] cycle. S = Successful, F = Fail. α/β stereodescriptors not implemented in CD, MS or CS and *parvo/amplo* stereodescriptors not yet accepted by IUPAC.

b. experimental structure determination; X = X-ray crystallography, N = NMR methods, M = molecular modelling.

5.3.2 Akamptisomerism possibilities for oxabicyclo compounds and their naming

In Chapter 4, four possible stereoisomeric arrangements were considered pertaining to akamptisomerism in the B_2OF_2pqx system. Here, all possible arrangements are considered including where the bridgehead atoms do not have substituents that are not part of the embedding. Tetrahedral bridgehead atoms may be describable as either (*R*) or (*S*) (or by some other absolute or relative configuration stereodescriptor), and their substituents can exhibit either a *transoid* or *cisoid* relationship to each other, *e.g.*, the *transoid* autakamptisomers **2** and their *cisoid* isomeric variants **3a** and **3b**. Whereas *transoid* akamptisomers may be indistinguishable, as in **2**, or distinguishable, as in the lower-symmetry pair **1a** and **1b**, *cisoid* akamptisomers are always distinguishable; the 4 general *cisoid/transoid* possibilities are sketched in Figure 5.4, along with 4 more possibilities depicting situations in which just one or none of the bridgehead atoms are tetrahedral. If molecules are of C_1 symmetry, then the relative orientation (relative, absolute, or pseudo-symmetric *R/S*)[4] of X–Y–Z to its surroundings is also significant, doubling the number of possible isomers. A wide scope of stereochemical possibility is thus intrinsically associated with akamptisomerism.

An alternative approach to the description of the stereochemistry of these compounds is as a composite rotamerism involving the X–Y and Y–Z bonds. Under the assumption of tetrahedral X and Z, there would be 6 general conformations (eclipsed, gauche and staggered) for each torsion. This gives 36 possible composite rotamers. Embedding of X–Y–Z is a significant constraint leading to only four possibilities. Considering these stereoisomers as akamptisomers recognises this critical structural feature.

As with all aspects of stereochemistry, stereodescriptors are needed to unambiguously name isomers; existing descriptors are listed in the IUPAC "Gold Book",[4] with these plus new ones proposed (Chapter 4, Section 4.4.4) that pertain to akamptisomerism. Figure 5.4 considers the basic principles involved and applies them to some representative compounds. If both bridgehead atoms X and Z are three-coordinate, then the established stereodescriptors α and β readily apply to Y. For the situation in which one or more are tetrahedral (or of higher coordination number), it would be possible to apply the Klyne-Prelog system, which focuses on the torsional angles about the X–Y and Y–Z bonds, defining *syn* and *anti* relationships. This system becomes undefined for certain structures typical of akamptisomerisation (Chapters 3 and 4) and hence is not robust and, additionally, the

resulting labels are not particularly human readable. To have an intuitive and robust nomenclature that is independent of coordination and structure, the new stereodescriptors *amplo* and *parvo* were introduced in Chapter 4 to depict the relationship between the bridgehead atoms X and Z with respect to their surroundings (Figure 5.4). This nomenclature captures the critical features and strengths of both the α/β and *syn*/*anti* nomenclatures and, in addition, making key 3D structural relationships immediately apparent from just the name.

(a) Number of tetrahedral bridgehead atoms:

0   **1**   **2 - *cisoid***   **2 - *transoid***

16a, 19a

3a, 6a, 6b, 10a,11a,14a, 17a, 18a

4a, **4b**, **7**, **8a**, **8b**, **9a**, **9b**, **15a**, **15b**   **16b, 19b**   **3b, 10b, 11b, 14b, 17b, 18b**   **1a, 1b, 2, 12a, 12b, 13a, 3b**

(b)

8a: (α,α) isomer
(15α,16α)-15,16-dioxatricyclo[8.4.1.1^{3,8}] hexadeca-1,3,5,7,9,11,13-heptaene
referred to in the literature and CAS as "*syn*"

8b: (α,β) isomer
(15α,16β)-15,16-dioxatricyclo[8.4.1.1^{3,8}] hexadeca-1,3,5,7,9,11,13-heptaene
referred to in the literature and CAS as "*anti*"

(c)

11a: (*amplo,amplo*) isomer
(3a*S*,4*R*,5*R$_{amplo}$*,6*S*,9*R$_{amplo}$*,11a*S*)-4,6-dihydroxy-6-methyl-3,10-dimethylenedecahydro-5,9-epoxycyclodeca[*b*]furan-2(3*H*)-one

11b: (*parvo,parvo*) isomer
(3a*S*,4*R*,5*R$_{parvo}$*,6*S*,9*R$_{parvo}$*,11a*S*)-4,6-dihydroxy-6-methyl-3,10-dimethylenedecahydro-5,9-epoxycyclodeca[*b*]furan-2(3*H*)-one

12a: (*parvo,amplo*) isomer
(3a*S*,4*R*,5*R$_{parvo}$*,6*S*,9*S$_{amplo}$*,11a*S*)-4,6-dihydroxy-6-methyl-3,10-dimethylenedecahydro-5,9-epoxycyclodeca[*b*]furan-2(3*H*)-one

12b: (*amplo,parvo*) isomer
(3a*S*,4*R*,5*R$_{amplo}$*,6*S*,9*S$_{parvo}$*,11a*S*)-4,6-dihydroxy-6-methyl-3,10-dimethylenedecahydro-5,9-epoxycyclodeca[*b*]furan-2(3*H*)-one

Figure 5.4   Stereochemical descriptors associated with akamptisomerism: (α,β) and *parvo* and *amplo* definitions. (a) Possible arrangements of substituents A and B at tetrahedral bridgehead atoms. Configurations are labelled with the new stereodescriptors (Chapter 4) *amplo* and *parvo* or else the existing descriptors α and β, with surrounding atoms represented as a ring for simplicity. Example isomers from Figure 5.2 and Table 5.1 are listed for each case, but sometimes require the surrounding ring atoms to be asymmetric so that the different forms are distinguishable; how these descriptors can be used in isomer naming is exampled for **8**, **11**, and **12**. (b) examples of (α,β) usage along with 3D models and (c) (*parvo/amplo*) usage. In (c) coordinates for the nine highlighted atoms in the 2D representation are depicted as in (a).

166

5.3.3 *Structural assignments of oxabicyclo akamptisomers*

With up to 8 general isomeric variants possible for molecules displaying akamptisomerism, knowing what stereoisomers constitute some synthesised or natural product may be a difficult task. Certainly, X-ray crystallography and 2D NMR techniques can, in principle, reveal all 3D information, but often such methods have not been, or cannot be, applied. Further, they only pertain to the conditions present for such specific measurements and different conditions could favour different arrangements and thermal distributions of stereoisomers. In Figure 5.2, 36 of the distinguishable isomers pertaining to 19 chemical systems are shown. All distinguishable akamptisomers are included, but for brevity, the selection shown is limited to only *cisoid/transoid etc.* variations that have previously been considered as structural possibilities.

Figure 5.2 indicates that, of **1** - **19**, the observed compound(s) have had the stereochemistry of their akamptisomeric centre unambiguously identified in only 13 cases. For the 17 systems with differentiable akamptisomers, both possible akamptisomers have been identified (or indeed postulated) in only 5 cases. For **9**, **10**, **13**, and **16**, the nature of the akamptisomer present has not been identified, with, in the case of **9** (at least), it being likely that the compound is an approximately equal mixture of both possibilities. The other poorly stereochemically defined compounds **11** and **12** pertain to badgerin and are of particular note. Initially[195] a natural product was identified by NMR as **11**, without distinguishing between **11a** and **11b**. However, recent analysis[196] of a sample from a different, but related, source plant species indicated either **12a** or **12b** instead. These possibilities differ in that **11** are *cisoid* whereas **12** are *transoid*. As the reported NMR data do not show a 1:1 correspondence, it is suggested that indeed both the *cisoid* and *transoid* stereoisomers may have been isolated; such a situation would parallel the recent identification of isomers **2** and **3a**. Most related natural products containing tetrahedral bridgehead atoms, and their synthetic analogues, are *cisoid*, making the identification of *transoid* forms like **2**, **12**, and the badgerin derivative **13,** unusual.

In the absence of X-ray crystallographic data, there is an explicit need to know all possible stereoisomeric configurations to fully inform chemical intuition and/or artificial intelligence approaches. Considering oxygen-bridged oxabicyclo compounds as akamptisomers provides the necessary conceptual basis for doing so.

5.3.4 Reliability of stereochemical information pertaining to oxygen-bridged oxabicyclo compounds stored in chemical databases

Complete description of stereochemistry in chemical databases requires a complete conceptual basis supporting the naming of compounds, understanding chemical reactions and transformations, and 2D diagrammatic representations that are stereochemically unambiguous. The IUPAC recommendations aim to provide this basis.[3-4, 89, 91, 170, 211] That IUPAC recommendations cannot account for all structures involved in stereoisomerism and stereoisomerisation for ML$_2$ stereogenic units has already been described (Chapter 4 Section 4.4), as too for the inability of existing stereodescriptors α and β combined with the Klyne-Prelog system to be intuitive and universally descriptive (Chapter 4).

Here, it is emphasised that the 2D drawn structures for **1** – **3** and **10** – **19** in Figure 5.2 do not follow IUPAC hash/wedge standard usage (Ref.[4] page 1926 rule ST.1.3.3) as these rules do not readily support the required stereochemical features. These recommendations are reproduced here.

From "Graphical representation of stereochemical configuration (IUPAC Recommendations 2006)",[91] page 1908:

ST-0.5 Stereobonds between stereocenters

Stereobonds between stereocenters should be avoided at all costs. This was stated in ST-0.3 but bears repeating because it is so important…

In rare cases—for example, when one stereocenter is completely surrounded by four other stereocenters—a stereobond must be present between two stereocenters. There is no ideal solution in such cases, as some ambiguity is unavoidable. In this case, solid wedged bonds should be used exclusively (as much as possible), and as always only the atom at the narrow end of the wedge should be considered as having a specified configuration. Hashed bonds connecting two stereogenic centers are always ambiguous, as discussed in ST-0.3.

Preferred   *Acceptable*
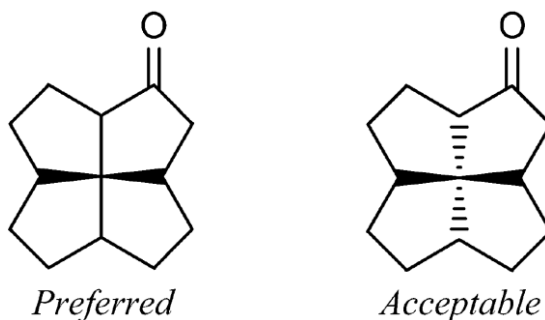
If the central atom has a specific configuration and each of the surrounding atoms also has a specific configuration, additional stereobonds must be added to the surrounding atoms. It is preferred to select a depiction style for the central atom that has as few stereobonds as possible, since that will eliminate completely any ambiguity regarding the intended configuration at some of the adjacent atoms.

Preferred   *Acceptable*
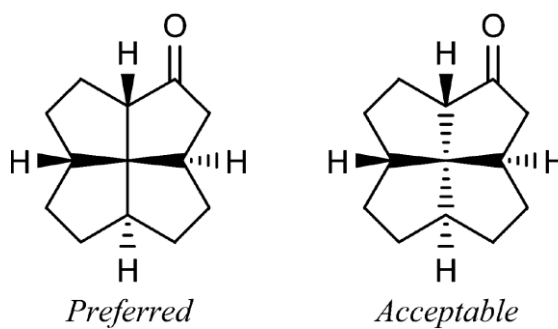
The introduced convention is to view bicyclo systems from one "face" and consider the entire X(A)–Y–Z(B) as a *single stereogenic unit*, to make obvious the critical information concerning 3D structure. In Table 5.3, 2D structures for **1** - **19** are presented adjacent to 3D models, making it clear that they are both unambiguous and intuitively interpretable; some details are not always embodied, however, *e.g.*, the B(F)–O–B(F) atoms in **1** – **3** are actually planar (and approximately orthogonal to the plane of the page). The reader is encouraged to look at the 2D representations and their comparison with 3D images, in some detail. The differences from current IUPAC practice are small and may appear familiar, yet these differences present profound consequences.

Indeed, failure of current IUPAC drawing recommendations is not an isolated issue, with other deviations from standard practice now being advocated in web-based undergraduate teaching environments.[212] The situation for akamptisomers is somewhat analogous to current issues[213] concerning Bredt's rule: by considering ring-size dependences, natural products labelled as "wrong" by IUPAC[4] are being increasingly identified. Indeed

many compounds discussed in that context can now be categorised as akamptisomers, with recent work highlighting the difficulties associated with stereochemical aspects of oxabicyclo[*m.n*.1] compounds.[214]

From "Graphical representation of stereochemical configuration (IUPAC Recommendations 2006)",[91] page 1926:

ST-1.3.3: Stereogenic centers at bridgehead atoms

Stereogenic centers at bridgehead atoms pose several difficult problems. When depicting substituents at bridgehead atoms, some chemists have been tempted to depict the substituent atom using the wedged bond style that is opposite to the one used for the bridging bonds, apparently on the assumption that "if the bridge goes 'up', then the other substituent must go 'down'." In fact, if the ring system is oriented in the plane of the paper and the bridging bonds are oriented upwards, then the substituents are oriented upwards as well in systems where the bridge contains fewer atoms than any other path between the bridgehead atoms (as in the usual depiction of bicyclo[2.2.1]heptane). Such compounds should be depicted using solid wedged bonds or hashed wedged bonds for the bridging bonds, with the substituents depicted using plain bonds. The substituents may also be depicted using the same bond type as the bridging bonds. Although it would be technically unambiguous to depict the bridging bonds using plain bonds and reserve the solid wedged bonds and hashed wedged bonds for the substituents, that style should be avoided due to the longstanding confusion over its intended meaning
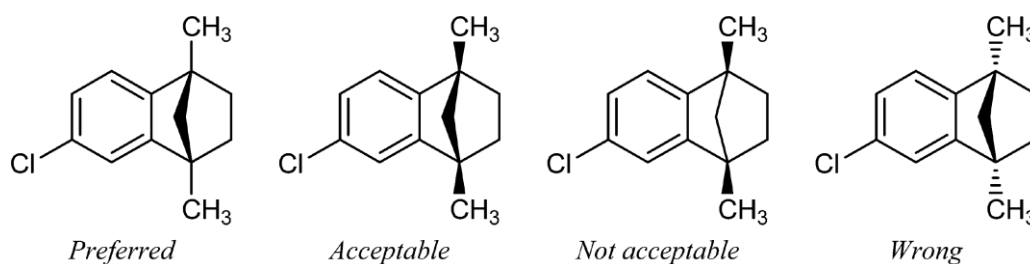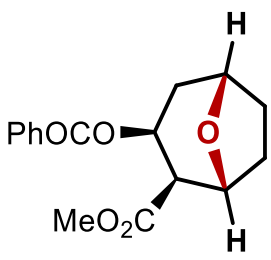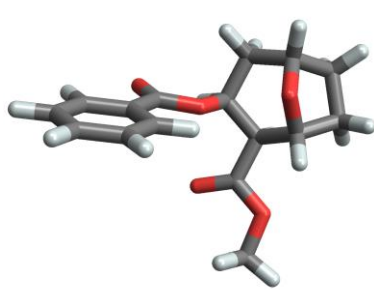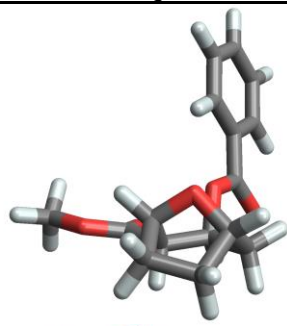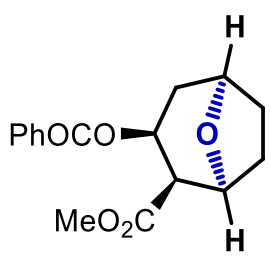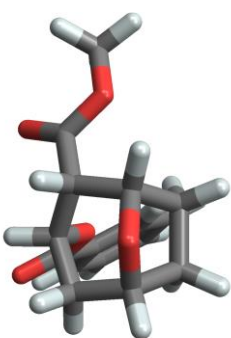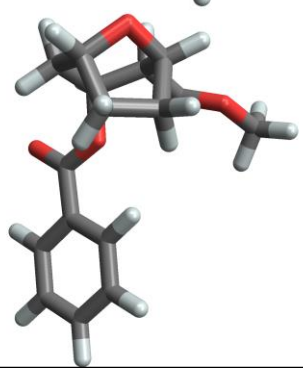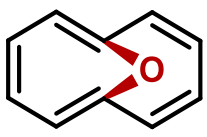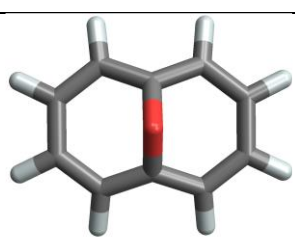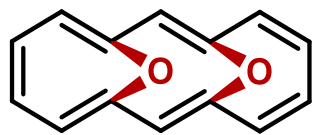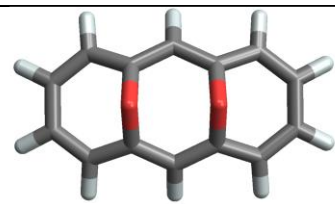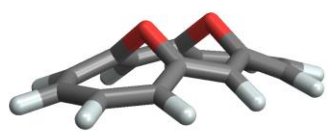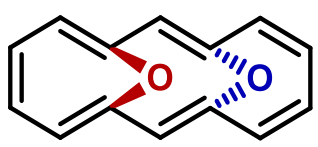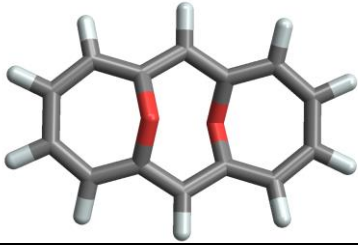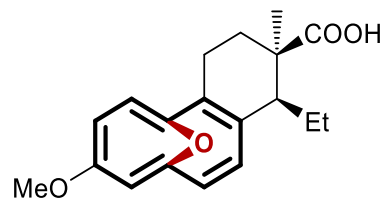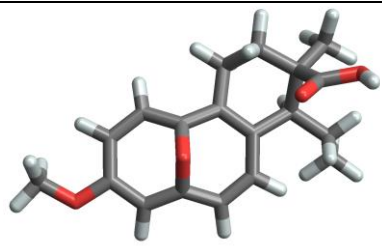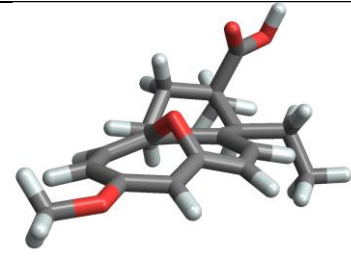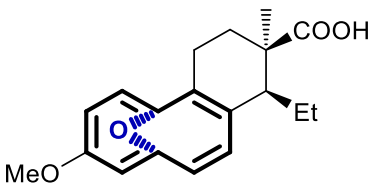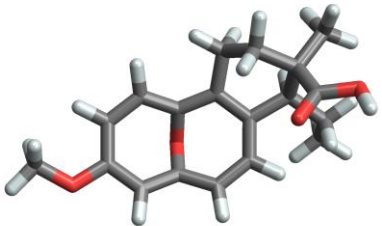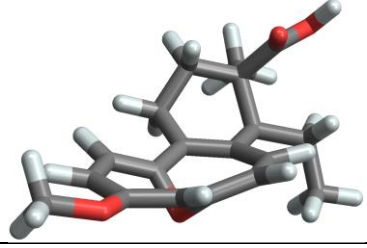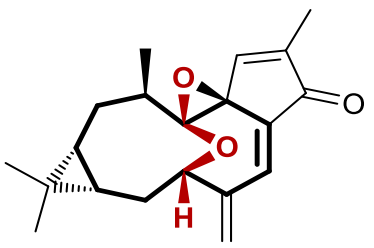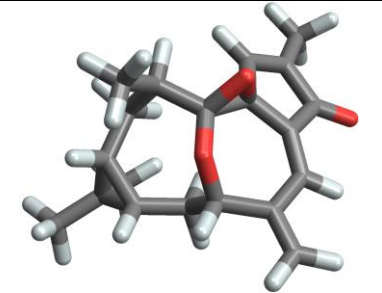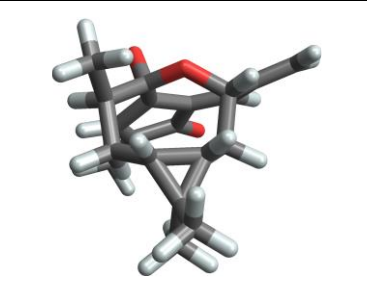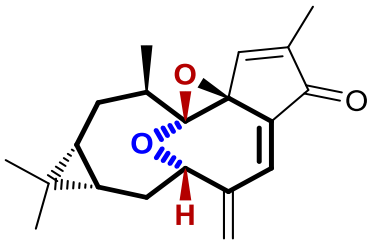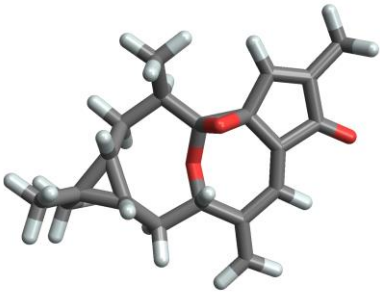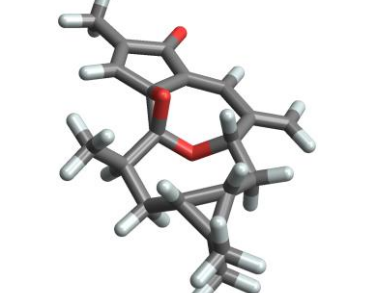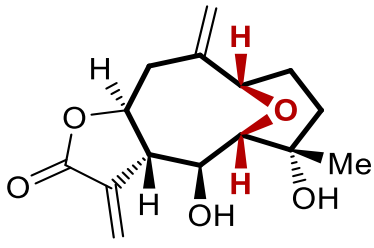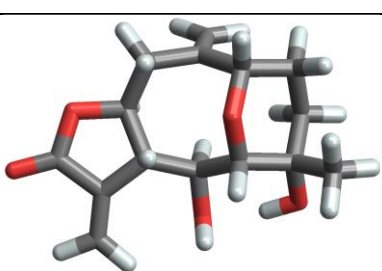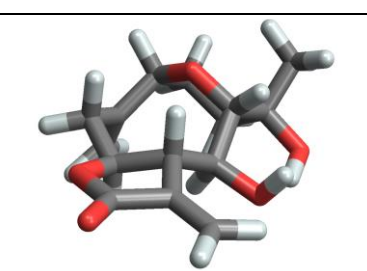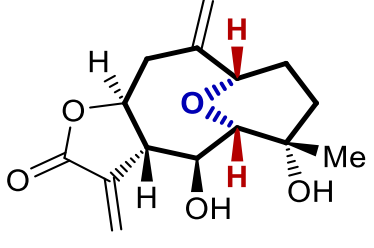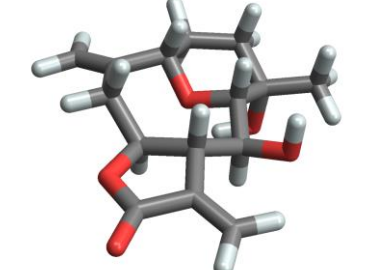
As mentioned, Figure 5.2 summarises the stereochemical experimental raw data of **1** – **19** pertaining to the 3D structure of the observed compounds. Following on, the 2D representations derived from the data presented in the original publications are analysed in Table 5.1 (with description of all relevant information presented in Table 5.3).

Table 5.3 Comparisons of 2D representations with 3D structures.
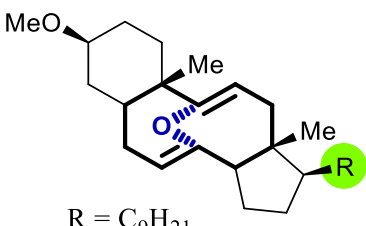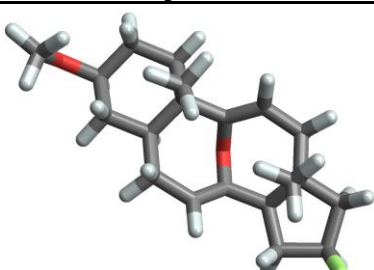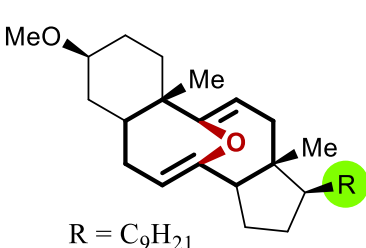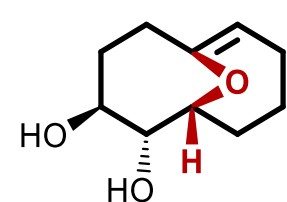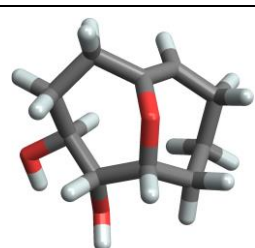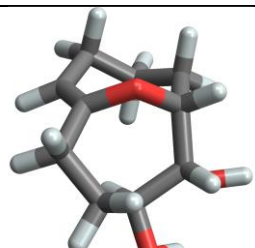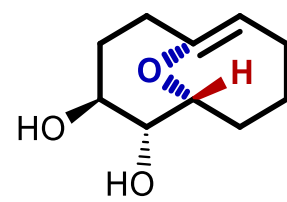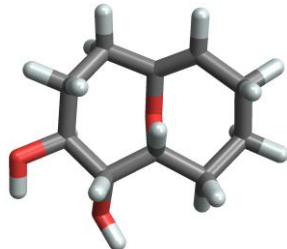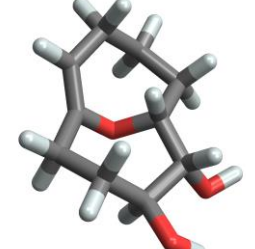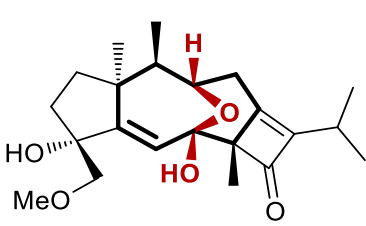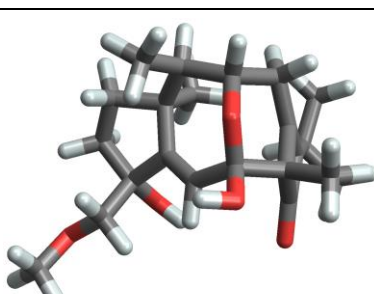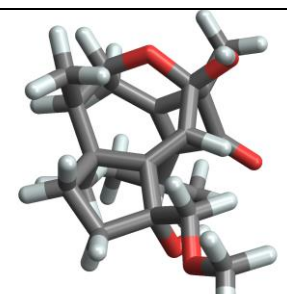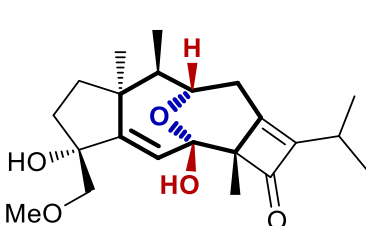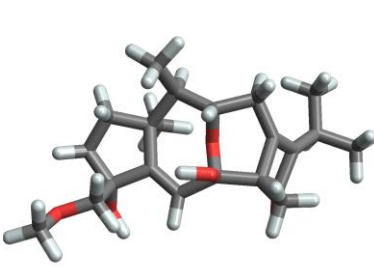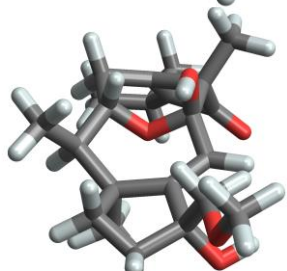
Mol.	2D representation	top view	oblique
1a	with Ar = H		
1b	with Ar = H		
2		Ethyl groups shown as H for clarity	
3a		Ethyl groups shown as H for clarity	

Mol.	2D representation	top view	oblique
3b		Ethyl groups shown as H for clarity	
4a			
4b			
5a			
5b			

Mol.	2D representation	top view	oblique
6a			
6b			
7			
8a			
8b			
9a			

Mol.	2D representation	top view	oblique
9b			
10a			
10b			
11a			
11b			

Mol.	2D representation	top view	oblique
12a			
12b			
13a			
13b			
14a			
14b			

Mol.	2D representation	top view	oblique
15a	R = C$_9$H$_{21}$		
15b	R = C$_9$H$_{21}$		
16a			
16b			
17a			
17b			

Mol.	2D representation	top view	oblique
18a			
18b			
19a			
19b			

In general, 2D representations could either accurately portray the available information (*i.e.*, precise stereochemical information or associated ambiguities), could fail to represent features, or could depict additional features not supported by the data. It is found that for **1** – **19**, the stereochemical information presented in the original literature accurately represented the raw experimental data. When the same structures were depicted in chemical databases such as CAS,[171] they were drawn *only* in accordance with IUPAC recommendations, but the chosen recommendation may not have been appropriate or the required recommendation may not have been available. In only 6 of the 19 cases (31%) is it found that the CAS structure accurately reflects the raw experimental information, otherwise either deleting stereochemical information or else adding unjustified features. Again, it is the small details that are easy to gloss over that can be the most important features.

To investigate how common it is that databases do not identify akamptisomerism-related stereochemical information, in Table 5.4, records are summarised for carbocyclic, metal-free, "ring-locked" oxabicyclo compounds found in CAS[171] that are listed in **E_Files_7** (see page xxv).

Table 5.4 Numbers of metal-free oxabicyclo[*m.n.*1] structures currently listed in CAS : either the total number[a] or those restricted to carbocyclic ring-locked compounds,[b] as well as the percentage[c] of them that have the isomeric features shown in Figure 5.4a fully characterised.

oxabicyclo class	total	Lock ring	
		Number	**% characterised**
[5.4.1]	2,379	11	0
[5.3.1]	5,633	52	0
[4.4.1]	3,740	84	15
[4.3.1]	6,055	318	0.3
[4.2.1]	10,385	277	83[e]
[3.3.1]	15,903	944	60[e]
[3.2.1][d]	101,484	43,765	
[2.2.1][d]	297,749	60,237	
[2.1.1][d]	479	29	69
total	443,807	105,717	

a: includes some over-counting for large rings as, *e.g.*, **8** which, as well as categorised as [4.4.1], can also be unhelpfully classified as [6.4.1]; sampled August 2019. *b*: ring-locking is very restrictive, hence producing human-manageable subsets, and prevents over-counting by excluding many compounds such as **8**, *etc*.; single-component structures without isotopes only are included, sampled March 2019. *c*: see **E_File_9** for a listing of all compounds analysed. Note that the same compound may appear multiple times in the data base, distorting the analysis. This may include instances in which it is depicted without stereochemical information and ones in which it is; such occurrences would result in underestimation of the fraction that are fully characterised. In contrast, the data may include instances where the 2D structure presented depicts full characterisation, yet the original data and published structure include undetermined features.
d: all-convex bicyclo. *e*: intermediate case, obtained assuming that only isomers typical of all-convex bicyclo structures are feasible.

No isomers with ring sizes of [5.4.1] or [5.3.1] are found to be differentiated according to the possible variations shown in Table 5.4, with 15% differentiated for [4.4.1], and 0.3% for [4.3.1]. In contrast, characterisation of the more numerous small-ring all-convex compounds is good, *e.g.*, 83% for [4.2.1], 60% for [3.3.1] and 69% for [2.1.1]. it is contended that this stark difference in characterisation stems from shortcomings in the current stereochemistry conceptual framework.

There exists a notable size dependence in stereoisomeric scope as shown in Figure 5.5. For smaller and more common bicyclo systems with systems with $m + n + 3 \leq 8$, the resulting structures are all convex with only *cisoid* substituents at the bridgehead atoms possible. Further, a BAR local operator applied to the X(A)–Y–Z(B) grouping results in Walden inversion of each of the bridgehead-atom configurations. For the larger systems with $m + n + 3 \geq 10$, the resulting structures are concave, at least, at the ring fusion points. In these cases, all *transoid* and *cisoid* configurations are possible.

For the intermediate oxabicyclo[3.3.1]nonane case where $m + n + 3 = 9$, the *transoid* stereoisomers **21**, and the (*parvo,parvo*) *cisoid* stereoisomer **22b** are each highly strained structures. These compounds were modelled in Gaussian16[153] at the B3LYP-D3(BJ)/6-311G(d,p) level. Shown in Figure 5.6, for **21** the *parvo* bridgehead atom features a bond angle of 143° and for **22b**, both *parvo* bridgehead atoms feature a bond angle at 153°. Relative to the (*amplo,amplo*) *cisoid* compound **22a**, compounds **21** are 150 kJ mol^{-1} higher in energy and **22b** is 379 kJ mol^{-1} higher in energy casting doubt over their intrinsic stability.

This size dependence on the scope of stereoisomerism at the X(A)–Y–Z(B) grouping is closely related to some instances of "in-out isomerism" (see Section A.20) which has been reviewed.[215] Indeed, this behaviour also parallels that in the discussion[213-214] of Bredt's rule, which fails to hold when applied to bicyclo[*m.n.l*] compounds with $S = m + n + l + 2 \gtrsim 8$, owing to the concave-convex geometry crossover of the bicyclo periphery. It is suggested that the prevalence of small-ring all-convex *cisoid* compounds has prejudiced the IUPAC recommendations as nomenclature and terminology development have been based on the requirements of commonly synthesised compounds.

saturated oxabicyclo[*m.n.*1]
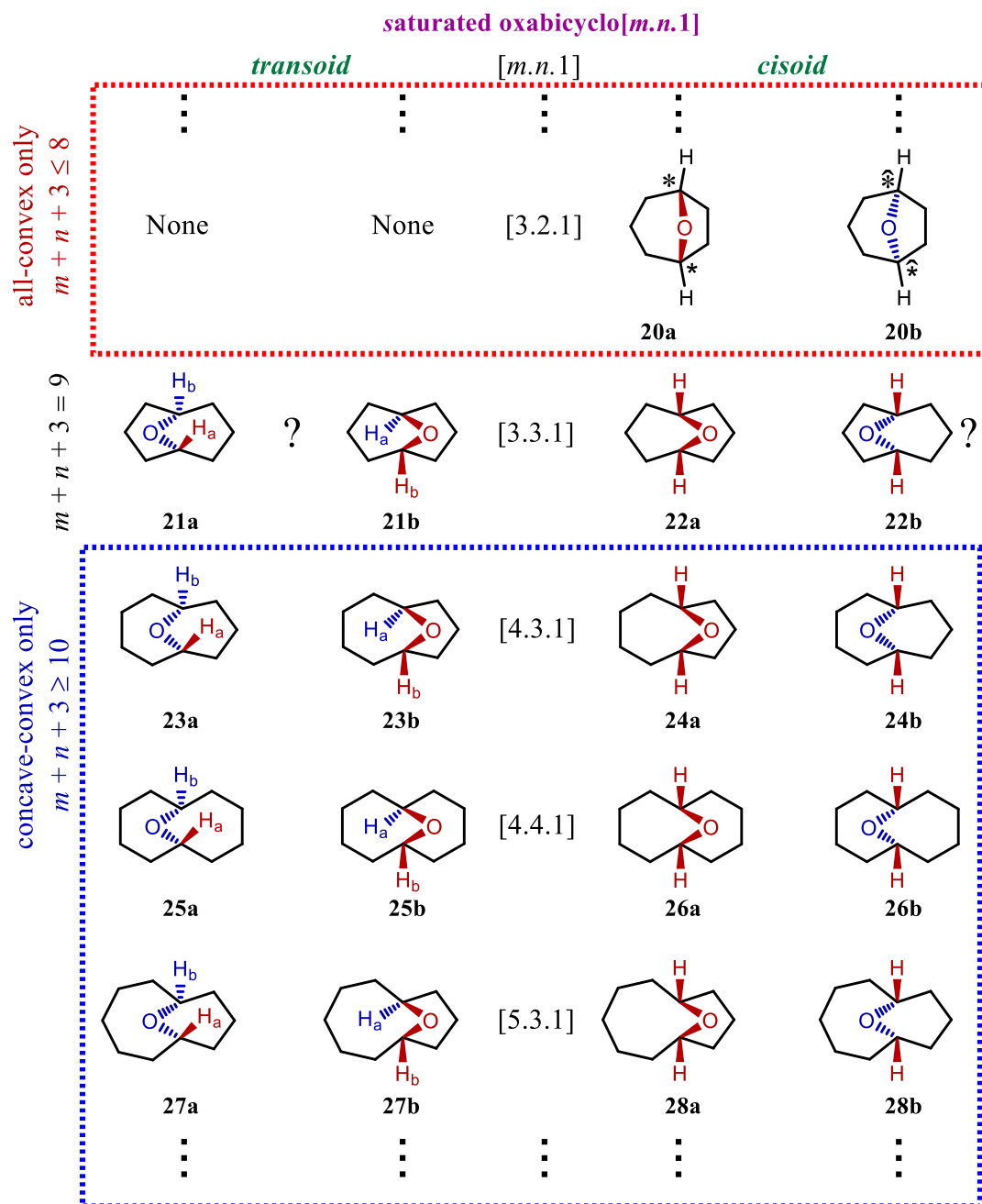
Figure 5.5   Convex – concave crossover in saturated oxygen-bridged oxabicyclo[*m.n.*1] systems. For the smaller and more common systems with $m + n + 3 \leq 8$, the resulting structures are all convex with only *cisoid* substituents at the bridgehead atoms possible. Further, a BAR local operator applied to the X(A)–Y–Z(B) grouping results in Walden inversion of the bridgehead configuration. For the larger systems with $m + n + 3 \geq 10$, the resulting structures are concave, at least, at the ring fusion. All *transoid* and *cisoid* configurations are possible. For the intermediate oxabicyclo[3.3.1]nonane case with $m + n + 3 = 9$, the *transoid* stereoisomers **21a** and **21b**, and the (*parvo,parvo*) *cisoid* stereoisomer **22b** are each highly strained structures casting doubt on their stability.

21b   **22b**
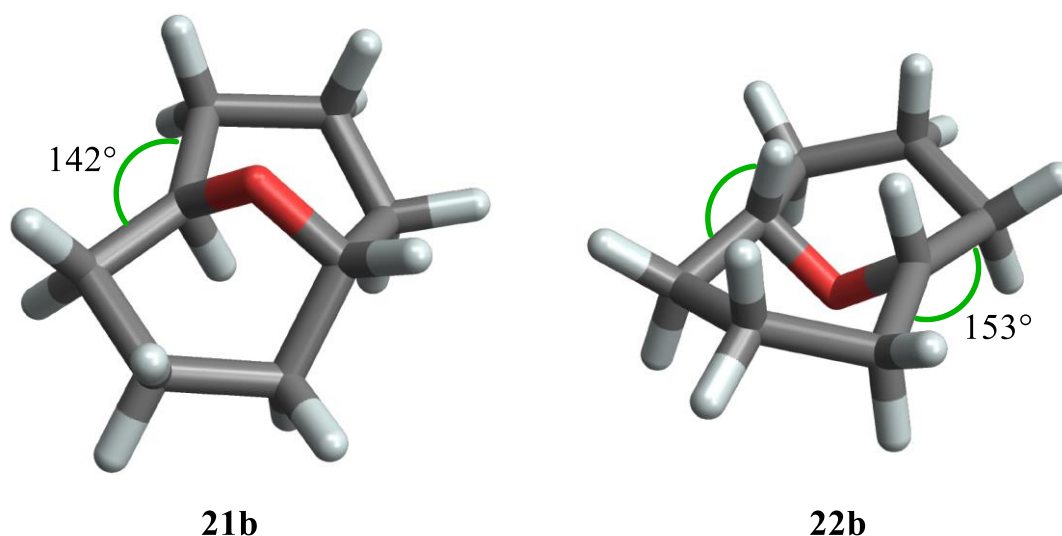
Figure 5.6 DFT structures of **21b** and **22b** with the highly strained bond angles at the *parvo* atoms as indicated. Structures calculated using Gaussia16[153] at the B3LYP-D3(BJ)/6-311G(d,p) level.

5.3.5 *Reliability of chemical infrastructure software for the processing of oxabicyclo compounds*

Chemical software, including drug-discovery platforms, process and manipulate 2D structures and chemical names that are assumed as being unique to each stereoisomer. Such software, in general, deals well with stereochemical subtleties (*e.g.*, axial isomerism in general, and atropisomerism in particular, as well as pyramidal "inversion"), being based on the encoding of IUPAC recommendations. In Table 5.1, as detailed in Table 5.2, the performance of the ChemDraw,[208] MarvinSketch,[209] and ChemSketch[210] software packages is considered in carrying out a cyclic operation, that of converting a 2D structure to a systematic IUPAC chemical name and then back to a 2D structure. Despite package differences, all returned 2D structures for all isomers equivalent to the starting ones for only 2 of the 19 example systems, with, as one would expect, no package explicitly recognising akamptisomers (see Table 5.2). Also, no chemical database currently explicitly recognises akamptisomers, nor chemical contextual identifiers such as SMILES[216] and IUPAC's InChI,[217] even though they can include stereochemical information. In addition, it is noted that the software packages considered offer "structure checking" and "structure cleaning" functions that do not recognise critical stereochemical aspects and hence can destroy valuable information under the currently deficient IUPAC recommendations. Indeed, and of particular note, none of the software package considered currently implement the α/β specifications

defined by IUPAC, which would have proved useful in many cases. The stereodescriptors, existing or new, needed to accurately depict each example molecule are listed in Table 5.2. Immediate consequences of this are that modern machine-learning software designed for target design and synthesis for the pharmaceutical and agrochemical industries[218-219] is often based on SMILES[216] specifications and therefore is currently not capable of developing explicitly *akamptisomerism*-selective targets or syntheses.

5.3.6 *Quantum chemistry and spectroscopic software*

In principle, IUPAC stereochemical recommendations should form the basis for modern quantum-chemical software that seeks to evaluate all properties concerning stereochemistry and stereoisomerisation by first-principles means. Instead, most software either avoids stereochemical descriptions altogether by using Cartesian coordinates, or else embed quantitative analyses based on the redundant-internal-coordinate stereochemical descriptors introduced by Wilson, Decius and Cross in 1955.[220] These implement all ML_1 stereodescriptors plus some pertinent to ML_2 but not those shown in Figure 5.4. With the rarely needed addition of improper torsions and extended linear atomic chains (features not supported by IUPAC), this approach provides a widely useful description of general stereochemical phenomena in terms of combinations of composite atropisomerism and akamptisomerism. Key to this is the internal determination of the best-possible composite description based on the complete set of available possibilities. All properties of ML_3, ML_4, and higher stereocentres are described this way, *e.g.*, Walden inversion in ML_4 is described as resulting from compositing two BAR-operator instances as demonstrated in Chapter 3, Figure 3.5.

As the Polytope Formalism indicates, all such descriptions of higher stereocentres embody singularities and are not generally applicable for the description of phenomena, just as two coupled torsions cannot universally describe all properties associated with akamptisomers. Wilson, Decius and Cross[220] showed how such anomalies can be internally detected, *e.g.*, by counting the number of non-zero eigenvalues of the nuclear kinetic-energy operator, and following detection widely-used software such as Gaussian16[153] stops with a "coordinate system error" message. Software like Gaussian16 typically checks for the singularity associated with the description of akamptisomerism as composite M–L torsions, attempting to mitigate its effects, and, *e.g.*, Reimers' DUSHIN package[221], that considers how nuclear motion controls UV/visible spectroscopic band shapes, does likewise. Nevertheless,

the intrinsic problem can only be solved by implementing the full set of ML_2 stereodescriptors and their geometrical bases (Figure 5.4, see also Chapter 3). Most user help-line issues associated with DUSHIN pertain to the effects of singularities in the Wilson-Decius-Cross approach. Only through the implementation of all fundamental isomerism forms and the exclusion of stereochemical representations that generate singularities can optimal, generally applicable, coordinate systems be obtained for use in quantum chemical software.

5.4   Discussion

The results obtained from looking at the 443,807 oxabicyclo compounds vindicates the hypothesis that the IUPAC recommendations for nomenclature and 2D depiction are incomplete. Some recommendations reflect a narrower scope of chemistry than is now available, *e.g.*, favouring the implicit depiction of isomers with *cisoid* bridgehead configurations at the expense of *transoid* ones despite analogous systems being known since 1968.[215] The current recommendations[91] were designed largely with small-ring examples in mind, *e.g.*, the cocaine derivatives **6**, and work for many natural products, *e.g.*, **10**, **14**, **17** and **18**, but fail for larger rings, *e.g.*, **1**, **2**, **12** and **13**. It remains to be determined if alternative *cisoid*/*transoid* stereoisomers can be synthesised, including many forms not explicitly represented in Figure 5.2. In addition, for many examples, either the akamptisomeric composition of the identified compound is uncertain (*e.g.*, **4**, **10** - **13**) or else alternative akamptisomers are possible that also have gone largely unforeseen or unrecognised (*e.g.*, **3b**, **14b, 15b, 17b, 18b, 19b**). Further, the scope of available oxabicyclo chemistry for large-ring compounds appears to be poorly explored, *e.g.*, of **1** - **19**, in only 5 instances (26%), **4 − 8**, could the available useful stereochemical space be considered to have been fully explored.

The results obtained can be summarised as:

1. the currently accepted basis for chemical intuition is incomplete and biased towards past achievements,

2. modern chemical databases fail to correctly store akamptisomerism-associated stereochemical information, and

3. chemical infrastructure software fails to uniformly process compounds correctly.

Here, ramifications of this are discussed, first by considering possible small-scale revisions to IUPAC recommendations. Additionally, significance for patents and government regulatory authorities, and future developments in the understanding of akamptisomerism are discussed. Finally, these deficiencies draw attention to a deeper issue with the historical approach to stereoisomerism with a call to overhaul the subject matter in favour of a fully systematic, and hence comprehensive, approach.

5.4.1 Small-scale revised IUPAC recommendations

In keeping with the tradition of adding new terminology and nomenclature for describing novel stereoisomerism, Chapter 2 lists several new terms for akamptisomerism and related needs (see Sections 2.4.1 – 2.4.6) and Chapter 4 describes the new supplementary stereodescriptors *parvo* and *amplo* (see Section 4.4.4).

5.4.2 New opportunities for patent protection

Issues of inadequately described stereochemistry can have important ramifications for patents, including validity as well as the opportunity to protect new discoveries and innovation. The material in this subsection was developed in collaboration with Dr Linda Govenlock, Partner, Head of Allens Patent & Trade Mark Attorneys, Patent Attorney. In patent law, the concept of "obviousness" or "inventiveness" is a fundamental concept, the assessment of which can vary across jurisdictions, and stereochemistry can be critical in this this.[222] As such, IUPAC definitions are important aspects of the underlying conceptual basis. Additionally, regulatory authorities such as the United States Foods and Drug Administration (US FDA) and the European Medicines Agency (EMA), encourage evaluation of single-stereoisomer active pharmaceutical ingredients (API), implicitly demanding full recognition of akamptisomers.[223-224]

In Table 5.1, 7 patent-relevant examples are considered. Within the available knowledge, the only existing patent in which akamptisomerisation is discussed[193] includes technological applications of the dipole switches **1** and **2**, but excludes the *cisoid* variant **3.** Existing patent definitions fail to describe the full range of stereoisomers associated with akamptisomerism for **3**, the estrogenic agent **9**,[189] the ovarian cancer active agent **10**[194], the internal-*seco*-steroid model **15**, the principle of which may be applied to the many steroid patents, and the kinase inhibitor staurosporine **17**.[204] For example, the original staurosporine patent[204] shows one general molecule that implicitly includes all stereoisomers, yet

subsequent specific embodiments failed to recognise the *cisoid/transoid* aspect associated with akamptisomerism.   Likewise, for **9**, **10**, and **17**, the stereoisomerism concept that is akamptisomerism was not explicitly discussed. Whilst these patents broadly cover all forms, this still leaves open the potential for new "follow-on" patents focused on a previously unrecognised stereoisomeric drug form that could prolong patent life – a strategy that has been used effectively in the pharmaceutical industry, *e.g.,* in relation to active enantiomeric forms.   Applied to a mature area such as steroid research, this could lead to a significant reinvigoration.

5.4.3   *The future of Akamptisomerism*

The discovery of akamptisomerism presents many immediate challenges, presenting unfettered synthetic goals, appropriate classification schemes, and enhancements to existing software, including:

1) *Additional classification* of all compounds meaningfully recognisable as akamptisomers.

2) *Improved structural rules* for 2D representations of akamptisomers. The entire akamptisomeric centre X(A)–Y–Z(B) needs to be considered as a single stereogenic unit.

3) *Revisions to conformational-analysis software,* including structure drawing, structure cleaning and naming functions, for both traditional chemical and drug-discovery applications. All existing software requires updating to recognise akamptisomers and their stereochemical implications.

4) *Data mining* of compound databases in search of incompletely characterised akamptisomers, including:

i)   mixtures of akamptisomers reported as single compounds,

ii)   compounds with incompletely described stereochemistry,

iii)   unrecognised akamptisomeric pairs, and

iv)   possible unexploited *cisoid/transoid* variations.   More sophisticated search algorithms need to be developed to facilitate this.

5) *Machine-learning high-throughput virtual-screening studies* of potential akamptisomeric ligands docked to proteins, DNA, sugars, *etc*., seeking ligands with improved binding.

6) *Patent specifications* need to recognise akamptisomers and related systems. Patented molecules may not have been sufficiently represented or specified in terms of stereochemistry. Consequentially the potential to pursue patent protection for stereoisomers not previously protected may have been overlooked, which presents an important opportunity for innovation, especially for new isomeric forms that exhibit new properties or advantages. In this respect, patent applications directed to new isomeric forms will need to satisfy legal requirements, including "obviousness" or "inventiveness", which can vary subtly across jurisdictions.

7) *Synthesis and testing* of targeted akamptisomer pairs and other newly conceivable isomers (*e.g.*, *transoid* structures) for possible medicinal and technological applications, *e.g.*, see Chapter 7

Such opportunities are rare for seemingly fully understood chemistry topics, and it remains to be seen how many new compounds can be conceived, synthesised, characterised, depicted, and catalogued. Enhanced IUPAC stereochemical recommendations can play a critical role in advancing these goals. In particular, Machine Learning provides a powerful diagnostic tool for processing large and complex datasets to identify trends, patterns, uses and effects, many of which were previously indiscernible. Concerning drug discovery, while such human endeavours can break free of the conceptual space provided by Nature, IUPAC recommendations, and current frameworks, all Machine Learning design methods are *strictly* limited to work within the internal conceptual space provided.[218-219] The ability to factor in stereochemical nuance, including akamptisomerism, as an additional parameter provides a powerful opportunity to further interrogate these datasets.

5.4.4 A broader perspective: full stereochemistry overhaul

To this point, the contention has been that the existing IUPAC system needs to recognise akamptisomers and associated phenomena to be able to describe a wider range of modern chemistry. Taking a broader perspective, it is worth recognising that, to date, the subject that is stereochemistry has adapted phenomenologically resulting in an *ad hoc* collection of concepts, terms, and nomenclature. The recent discovery, description and reporting[7] of akamptisomerism as a *novel* form of stereoisomerism highlights the fact that the current *ad hoc* approach, whilst practical, is not systematic. In short, it does not provide a

means to rigorously search for *all* stereochemical possibilities and can not answer the question: *what other forms of stereoisomerism remain to be discovered?*

Any solution seeking to answer that question requires a unified and mathematically rigorous, *a priori* model of all stereoisomerism – nothing short of a "Grand Unified Theory of stereoisomerism". Such an approach would naturally accommodate all possibilities – discovered or otherwise. The Polytope Formalism of stereoisomerism is just such a model, and it is proposed that it be implemented as such.

The Polytope Formalism of stereoisomerism maps out and categorises all structural possibilities for an ML_n stereogenic unit and naturally separates out geometric relationships from unimolecular stereoisomerisation processes. It does not bias the analysis towards structures that are "common" nor phenomena that is human-centric.

Superseding the current *ad hoc* IUPAC nomenclature with a systematic, mathematically based one will not only lead to dramatic simplification from the perspective of qualitative understanding and chemistry teaching but will also provide a rigorous basis for the development of chemistry software and the digital handling of chemical structure. Modern software has wide-ranging applications from molecular-structure drawing to molecule naming, to patent specifications and conformational searching, to Machine Learning and data mining, to materials and molecule design and drug discovery.

5.5   Conclusions

It may seem that the existing *ad hoc* IUPAC stereochemistry prescriptions, and their implementation into modern software, are adequate for most needs. Essentially, one does not miss what one does not know about. This is especially true for chemical products and reactions designed by computers – *digital chemistry*. The future is not about what is known but rather about discovering the unknown. Only mathematically rigorous stereochemical descriptors can direct the future, defining and enabling the exploration of chemical space.

Perhaps unsurprisingly, most of the compounds investigated were found to have been amply described within existing IUPAC recommendations and by modern software as these have been developed in order to account for commonly observed features. Focus here, however, is placed on the exceptions, which are shown to be numerous, represent many families of molecules of pharmaceutical and technological interests, and present significant

issues and new possibilities pertaining to patent protection and the fulfilment of the requirements of regulatory authorities. It is contended that only the full recognition of akamptisomers, and its classification within the larger context of stereochemistry, could allow the demonstrated systematic software weaknesses to be overcome. Taking a bigger perspective, the whole current *ad hoc* nomenclature for arbitrary ML_n stereogenic units is considered, calling for these current IUPAC stereochemical recommendations to be superseded with one based on the Polytope Formalism of stereoisomerism.

Conceptually, these two issues become critical when it comes to writing computer software designed to implement definitions pertaining to stereoisomerism. Software provides not only the means to store, draw, and name molecules, but also the conceptual basis in which molecules are conceived and designed by both human and artificial-intelligence (including machine-leaning) means.[218-219] To be most effective, software must seamlessly describe all possible geometrical scenarios that isomers can adopt using a universal description that *does not* embody singularities. Modern software has been generated based on the existing *ad hoc* IUPAC classification scheme in which it is implicitly assumed that akamptisomers can be treated as some form of composite rotamerism. These rules have grown adaptively based on the properties of discovered isomers, but, by not being systematic, there is no guarantee that they can encompass all possible ones or provide widely applicable, human-readable nomenclature.

Databases and software now provide essential tools for the exploration of new chemical space, facilitating both manual and Machine-Learning efforts. The experimental identification of akamptisomerism (Chapter 4) and the resulting deeper understanding of stereoisomerism relationships now demands the establishment of a complete set of rules and terms by IUPAC and its full implementation in databases, chemical software, and drug-discovery applications. Only a universal *a priori* conceptual basis for stereoisomerism, and the systematic approach that it guarantees, allows for the complete-in-principle search of chemical space. This is critical for the advancing of chemistry from a "creative art" further towards an "exact science".[225]

6 The Polytope Formalism of constitutional isomerism and isomerisation

6.1 Summary

A unifying and systematic framework for all isomerism is provided by application to constitutional isomerism of the underlying mathematics of the Polytope Formalism of stereoisomerism. All possible constitutional arrangements of a set of atoms are generated and the permitted pathways between any pair are also defined. This is a powerful tool to elucidate unimolecular isomerisation mechanisms. Cataloging constitutional isomerism is limited by superexponential scaling associated with bond permutations. Previously, isomeric structures were generated without consideration of unimolecular isomerisation intermediates. Our Polytope Formalism approach explicitly addresses this deficiency. Despite added complexity, problems of practical interest to chemists are made substantially easier by then focusing on a subset of atoms of interest that undergo bond permutation. This is achieved by partitioning the chemical system into active components $\mathcal{S}_m \mathcal{B}_n$, comprising m bonding sites $\mathcal{S}$ combined with n bonders $\mathcal{B}$, and non-active spectator atoms whose bonding topology remain unchanged. Spectator atoms limit motion of bonding sites $\mathcal{S}$, but not that of the bonders $\mathcal{B}$, and thus are ignored to increase mathematical tractability. Results are exemplified for the $\mathcal{S}_4 \mathcal{B}_2$ class by examination of tautomerism in free-base porphyrin. A graph corresponding to the full potential-energy surface (PES) is generated. This treatment for the $\mathcal{S}_4 \mathcal{B}_2$ class reveals that 28 distinct general structural forms called genera exist if all $\mathcal{S}$ are equivalent and form a topologically-cyclic set and all $\mathcal{B}$ are equivalent. Of these, 12 generic compounds correspond to already reported porphyrin derivatives. The structural forms increase to 225 when all $\mathcal{S}$ and $\mathcal{B}$ are distinguishable. The approach detailed here is broadly applicable to many chemical systems. Cataloging and graphing software tools are provided

6.2 Introduction

Knowledge of isomeric structures and chemical reactions is central to chemistry. Isomerism is composed of stereoisomerism and constitutional isomerism. To date, these two sub-categories of isomerism have been treated independently.

Recently, the utility of the polytopal-rearrangements model in providing a framework for the classification of stereoisomerism and associated stereoisomerisation processes was shown. In Chapter 3 the polytopal-rearrangements model was generalised to include all possible polytopal forms. Additionally, bond torsional isomerism was included, thus allowing the model to describe all types of stereoisomerism. The resulting model is comprehensive and mathematically rigorous, and we have named it the Polytope Formalism of stereoisomerism. It can be applied broadly to any bonding centre, organic or inorganic.

A key strength of the Polytope Formalism of stereoisomerism is that it does not only describe just the local minima structures. All possible *spatial configurations* for a system are treated on an equal footing with all this added complexity explicitly describing all the possible concerted unimolecular stereotropic rearrangements. When energies are assigned to each structure then the resulting graph is isomorphic to the *complete* potential-energy surface (PES) for that system. The Polytope Formalism not only describes all possible static structures but also all the allowed concerted unimolecular reaction mechanisms connecting the structures.

Recognising that constitutional isomerism can be viewed as the *permutation of bonds* for a collection of atoms, a direct analogy is drawn to stereoisomerism where the *spatial arrangements of atoms are permuted*. From this perspective, this allows for the direct adaptation of the mathematics embodied in the Polytope Formalism of stereoisomerism to be applied to constitutional isomerism. The result is a Polytope Formalism of constitutional isomerism. A consequence of implementing a common abstract framework to both stereoisomerism and constitutional isomerism is that these main branches of isomerism become unified.

The mathematics involved falls within the disciplines of Combinatorics and Graph Theory. Combinatorics concerns itself with the counting and categorisation of combinations and permutations or abstract objects. Graph Theory, as the name implies, is the study of *graphs* – representations of the relationships between pairs of abstract objects. Graph Theory's utility to the Polytope Formalism arises from the suite of theorems and other tools it provides for finding and describing properties of the graphs.

Explicit Combinatorics and Graph Theory approaches to enumerating (*i.e.,* envisaging, numbering and naming, to make a catalogue) constitutional isomers have been applied before, generally based on the Redfield-Pólya enumeration theorem or Burnside's

lemma, *etc*.[6, 226-233] Such approaches usually treat all atoms of the same element in a system at an equivalent level – typically following element valence.[6, 228, 234]. An unavoidable challenge for any such approach is that the problem scales superexponentially and so focus has been just on the number and description of the isomers, i.e., the number of potential-energy local minima structures that can accommodate zero-point vibration.

To apply the Polytope Formalism to constitutional isomerism, all analogous *bonding configurations* must be included in the enumeration. Consequently, the already difficult challenge of describing constitutional isomers is made that much more difficult with all this added complexity. Notwithstanding this, the resulting power that the Polytope Formalism of constitutional isomerism promises is worth the effort.

Whereas the Polytope Formalism of stereoisomerism is unavoidably descriptively complex, the Polytope Formalism of constitutional isomerism is, in its most general form, simple to describe and given in this Chapter. To address the superexponential scaling problem, as with earlier approaches for enumerating just the constitutional isomers, a simplification scheme is introduced to make the approach tractable. What differentiates the approach described in this work is that the simplification still retains all the information relating to the bond-rearrangement processes.

The approach is demonstrated in detail through the worked example of tautomerisation in free-base porphyrin which is particularly amenable to this method. Numerous adaptable software tools are provided to generate and manage the general implementation of the Polytope Formalism of constitutional isomerism. Additionally, necessary terminology directly analogous to that used in Polytope Formalism of stereoisomerism is introduced and used throughout.

6.3   The Polytope Formalism of constitutional isomerism

In broadest terms, constitutional isomerism is concerned with all the different arrangements of bonds between a set of atoms. An arrangement of bonds is called the connectivity (see Section A.8) or the *bond topology*. From a bond topology perspective, a "bond" can *only* exist between two atoms. In its most general form where there are no restrictions upon the bonding, the possible combinations of bond topologies, or *species*, for N atoms spans everything from the unique case when all N atoms are unbonded through to the

unique case when all N atoms are bonded to each other. As an example, Figure 6.1 shows the first two and the last two bonding topologies for an ordered listing for the 5-atom system.
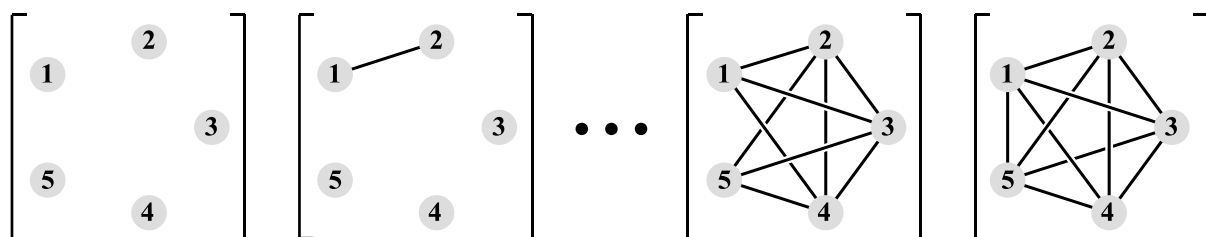
Figure 6.1 Unrestricted bonding topologies for the 5-atom system. The first two and last two bonding topologies for an ordered listing are shown as molecular graphs using a "circular layout". There are a total of 1024 distinct bond topologies for the 5-atom system.

Enumeration of all unrestricted bond topologies for N atoms is the sum of all possible arrangements of k bonds for N atoms, where k runs from 0 through to the maximum number of bonds.[233] For N atoms, the maximum number of bonds is equal to the binomial expression $\binom{N}{2} = N(N-1)/2$. Thus, the summation can be expressed by Equation (1) which simplifies to Equation (2):

$$\sum_{k=0}^{\binom{N}{2}} \binom{\binom{N}{2}}{k} = \sum_{k=0}^{N(N-1)/2} \binom{N(N-1)/2}{k} \tag{1}$$

$$= \sum_{k=0}^{N(N-1)/2} \frac{(N(N-1)/2)!}{k!\,([N(N-1)/2]-k)!}$$

$$= 2^{N(N-1)/2} \tag{2}$$

Table 6.1 shows the evaluation of Equation (2) for $N = 1$ through to 10, showing that, even for systems of modest size, the enumeration rises superexponentially. Indeed, the scaling is formally of the order $\mathcal{O}(2^{N^2})$, making it one of the most computationally difficult problems to solve.

Table 6.1 Total number of unrestricted bond-topology combinations for a given set of N atoms

N	$2^{n(n-1)/2}$	
1	1	$= 2^0$
2	2	$= 2^1$
3	8	$= 2^3$
4	64	$= 2^6$
5	1 024	$= 2^{10}$
6	32 768	$= 2^{15}$
7	2 097 152	$= 2^{21}$
8	268 435 456	$= 2^{28}$
9	68 719 476 736	$= 2^{36}$
10	35 184 372 088 832	$= 2^{45}$

As Equation (2) describes *all* mathematically possible bonding topologies, it places an upper bound on any enumeration as any imposed constraints will reduce the total number. Fortunably, for real chemical systems, valence and steric considerations place obvious limitations on the maximum number of bonds any particular atom in the set may form with this greatly reducing the number of bond-topology combinations worth considering. Typically, *valence* imposes the greater of these restrictions with previous attempts at enumeration of isomers based around this principle. [6]

An additional obvious constraint concerns minimal connectivity. All traditional isomer-enumeration approaches impose the requirement that the bonding topologies (species) must constitute a single connected entity thus complying with the IUPAC definitions for constitution (see Section A.9) and constitutional isomerism (see Section A.10).

Though these constraints are obvious ones and chemically sensible they restrict the structure enumeration only to potential energy local minima structures (isomers) and provide little or no information about processes that interconvert between these. A key strength of the Polytope Formalism, be it of stereoisomerism or constitutional isomerism, is that a far broader range of possibilities is considered (in principle, *all* possibilities) with this providing information about interconversion processes. In the Polytope Formalism of constitutional isomerism, sensible subvalent and hypervalent bonding topologies are explicitly included to account for bond topologies representing reaction intermediates.

The graph that represents all the relationships between the set of structures has many uses. The most obvious use is that each graph edge, interconnecting a pair of graph vertices, defines an allowed concerted unimolecular topotropic (see Section 2.4.28) rearrangement process. This type of rearrangement is given the symbol $R_{to}^c 1$, introduced here for the first time to the literature (see Section 2.4.21). An $R_{to}^c 1$ process is the bond topology equivalent of the concerted unimolecular *stereotropic* rearrangement $R_{st}^c 1$, as used for the Polytope Formalism of stereoisomerism (see Section 2.4.20 and Chapter 3).

As with stereoisomerism, for constitutional isomerism the graph of bond-topology species can have potential energies, gradients, and curvature assigned to all graph vertices to give a compact representation of the bong topology full potential energy surface for the system.

It is worth noting that in the context of stereoisomerism, the word "polytope" that forms part of the name of the formalism does refer to geometric polytope objects. In the context of constitutional isomerism, the word "polytope" now refers to a more abstract analogue, that being *topological polytopes* which represent specific bonding topologies.

6.4 The "$S_m B_n$ partitioning" approach to the Polytope Formalism of constitutional isomerism

As shown above, enumeration of bonding topologies for even relatively small systems rises superexponentially. To make practical use of the strengths of the Polytope Formalism of constitutional isomerism significant simplifications must be introduced. Whilst many approaches are conceivable, the one described here is inspired by the fact that many problems of practical interest to chemists involve molecular systems where only a small subset of atoms in a significantly larger molecule undergo bond topology changes. Typically, these "active" atoms can be further divided into two categories based upon the scope of bond topology changes. "Topologically mobile" fragments called "bonders" B, are free to move about and take up differing bonding arrangements with a similarly small number of bonding sites S. The bonding sites are less mobile and are otherwise bound to the remaining part of the molecular system, this being "bond-topologically static" and are called "spectator" atoms. If there are m bonding sites S and n bonders B, this partitioning approach is termed $S_m B_n$. This is depicted graphically in Figure 6.2a.

In Figure 6.2b, free-base porphyrin is partitioned with the inner two hydrogen atoms as the bonders and the four nitrogen atoms as the bonding sites with all other atoms as spectators. This $S_4 B_2$ partitioning allows the Polytope Formalism of constitutional isomerism to describe *all* possible tautomerism-related structures for free-base porphyrin. Similarly, in Figure 6.2c, the B(F)OB(F)-porphyrin is $S_4 B_2$ partitioned as in Figure 6.2b but with the boron atoms as the bonders. Whilst in free-base porphyrin the bonders are not topologically connected to each other, in the B(F)OB(F)-porphyrin example the bonders are topologically connected, though in this case this bonder – bonder connection only minimally limits motion of the bonders. Nevertheless, with the same partitioning, both systems are equivalent within the Polytope Formalism of constitutional isomerism.
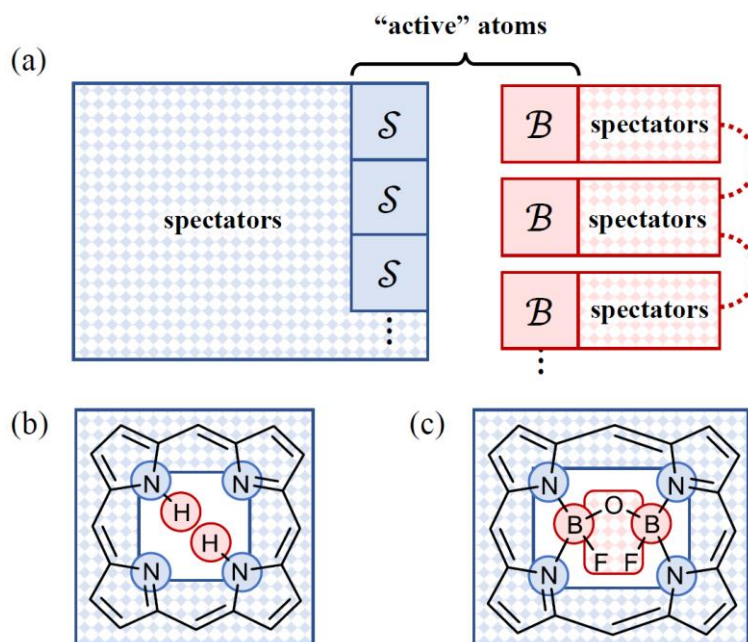
Figure 6.2 $\mathcal{S}_m\mathcal{B}_n$ partitioning of a molecular system. (a) The molecular system of interest is partitioned into collections of atoms or groups of atoms representing (1) bonding sites $\mathcal{S}$, for which the spectator atoms limit motion, and (2) bonders $\mathcal{B}$, for which the spectator atoms do not limit motion.[7] The bonding sites and bonders constituting the active atoms define the isomers or isomerisation of interest. (b)-(c) Two examples pertaining to free-base porphyrin: (b), see Figure 6.3, with unconstrained bonders, and (c) this after B(F)OB(F) substitution to establish a bonder to bonder through-spectator connection that, in this case, only minimally affects isomerism possibilities.

Symmetry is an additional property that can be used to further simplify the $\mathcal{S}_m\mathcal{B}_n$ partitioning approach. As the bonding sites and the network of spectator atoms to which they are bound remains a single unit (coloured blue in Figure 6.2) throughout the bond permuting process, the symmetry of this fragment can be used to group together species that are bond-topologically degenerate. Each set of bond-topologically degenerate species comprises a *genus* (plural: *genera*, adjectival form: *generic*, see Section 2.4.15). If there is no bond-topology degeneracy, the genus will only contain a single species (a "monotypic genus").

In Figure 6.3a is shown the two lowest energy forms involved in the tautomerism (see Section A.46) of free-base porphyrin. When R = H, the two *trans* species are degenerate, and the tautomerisation (see Section A.47) is a degenerate rearrangement (see Section A.12). When R ≠ H, the two *trans* species are distinct and not necessarily isoenergetic.

The tautomerism and tautomerisation processes can be described by partitioning free-base porphyrin (R = H) as $\mathcal{S}_4\mathcal{B}_2$ as shown in Figure 6.2b. The "porphyrindiyl" fragment, *i.e.*, the molecule less the two inner hydrogen atoms, exhibits the D_{4h} symmetry point group from a bond topology perspective. In Figure 6.3b are shown three bond-topology species **3a**, **6a**, and **2a** of free-base porphyrin (all free-base porphyrin bond topology structures will be

described in detail later in Section 6.6). Species **3a** is a *trans* isomer (potential energy local minimum structure) and **2a** a *cis* isomer (potential energy local minimum structure). Species **6a** is the first-order transition-state structure along the $R_{to}^{c}1$ reaction path connecting **3a** and **2a**.

(a)

(b)

Species:	**3a** *trans*	**6a** *cis ⇌ trans*	**2a** *cis*
Compact label:	LM21:23	TS21:μ22,23	LM21:22
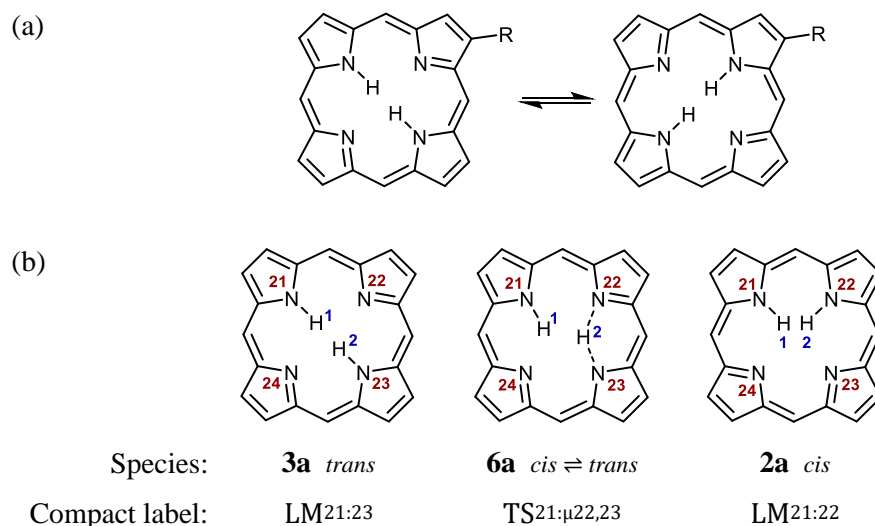
Figure 6.3 Free-base porphyrin tautomerism species. (a) Tautomerism in free-base porphyrins. Only lowest energy *trans* forms (shown here) are seen experimentally. When R = H, the structures are degenerate. When R ≠ H, the structures are distinct and generally non isoenergetic. (b) Shown is the *trans* species **3a**, *cis* species **2a**, and their associated interconversion first-order transition-state structure species **6a**. The labels 21 − 24 depict the standard atom locant numbering in porphyrins.[150] The compact labels describe the character of the structure as a potential energy local minimum (LM) or a first-order transition-state structure (TS) along with the colon-separated nitrogen atom locants that bond to H(1) and H(2), respectively.

Depending upon the specific implementation, all bonders may be constrained to form at least one bond with a bonding site to preserve the property of each bond topology structure comprising a single fully connected molecular unit. If this constraint is lifted this allows the bonders to dissociate and hence this approach can also describe dissociation mechanisms.

The structures in Figure 6.3 are *species* as the *specifically* labelled H atoms are bonded to the *specific* nitrogen atoms as indicated. Species **3a** is an example of a *trans* isomer and, due the D_{4h} symmetry of the porphyrindiyl fragment, there are four different ways of permuting the bonding of the two hydrogen atoms to the four nitrogen atoms whilst preserving a *trans* relationship. All four *trans* species, without the atom labelling, are structurally degenerate and comprise the *trans* genus **3**. Similarly, species **2a** is one of eight *cis* species that comprise the *cis* genus **2**, and species **6a** is one of 16 *cis* ⇌ *trans* transition-state structure species that comprise the genus **6**. In this work, genera are indicated simply by a number whilst species are given a number and appended letter. Further, no atom locant

numbers should appear when depicting a generic structure. Figure 6.4 shows the corresponding genera for the species in Figure 6.3.

Genus: **3** *trans*   **6** *cis ⇌ trans*   **2** *cis*
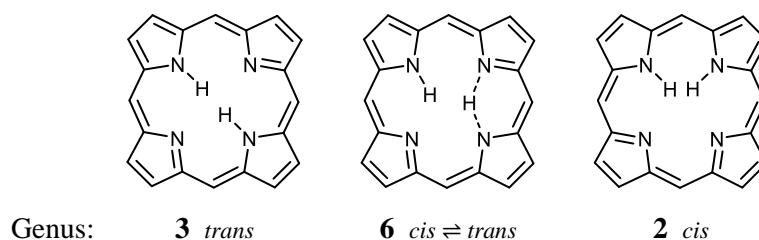
Figure 6.4 Examples of free-base porphyrin tautomerism genera corresponding to species from Figure 6.3.

Genera can further be grouped into the higher taxonomic rank of *family* based upon the bonding topology character of the bonders. For example, in free-base porphyrin, both the *trans* **3**, and *cis* **2**, genera and all their constituent species, feature hydrogen atom bonders that only form a single bond to nitrogen atom bonding sites. Thus, the genera **3** and **2** belong to the same family. Genus **6** features one hydrogen bonder that forms a single bond and one hydrogen bonder that forms two bonds and so belongs to a different family. Later in the Chapter, a full tabulation of species, genera, and families for the D_{4h} symmetric, $\mathcal{S}_4\mathcal{B}_2$ partitioning for tautomerism in free-base porphyrin is presented.

As a final taxonomic rank, all families for a set of atoms form a *class*. In general and for the context of constitutional isomerism for N atoms, this is referred to as the N atom class. For the $\mathcal{S}_m\mathcal{B}_n$ partitioning approach described here, "$\mathcal{S}_m\mathcal{B}_n$" is the class.

6.5   Enumeration and cataloguing $\mathcal{S}_m\mathcal{B}_n$ partitioned bonding-permuted chemical species

For a $\mathcal{S}_m\mathcal{B}_n$ partitioned chemical system where all bonders must form at least one bond to a bonding site, the total number of species N_s, and number of families N_f, is given by

$$N_s = (2^m - 1)^n \tag{3a}$$

and

$$N_f = \binom{m + n - 1}{n} = \frac{(m+n-1)!}{m!\,n!} \tag{4a}$$

More generally with bonder dissociation permitted, the total number of species N_s, and number of families N_f, is given by

$$N_s = 2^{mn} \tag{3b}$$

and

$$N_f = \binom{m+n}{n} = \frac{(m+n)!}{m!\,n!} \tag{4b}$$

As described in 6.4, structures may be chemically degenerate due to symmetry of the bonding sites, allowing for numerous species to be collected into a single genus. Following, further description of the $\mathcal{S}_m\mathcal{B}_n$ partitioned chemical system is focussed on the example of tautomerism in free-base porphyrin where the symmetry is D_{4h} and the partitioning is $\mathcal{S}_4\mathcal{B}_2$.

To provide a catalogue of bond topology species, software for generating families, genera, and species for $\mathcal{S}_m\mathcal{B}_n$ partitioned chemical systems, under the assumption of non-dissociation, is provided in **E_File_11** with a brief description and instructions given in Appendix G. Also in Appendix G are tables showing the enumeration of families, genera (given D_{nh} site symmetry), and species for m and $n \leq 8$. Note that this software is analogous to that which could be used to find a numerical solution to a so-called "intractable" NP-complete problem.[233]

6.5.1 *Compact notation labels for $\mathcal{S}_4\mathcal{B}_2$ partitioned bond-permuted free-base porphyrin tautomerism species*

To help concisely and informatively label a species, the following compact label notation is used throughout. The label comprises a two-letter descriptor of the PES critical point type for the species with LM being a local minimum, TS, a first-order transition-state structure, kS a kth order transition-state structure, and NC a noncritical point. Additionally, kI can represent a kth order inflexional critical point though these are not found for free-base porphyrin. Appended to right of this is a colon-separated listing of the bonding site atom locant numbers that form bonds to the bonders and is ordered by bonder number. If a bonder is bonded to more than one bonding site, and adaptation of the IUPAC "μ" notation[3] is used. For example, the *cis* $\rightleftharpoons$ *trans* transition-state species **6a** has the second H atom bonded to nitrogen atoms 22 and 23 – this bond topology being written as μ22,23. If a bonder is bonded to more than two bonding sites the μ notation follows IUPAC nomenclatural form with a

right subscripted number indicating the bridging number. For example, a hydrogen atom bonder triply bridging nitrogen atoms 21, 22, and 23 would use μ₃21,22,23.

For free-base porphyrin, the hydrogen atom bonders are seen to be physically situated in the quasi-plane of the macrocycle or above or below the quasi-plane. This property is utilised in the construction of the label with in-plane H atom bonders having their bonding locant lists positioned to the right and in the middle of the line. For example, the *trans* local minimum species **3a**, has both H atoms in the plane of the porphyrin with H(1) bonded only to N(21) and H(2) only to N(23), thus the symbol is LM21:23. If an H atom is bonded above or below the quasi-plane of the porphyrin, its bonding locant list is right superscripted or subscripted, respectively. For example, a local minimum species with both H atoms bonded to N(21) is given the label LM_{21}^{21}.

6.5.2 *Allowed concerted unimolecular reactions that interconvert species.*

As noted earlier, the critical feature of the Polytope Formalism of stereoisomerism where all concerted unimolecular reactions that interconvert species are identified is maintained within this adaptation to constitutional isomerism. The bond-topology species are represented as graph vertices on a graph, with the allowed $R_{to}^c 1$ processes represented as graph edges connecting the graph vertices. Software is provided in **E_File_12** for constructing such graphs from the species lists generated by the software in **E_File_11**. Detailed usage instructions and sample outputs are provided in Appendix H. This graphing software calculates various graph properties and generates both labelled pictorial representations of the graph objects and graph-object specific output file types of general utility to the user.

An example of the pictorial graphing software is shown in Figure 6.5a. This shows the first-order motions (only one bonder at a time is allowed to make a single bond-topology change – this will be discussed further in Section 6.6.2) for all species belonging to genera **2**, **3**, and **6** for the D_{4h} site symmetric $S_4 B_2$ family. A degenerate *trans* ⇌ *trans′* rearrangement of **3a** and **3b** representative of experimentally observed tautomerisation of free-base porphyrin is shown in Figure 6.5b. In Figure 6.5c a shortest allowed elementary pathway interconnecting **3a** and **3b** involving three intermediate species belonging to genera **6** and **2** is shown. This pathway has been highlighted (red and blue) and annotated on the graph shown in Figure 6.5a.

Also, in **E_File_13** is "S4B2_genera{1-28}_1_2.pdf"; a large, high-resolution pictorial representation of the full graph for the $\mathcal{S}_4\mathcal{B}_2$ partitioning implemented for free-base porphyrin.

(a)

```
Class: S4B2 with site symmetry point group D4h
Genera included = {2, 3, 6}
1st order motions graph:
Graph vertex layout: TutteEmbedding
Hamiltonian graph? False
Eulerian graph? True
Planar graph? True
Graph radius = 8
Graph diameter = 8
Graph density = 0.0846561
```


(b)

3a: LM21:23   **3b**: LM24:22

(c)

3a: LM21:23   **6a**: TS21:μ22,23   **2a**: LM21:22   **6b**: TSμ21,24:23   **3b**: LM24:22
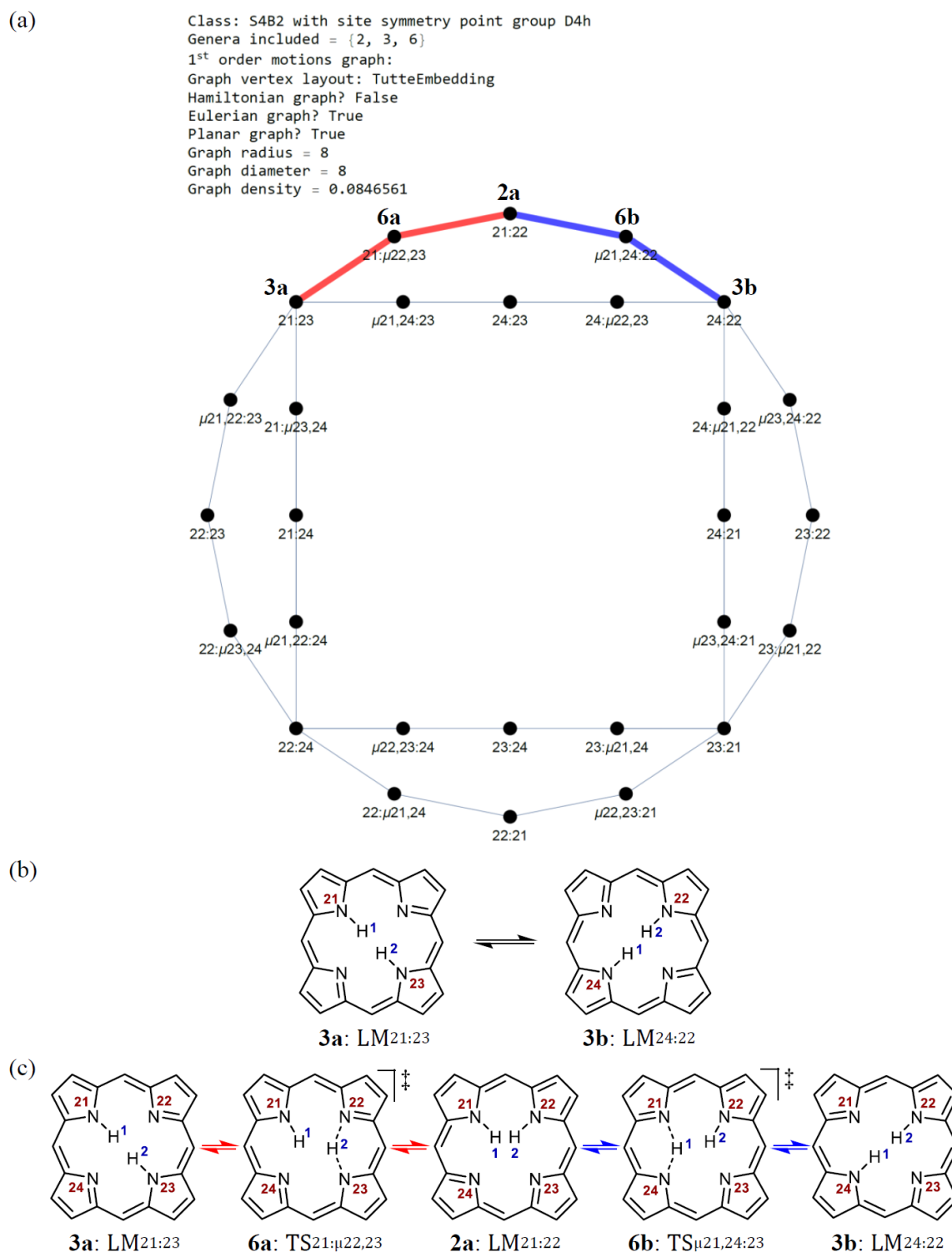
Figure 6.5 Example output from the pictorial graphing software and how it relates to isomerisation processes. (a) Output for the first-order motions graph of all species belonging to genera **2**, **3**, and **6** from the D_{4h} site symmetric $\mathcal{S}_4\mathcal{B}_2$ family. (b) A degenerate *trans* ⇌ *trans'* rearrangement of **3a** and **3b** representative of experimentally observed tautomerisation of free-base porphyrin. (c) A shortest allowed elementary pathway involving three intermediate species belonging to genera **6** and **2**. This pathway has been highlighted (red and blue) and annotated on the graph (a).

200

6.5.3   From a catalogue to representative molecular structures.

Starting with the intrinsic form of each bond topology species (see Figure 6.3), it is possible to generate a simple atomic representation of the sites and bonders. More than one geometrical structure may be possible for each bond topology species as isomers are defined not only by the bond topology but also by site stereoisomerism, bonder stereoisomerism, and global properties of the whole molecular system. As each aspect presents a finite number of possible structural variations, the total number of conceived chemical structures remains countable.

All structures envisaged by this procedure may not be chemically feasible, however, owing to steric constraints and the need to link the bonding sites together (and possibly also the bonders), see Figure 6.2. If a realistic complete chemical model can be constructed, then low-level and/or high-level computational schemes could be used to convert this into a chemical structure worthy of consideration.

It is possible that the entire process of determination of the bond topology species, adding of variations associated with site and bonder stereochemistry, and implementation of global molecular properties can be computer automated. This applies also to optimisation procedures for obtaining quantitatively useful structures. Hence it is possible to fully automate computational searches for isomers and reactions, with very broad applicability throughout chemical research.

In databases used for searching chemical structures and sub-structures like the CAS database, the "chemical structures" are stored as a connectivity list or matrix, i.e., they are digitally represented as a bond topology list augmented by bond multiplicities and stereodescriptors. Further, systematic chemical names can be automatically generated from this same digital representation. The bond-topology permuted structures that the Polytope Formalism of constitutional isomerism (or its $\mathcal{S}_m\mathcal{B}_n$ partitioned approach described here) generates provides "structures" using the same basic representations as already used in structure databases. Consequently, many existing cheminformatics tools are expected to be readily adaptable to implement the principles and concepts described here.

6.6   Detailed results for tautomerism in free-base porphyrin.

The appropriate class for tautomerism in free-base porphyrin is $\mathcal{S}_4\mathcal{B}_2$. For this, Equations (3a) and (4a) indicate that there are $N_s = 255$ species and $N_f = 10$ families. The site topology symmetry appropriate to this molecule is the D_{4h} symmetry point group. Assuming this, there are $N_g = 28$ possible genera, each of which corresponds to a different qualitative bonding scenario. Previously, three example species were presented in Figure 6.3, their corresponding genera in Figure 6.4, and the description of their family relationships in the accompanying discussion. Here, a complete tabulation of all 28 genera is shown in Figure 6.6

In Figure 6.6 the genera are grouped into families F_1 through to F_{10} by the different general bonder linking patterns, with this being indicated by colour coding and descriptive legend. For each genus two representations are given: (i) the general $\mathcal{S}_4\mathcal{B}_2$ partitioning form and (ii) this depicted as a 2D molecular structure.

All $N_s = 225$ possible species are listed in Table 6.2 grouped by genus and family. The set definition of each genus is listed using the following approach:

For $\mathcal{S}_m\mathcal{B}_n$, let S be the set of bonder site atom locants s_i. For tautomerism in free base porphyrin partitioned as $\mathcal{S}_4\mathcal{B}_2$, $m = 4$, thus:

$$S = \{s_0, s_1, s_2, s_3\} \equiv \{s_{i \bmod 4}\} \text{ where } i \in \mathbb{Z}$$

For free base porphyrin and using standard IUPAC porphyrin atom locants:

$$s_0 = 21, s_1 = 22, s_2 = 23, \text{ and } s_3 = 24.$$

Given the D_{4h} topological site symmetry, the following relations apply:

$$s_{i+1} \not\equiv s_{i-1} \text{ and } s_{i+2} \equiv s_{i-2} \text{ with } i = i \bmod 4$$

For convenience let $S \backslash i$ be the set difference of S and i. For example, $S \backslash s_1 = \{s_0, s_2, s_3\}$ or specifically for free base porphyrin, $S \backslash 22 = \{21, 23, 24\}$. This is "hole" notation.

From the genus definitions, bonder-site locant lists using the adapted μ notation as outlined in Section 6.5.1 are generated and given in the Table 6.2.

Figure 6.6 All bond topology genera for free-base porphyrin tautomerism. This ustilises the $\mathcal{S}_4\mathcal{B}_2$ partitioning with D_{4h} site symmetry and assuming non-dissociative topologies. All $N_g = 28$ possible genera are shown as both their general topology-only $\mathcal{S}_4\mathcal{B}_2$ partitioned form and as feasible free-base porphyrin 2D chemical structures. The genera are color-coded into $N_f = 10$ families F_1 through to F_{10}, depicting different bonder linking patterns as described in the legend.

Table 6.2 Listing of families, genera (D_{4h} site symmetry), and species for $\mathcal{S}_m\mathcal{B}_n$, partitioning as is appropriate for describing tautomerism of free-base porphyrin. For each genus definition, $i \in \mathbb{Z}$ with $i = i$ mod 4.

Genus entry	Family	Canon. genus	N'ber	Genus definitions	Species locant lists
Both H^1 and H^2 non-bridging: total $= 1 \binom{4}{1}\binom{4}{1} = 1 \times 4 \times 4 = 16$					
1	F_1	1	4	$(i{:}i)$	(21:21) (22:22) (23:23) (23:23)
2	F_1	2	8	$(i{:}i{\pm}1)$	(21:22) (22:23) (23:24) (24:21) (22:21) (23:22) (24:23) (21:24)
3	F_1	3	4	$(i{:}i{+}2)$	(21:23) (22:24) (23:21) (24:22)
One H non-bridging, one H μ bridging: total $= 2 \binom{4}{1}\binom{4}{2} = 2 \times 4 \times 6 = 48$					
4	F_2	1	16	$(i{:}\mu i,i{\pm}1)$ $(i{:}\mu i{\pm}1,i)$	(21:μ21,22) (21:μ21,24) (22:μ22,23) (22:μ21,22) (23:μ23,24) (23:μ22,23) (24:μ21,24) (24:μ23,24) (μ21,22:21) (μ21,24:21) (μ22,23:22) (μ22,21:22) (μ23,24:23) (μ22,23:23) (μ21,24:24) (μ23,24:24)
5	F_2	2	8	$(i{:}\mu i,i{+}2)$ $(\mu i,i{+}2{:}i)$	(21:μ21,23) (22:μ22,24) (23:μ21,23) (24:μ22,24) (μ21,23:21) (μ22,24:22) (μ21,23:23) (μ22,24:24)
6	F_2	3	16	$(i{:}\mu i{\pm}1,i{+}2)$ $(\mu i{\pm}1,i{+}2{:}i)$	(21:μ22,23) (21:μ23,24) (22:μ23,24) (22:μ21,24) (23:μ21,24) (23:μ21,22) (24:μ21,22) (24:μ22,23) (μ23,22:21) (μ23,24:21) (μ23,24:22) (μ21,24:22) (μ21,24:23) (μ21,22:23) (μ21,22:24) (μ22,23:24)
7	F_2	4	8	$(i{:}\mu i{+}1,i{-}1)$ $(\mu i{+}1,i{-}1{:}i)$	(21:μ22,24) (22:μ21,23) (23:μ22,24) (24:μ21,23) (μ22,24:21) (μ21,23:22) (μ22,24:23) (μ21,23:24)
Both H^1 and H^2 μ bridging: total $= 2 \binom{4}{2}\binom{4}{2} = 1 \times 6 \times 6 = 36$					
8	F_3	1	4	$(\mu i,i{+}1{:}\mu i,i{+}1)$	(μ21,22:μ21,22) (μ22,23:μ22,23) (μ23,24:μ23,24) (μ21,24:μ21,24)
9	F_3	2	16	$(\mu i,i{\pm}1{:}\mu i,i{+}2)$ $(\mu i,i{+}2{:}\mu i,i{\pm}1)$	(μ21,22:μ21,23) (μ21,24:μ21,23) (μ22,23:μ22,24) (μ21,22:μ22,24) (μ23,24:μ21,23) (μ22,23:μ21,23) (μ21,24:μ22,24) (μ23,24:μ22,24) (μ21,23:μ21,22) (μ21,23:μ21,24) (μ22,24:μ22,23) (μ22,24:μ21,22) (μ21,23:μ23,24) (μ21,23:μ22,23) (μ22,24:μ21,24) (μ22,24:μ23,24)
10	F_3	3	8	$(\mu i,i{\pm}1{:}\mu i,i{\mp}1)$	(μ21,22:μ21,24) (μ22,23:μ21,22) (μ23,24:μ22,23) (μ21,24:μ23,24) (μ21,24:μ21,22) (μ21,22:μ22,23) (μ22,23:μ23,24) (μ23,24:μ21,24)
11	F_3	4	4	$(\mu i,i{+}1{:}\mu i{-}1,i{+}2)$	(μ21,22:μ23,24) (μ22,23:μ21,24) (μ23,24:μ21,22) (μ21,24:μ22,23)
12	F_3	5	2	$(\mu i,i{+}2{:}\mu i,i{+}2)$	(μ21,23:μ21,23) (μ22,24:μ22,24)
13	F_3	6	2	$(\mu i,i{+}2{:}\mu i{+}1,i{-}1)$	(μ21,23:μ22,24) (μ22,24:μ21,23)
One H non-bridging, one H μ_3 bridging: total $= 2 \binom{4}{1}\binom{4}{3} = 2 \times 4 \times 4 = 32$					
14	F_4	1	8	$(i{:}\mu_3 S{\backslash}i)$ $(\mu_3 S{\backslash}i{:}i)$	(21:μ$_3$22,23,24) (22:μ$_3$21,23,24) (23:μ$_3$21,22,24) (24:μ$_3$21,22,23) (μ$_3$22,23,24:21) (μ$_3$21,23,24:22) (μ$_3$21,22,24:23) (μ$_3$21,22,23:24)
15	F_4	2	16	$(i{:}\mu_3 S{\backslash}i{+}1)$ $(\mu_3 S{\backslash}i{+}1{:}i)$	(21:μ$_3$21,22,23) (22:μ$_3$22,23,24) (23:μ$_3$21,23,24) (24:μ$_3$21,22,24) (21:μ$_3$21,23,24) (22:μ$_3$21,22,24) (23:μ$_3$21,22,23) (24:μ$_3$22,23,24) (μ$_3$21,22,23:21) (μ$_3$22,23,24:22) (μ$_3$21,23,24:23) (μ$_3$21,22,24:24) (μ$_3$21,23,24:21) (μ$_3$21,22,24:22) (μ$_3$21,22,23:23) (μ$_3$22,23,24:24)
16	F_4	3	8	$(i{:}\mu_3 S{\backslash}i{+}2)$ $(\mu_3 S{\backslash}i{+}2{:}i)$	(21:μ$_3$21,22,24) (22:μ$_3$21,22,23) (23:μ$_3$22,23,24) (24:μ$_3$21,23,24) (μ$_3$21,22,24:21) (μ$_3$21,22,23:22) (μ$_3$22,23,24:23) (μ$_3$21,23,24:24)

Table 6.2 *continued*. Listing of families, genera (D_{4h} site symmetry), and species for $\mathcal{S}_m\mathcal{B}_n$, partitioning as is appropriate for describing tautomerism of free-base porphyrin. For each genus definition, $i \in \mathbb{Z}$ with $i = i$ mod 4.

Genus entry	Family	Canon. genus	N'ber	Genus definitions	Species locant lists
One H μ bridging, one H μ_3 bridging: total = $2\binom{4}{2}\binom{4}{3} = 2 \times 6 \times 4 = 48$					
17	F_5	1	16	$(\mu i, i\pm1 : \mu_3 S\backslash i\pm1)$ $(\mu_{3L} S\backslash i\pm1 : \mu i, i\pm1)$	$(\mu21,22:\mu_321,23,24)$ $(\mu22,23:\mu_321,22,24)$ $(\mu23,24:\mu_321,22,23)$ $(\mu21,24:\mu_322,23,24)$ $(\mu21,24:\mu_321,22,23)$ $(\mu21,22:\mu_322,23,24)$ $(\mu22,23:\mu_321,23,24)$ $(\mu23,24:\mu_321,22,24)$ $(\mu_321,23,24:\mu21,22)$ $(\mu_321,22,24:\mu22,23)$ $(\mu_321,22,23:\mu23,24)$ $(\mu_322,23,24:\mu21,24)$ $(\mu_321,22,23:\mu21,24)$ $(\mu_322,23,24:\mu21,22)$ $(\mu_321,23,24:\mu22,23)$ $(\mu_321,22,24:\mu23,24)$
18	F_5	2	16	$(\mu i, i\pm1 : \mu_3 S\backslash i\mp1)$ $(\mu_{3L} S\backslash i\mp1 : \mu i, i\pm1)$	$(\mu21,22:\mu_321,22,23)$ $(\mu22,23:\mu_322,23,24)$ $(\mu23,24:\mu_321,23,24)$ $(\mu21,24:\mu_321,22,24)$ $(\mu_321,22,23:\mu21,22)$ $(\mu_322,23,24:\mu22,23)$ $(\mu_321,23,24:\mu23,24)$ $(\mu_321,22,24:\mu21,24)$ $(\mu21,24:\mu_321,23,24)$ $(\mu21,22:\mu_321,22,24)$ $(\mu23,22:\mu_321,22,23)$ $(\mu21,24:\mu_322,23,24)$ $(\mu_321,23,24:\mu21,24)$ $(\mu_321,22,24:\mu21,22)$ $(\mu_321,22,23:\mu22,23)$ $(\mu_322,23,24:\mu21,24)$
19	F_5	3	8	$(\mu i, i+2 : \mu_3 S\backslash i)$ $(\mu_3 S\backslash i : \mu i, i+2)$	$(\mu21,23:\mu_322,23,24)$ $(\mu22,24:\mu_321,23,24)$ $(\mu21,23:\mu_321,22,24)$ $(\mu22,24:\mu_321,22,23)$ $(\mu_322,23,24:\mu21,23)$ $(\mu_321,23,24:\mu22,24)$ $(\mu_321,22,24:\mu21,23)$ $(\mu_321,22,23:\mu22,24)$
20	F_5	4	8	$(\mu i, i+2 : \mu_3 S\backslash i+1)$ $(\mu_3 S\backslash i+1 : \mu i, i+2)$	$(\mu21,23:\mu_321,23,24)$ $(\mu22,24:\mu_321,22,24)$ $(\mu21,23:\mu_321,22,23)$ $(\mu22,24:\mu_322,23,24)$ $(\mu_321,23,24:\mu21,23)$ $(\mu_321,22,24:\mu22,24)$ $(\mu_321,22,23:\mu21,23)$ $(\mu_322,23,24:\mu22,24)$
Both H¹ and H² μ_3 bridging: total = $1\binom{4}{3}\binom{4}{3} = 1 \times 4 \times 4 = 16$					
21	F_6	1	4	$(\mu_3 S\backslash i : \mu_3 S\backslash i)$	$(\mu_321,22,23:\mu_321,22,23)$ $(\mu_322,23,24:\mu_322,23,24)$ $(\mu_321,23,24:\mu_321,23,24)$ $(\mu_321,22,24:\mu_321,22,24)$
22	F_6	2	8	$(\mu_3 S\backslash i : \mu_3 S\backslash i\pm1)$	$(\mu_322,23,24: \mu_321,23,24)$ $(\mu_321,23,24: \mu_322,23,24)$ $(\mu_321,23,24: \mu_321,22,24)$ $(\mu_321,22,24: \mu_321,23,24)$ $(\mu_321,22,24: \mu_321,22,23)$ $(\mu_321,22,23: \mu_321,22,24)$ $(\mu_321,22,23: \mu_322,23,24)$ $(\mu_322,23,24: \mu_321,22,23)$
23	F_6	3	4	$(\mu_3 S\backslash i : \mu_3 S\backslash i+2)$	$(\mu_321,22,23:\mu_322,23,24)$ $(\mu_322,23,24:\mu_321,23,24)$ $(\mu_321,23,24:\mu_321,22,24)$ $(\mu_321,22,24:\mu_321,22,23)$
One H non-bridging, one H μ_4 bridging: total = $2\binom{4}{1}\binom{4}{4} = 2 \times 4 \times 1 = 8$					
24	F_7	1	8	$(i : \mu_4 S)$ $(\mu_4 S : i)$	$(21:\mu_421,22,23,24)$ $(22:\mu_421,22,23,24)$ $(23:\mu_421,22,23,24)$ $(24:\mu_421,22,23,24)$ $(\mu_421,22,23,24:21)$ $(\mu_421,22,23,24:22)$ $(\mu_421,22,23,24:23)$ $(\mu_421,22,23,24:24)$
One H μ bridging, one H μ_4 bridging: total = $1\binom{4}{2}\binom{4}{4} = 1 \times 6 \times 1 = 12$					
25	F_8	1	8	$(\mu i, i+1 : \mu_4 S)$	$(\mu21,22:\mu_421,22,23,24)$ $(\mu22,23:\mu_421,22,23,24)$ $(\mu23,24:\mu_421,22,23,24)$ $(\mu21,24:\mu_421,22,23,24)$ $(\mu_421,22,23,24:\mu21,22)$ $(\mu_421,22,23,24:\mu22,23)$ $(\mu_421,22,23,24:\mu23,24)$ $(\mu_421,22,23,24:\mu21,24)$
26	F_8	2	4	$(\mu i, i+2 : \mu_4 S)$	$(\mu21,23:\mu_421,22,23,24)$ $(\mu22,24:\mu_421,22,23,24)$ $(\mu_421,22,23,24:\mu21,23)$ $(\mu_421,22,23,24:\mu22,24)$
One H μ_3 bridging, one H μ_4 bridging: total = $2\binom{4}{3}\binom{4}{4} = 2 \times 4 \times 1 = 8$					
27	F_9	1	8	$(\mu_3 S\backslash i\pm1 : \mu_4 S)$ $(\mu_4 S : \mu_3 S\backslash i\pm1)$	$(\mu_321,22,23:\mu_421,22,23,24)$ $(\mu_324,22,23:\mu_421,22,23,24)$ $(\mu_321,23,24:\mu_421,22,23,24)$ $(\mu_321,22,24:\mu_421,22,23,24)$ $(\mu_421,22,23,24:\mu_321,22,23)$ $(\mu_421,22,23,24:\mu_322,23,24)$ $(\mu_421,22,23,24:\mu_321,23,24)$ $(\mu_421,22,23,24:\mu_321,22,24)$
Both H¹ and H² μ_4 bridging: total = $1\binom{4}{4}\binom{4}{4} = 1 \times 1 \times 1 = 1$					
28	F_{10}	1	1	$(\mu_{4L} S : \mu_4 S)$	$(\mu_421,22,23,24:\mu_421,22,23,24)$

6.6.1 Insight gained into properties of free-base porphyrin.

DFT-optimised Cartesian coordinates (see Methods Section 6.10 for details) depicting all 28 genera of porphyrin are listed in Appendix E. Chemical intuition indicates that structures of interest for free-base porphyrin and related molecules can be divided into two categories:

i) "distal structures", in which both H-atom bonders are distant from the ring centre, and

ii) "cross-bonded structures", in which at least one bonder is close to the ring centre.

Distal structures can be characterised primarily based on the rotation angle of the two hydrogens around the molecular plane. The technical details as to how to implement this concept for general molecules that may be warped and distorted in many ways are discussed in detail in the Methods Section 6.10, including the instructions needed for implementation of the adaptive coordinate system into Gaussian16[153] in Appendix I. Key features are presented in Figure 6.7, including the radii r_1 and r_2 of the two hydrogens from the ring centre and the angles ϕ_1 and ϕ_2 depicting their rotation within the macrocycle plane.
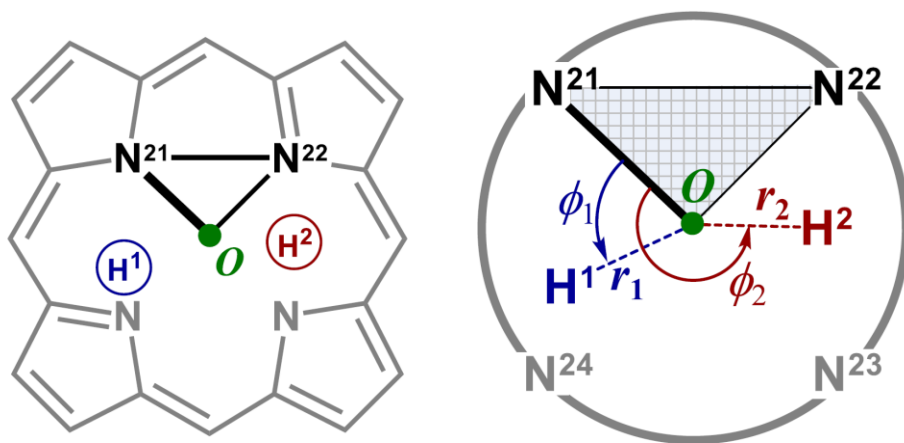
Figure 6.7 Critical polar coordinates r_1, r_2, ϕ_1, and ϕ_2 used to describe the locations of the two inner-coordinated hydrogen atoms. *O* is the origin, defined as the barycentre of the four nitrogen atoms, with the polar coordinates defined with respect to the N^{21}–N^{22}–*O* plane (using IUPAC porphyrinoid naming conventions[150]) and its *O*–N^{21} vector. The associated polar angles θ_1 and θ_2 (not shown) depict motion of the hydrogen atoms out of the reference plane.

The result of this coordinate transformation is that it is possible to construct a PES $E(\phi_1, \phi_2)$ provided that the distal constraint r_1, $r_2 > 0.95$ Å holds. In Appendix E, Table E.1

and Table E.2 list these distance values for DFT optimised structures representing all 28 genera demonstrating that this criterion for grouping structures as distal or cross-bonded holds.

Two representations of the resulting distal-structure PES are shown in Figure 6.8 with critical points labelled by genus numbers used for the $\mathcal{S}_4\mathcal{B}_2$ partitioning of free-base porphyrin. Details of the distal structures are tabulated in Figure 6.9 with results from high-level calculations of their structure and properties given.

To help understand the distal-structure relationships and the nature of the coordinate system, an interactive graphics-based application for exploring the distal-structure space is provided as **E_File_10** with details and instructions in Appendix F. This allows the user to "click" on any point in Figure 6.8a and have the corresponding energy, coordinates (ϕ_1, ϕ_2), and 3D molecular representation displayed.

It is not feasible to search for cross-bonded structures using an analogous systematic approach as the number of geometrical variables needed to define the PES is too great. Hence the general method described earlier to search for realistic chemical structures depicting the properties of the other polytopes is applied, using non-automated techniques. Detailed results and discussion obtained for six selected structures are presented in Figure 6.11, with results for the remaining 14 cross-bonded structures presented in Figure 6.12.

In general, it would not be possible to find chemically sensible example structures corresponding to all genera; the example of free-base porphyrin is exceptional in this regard as only two very small bonders are involved, and the physical space defined by the site cavity is large compared to what the bonders require. Also, the fully conjugated macrocycle π system adapts to all bond-topology configurations.
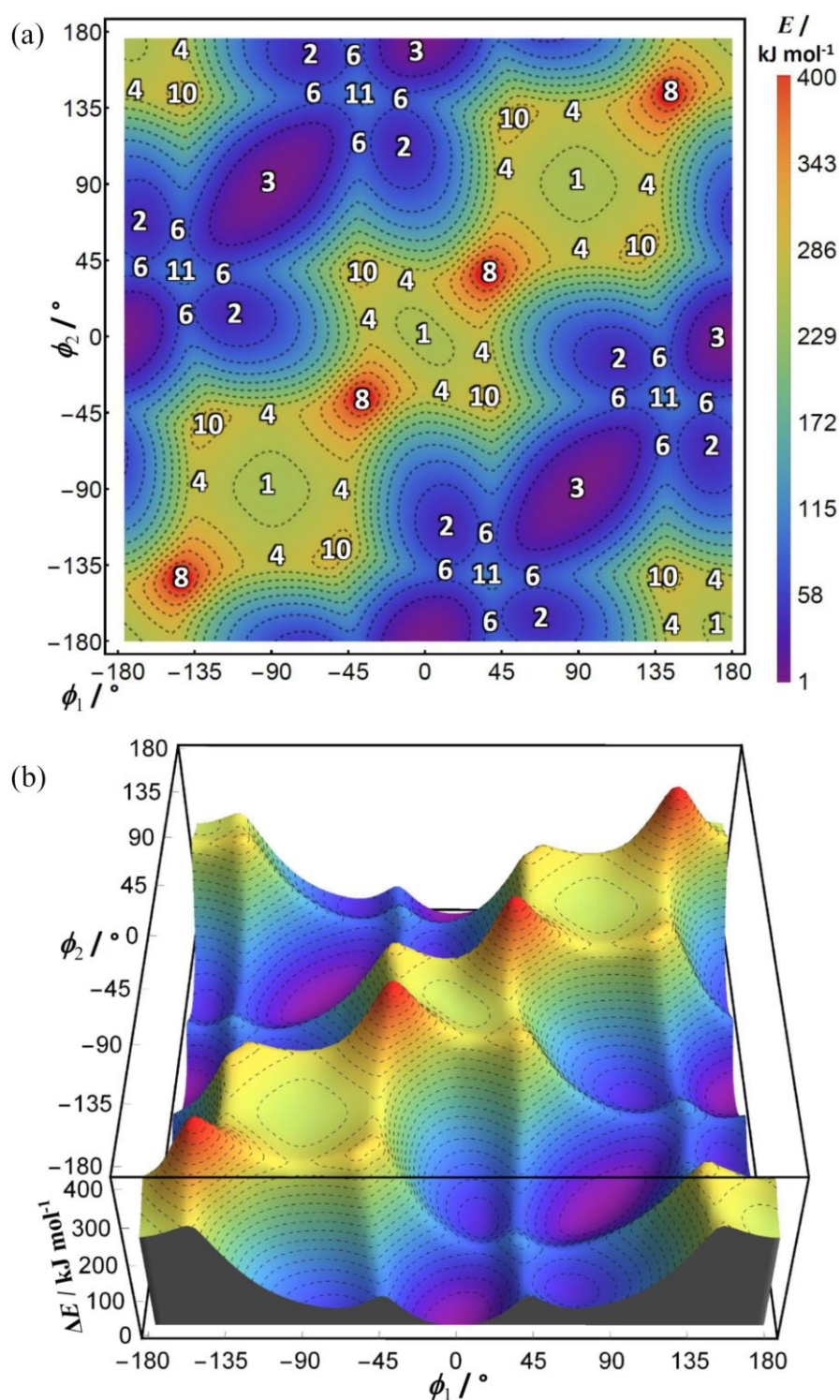
Figure 6.8   Equivalent representations of gas-phase B3LYP/6-31G(d) potential-energy surface $E(\phi_1, \phi_2)$, for free-base porphyrin distal structures with r_1, $r_2 > 0.95$ Å.   (a) The critical points, representing bond-topology species on the 2D contour map are labelled by their genus number.   (b) PES rendered as a 3D landscape with the vertical axis representing the energy values.   Colouring for (b) is the same as for (a).

Type	Description	ΔG / kJ mol^{-1}	Example	Model
Local–minima species (isomers)	**1** LM_{21}^{21} *gem* C_{2v}	225		
	2 $\mathrm{LM}_{21:22}$ *cis* C_{2v} ^{a}	22		
	3 $\mathrm{LM}_{21:23}$ *trans* D_{2h}	[0]		
1st order transition-state species	**4** $\mathrm{TS}_{21}^{\mu 21,22}$ *cis* $\rightleftharpoons$ *gem*$^{\mathrm{adj}}$ C_1	234		
	6 $\mathrm{TS}_{21:\mu 22,23}$ *cis* $\rightleftharpoons$ *trans* C_s	43		
2nd order transition-state species	**8** $2\mathrm{S}_{\mu 21,22}^{\mu 21,22}$ *cis* H-exch. *gem* $\rightleftharpoons$ *gem*$^{\mathrm{adj}}$ C_{2v}	365		
	10 $2\mathrm{S}_{\mu 21,24}^{\mu 21,22}$ gem $\rightleftharpoons$ trans$^{\mathrm{adj}}$ *cis* $\rightleftharpoons$ *cis*$^{\mathrm{adj}}$ C_2	252		
	11 $2\mathrm{S}_{\mu 23,24}^{\mu 21,22}$ trans $\rightleftharpoons$ trans$'$ *cis* $\rightleftharpoons$ *cis*$^{\mathrm{opp}}$ D_{2h} ^{a}	65		

Figure 6.9   Tabulation of the distal structures for free-base porphyrin based on DFT models. "Type" refers to the character of the critical point on the potential energy surface. "Description" gives the genus number, an example of a species within that genus, a relational description of the structure and the symmetry point group of the genus. ΔG is the DFT Gibbs free energy relative to **3**. "Example" gives a 2D representation of the species listed in the column to the left. "Model" is a 3D framework molecular model of the DFT calculated structure. Characterisation based on B2PLYP-D3(BJ)/cc-pVDZ calculations with implicit CHCl$_3$ solvation; grey- C, blue- N (the four sites $\mathcal{S}$), white- peripheral H, orange- the two bonder hydrogens ($\mathcal{B}$). *a*: Trivially small difference in symmetry between DFT methods - highest symmetry point group shown.
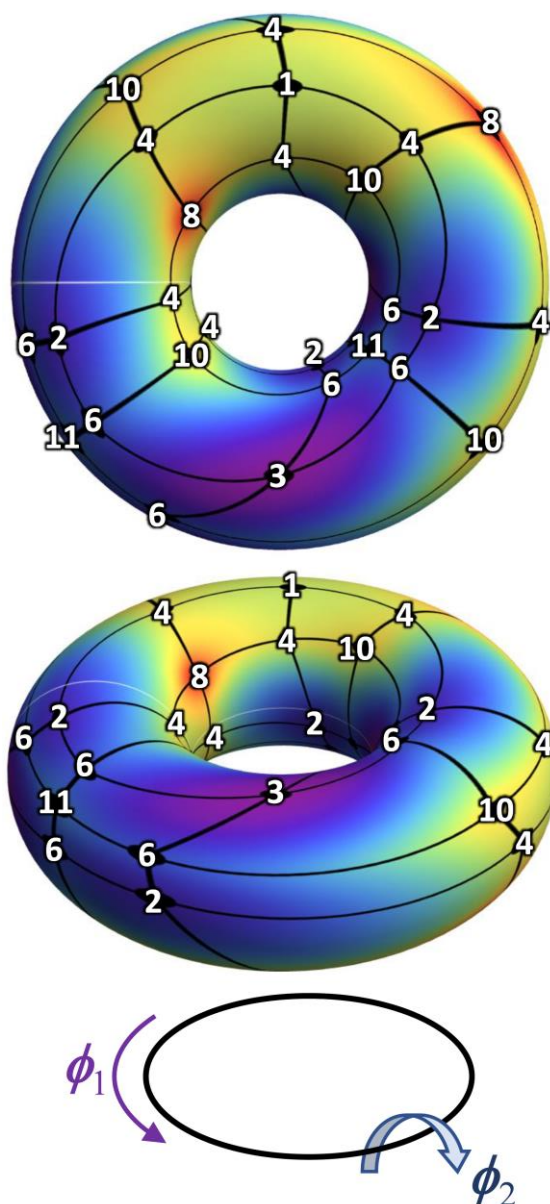
Figure 6.10 Toroidal mapping of the distal structure PES. As both ϕ_1 and ϕ_2 is periodic as indicated, the PES seamlessly maps onto the surface of a torus. The same genus-number labelling as used in Figure 6.8a is applied here for the critical points on the PES. Each critical point is a different bond-topology species and corresponds to a graph vertex. The black lines interconnecting pairs of critical points represent single bond-topology changes (first-order motions of the hydrogen bonders) for the connected species and correspond equally to reaction pathways and graph edges. This demonstrates the direct relationship between a bond-topology graph and the associated PES.

Type	Description	ΔG / kJ mol^{-1}	Example	Model
1st order transition-state species	**5** $\text{TS}_{21}^{\mu 21,23}$ *gem* $\rightleftharpoons$ *trans* C_s	366		
	7 $\text{TS}_{\mu 22,24}^{21}$ *cis* $\rightleftharpoons$ *cis*$^{\text{adj}}$ C_s	279		
2nd order transition-state species	**9** $2\text{S}_{\mu 21,23}^{\mu 21,22}$ *gem* $\rightleftharpoons$ *cis*$^{\text{opp}}$ *cis* $\rightleftharpoons$ *trans*$^{\text{adj}}$ C_1	390		
	12 $2\text{S}_{\mu 21,23}^{\mu 21,23}$ *trans* H-exch. *gem* $\rightleftharpoons$ *gem*$^{\text{opp}}$ D_{2h}	576		
	13 $2\text{S}_{\mu 22,24}^{\mu 21,23}$ $2 \times$ *cis* $\rightleftharpoons$ *cis*$^{\prime}$ D_{4d}	560		
3rd order	**24** $3\text{S}_{\mu_4 21,22,23,24}^{21}$ C_s	375		
4th order	**22** $4\text{S}_{\mu_3 21,22,23}^{\mu_3 21,22,24}$ C_2	604		
5th order	**28** $5\text{S}_{\mu_4 21,22,23,24}^{\mu_4 21,22,23,24}$ D_{4h}	652		

Figure 6.11 Select cross-bonded structures for free-base porphyrin based on DFT models. "Type" refers to the character of the critical point on the potential energy surface. "Description" gives the genus number, an example of a species within that genus, and the symmetry point group of the genus. ΔG is the DFT Gibbs free energy relative to **3**. "Example" gives a 2D representation of the species listed in the column to the left. "Model" is a 3D framework molecular model of the DFT calculated structure. Characterisation is based on B2PLYP-D3(BJ)/cc-pVDZ calculations with implicit CHCl$_3$ solvation; grey- C, blue- N (the four sites $\mathcal{S}$), white- peripheral H, orange- the two bonder hydrogens ($\mathcal{B}$).

211

Type	Example	Model	Type	Example	Model
14 $2S^{21}_{\mu_3 22,23,24}$ C_s			**20** $NC^{\mu_3 21,22,23}_{\mu 21,23}$ C_s		
15 $2S^{21}_{\mu_3 21,22,23}$ C_1			**21** $4S^{\mu_3 21,22,23}_{\mu_3 21,22,23}$ C_s		
16 $2S^{21}_{\mu_3 21,22,24}$ C_s			**23** $NC^{\mu_3 21,22,24}_{\mu_3 21,23,24}$ C_{2h}		
17 $NC^{\mu 21,22}_{\mu_3 22,23,24}$ C_1			**25** $4S^{\mu 21,22}_{\mu_4 21,22,23,24}$ C_s		
18 $3S^{\mu 21,22}_{\mu_3 21,22,23}$ C_1			**26** $4S^{\mu 21,23}_{\mu_4 21,22,23,24}$ C_{2v}		
19 $3S^{\mu_3 21,22,24}_{\mu 21,23}$ C_s			**27** $5S^{\mu_3 21,22,23}_{\mu_4 21,22,23,24}$ C_s		

Figure 6.12 Remaining cross-bonded structures not shown in Figure 6.11 for free-base porphyrin based on DFT models. "Description" gives the genus number, an example of a species within that genus, and the symmetry point group of the genus. ΔG is the DFT Gibbs free energy relative to **3**. "Example" gives a 2D representation of the species listed in the column to the left. "Model" is a 3D framework molecular model of the DFT calculated structure. Characterisation is based on B2PLYP-D3(BJ)/cc-pVDZ calculations with implicit $CHCl_3$ solvation; grey- C, blue- N (the four sites $\mathcal{S}$), white- peripheral H, orange- the two bonder hydrogens ($\mathcal{B}$).

6.6.2 *Graph and potential energy surface relationship*

For the distal-structure PES shown in Figure 6.8, each of the coordinates ϕ_1 and ϕ_2 is periodic and hence the resulting surface can be mapped onto a torus as shown in Figure 6.10. On this PES, all critical points, labelled with their corresponding genus numbers, represent different bond-topology species. Equally so, they represent graph vertices as required by the Polytope Formalism of constitutional isomerism. Figure 6.10 also shows black lines directly linking a species to its four bond-topologically nearest neighbouring species. These correspond to a *single* difference bond topology change for *just one* of the hydrogen bonders, this being referred to as a "first-order motion" $R_{to}^{c}1$ reactions. For example, in Figure 6.13 is shown the familiar example of *trans* species **3a**, changing to the first-order transition-state structure species **6a**, and then to *cis* species **2a**.

3a *trans*	$\rightleftharpoons$	**6a** *cis* $\rightleftharpoons$ *trans*	$\rightleftharpoons$	**2a** *cis*
LM21:23	$\rightleftharpoons$	TS21:μ22,23	$\rightleftharpoons$	LM21:22
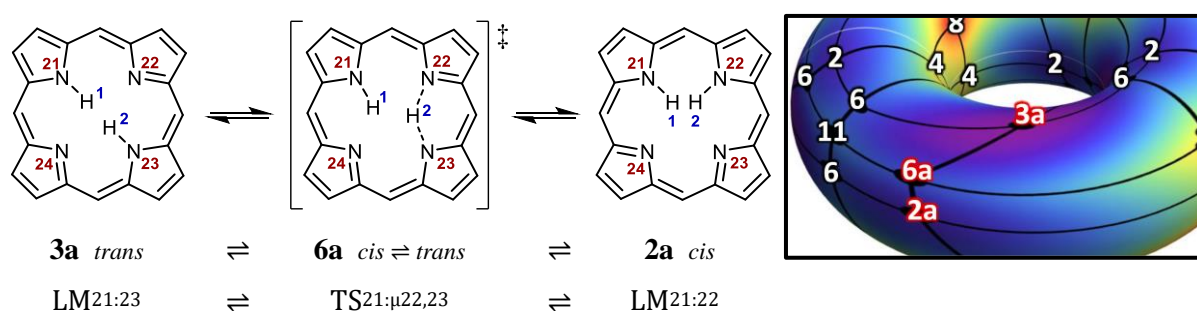
Figure 6.13 $R_{To}^{c}1$ reaction of *trans* species **3a** to *cis* species **2a** *via* the intermediate transition-state structure **6a** representative of first-order motions on the corresponding graph. Each step of this $R_{to}^{c}1$ reaction involves a single bond-topology change for just one hydrogen bonder, in this case H(2). The compact notation for the species makes the bond-topology changes obvious. The inset shows a portion of the toroidal mapping of the distal-structures PES from Figure 6.10 with these three species highlighted in red. The species on the PES are graph vertices and the black lines directly interconnecting them are graph edges. The graph edges here directly indicate first order motions $R_{to}^{c}1$ reactions.

In the forward direction, the first step of this reaction involves only a single bond-topology change for H(2) going from bonding to N(23) to bridging N(23) and N(22). The compact notation LM21:23 $\rightarrow$ TS21:μ22,23 makes this easy to track. Further, the number of imaginary frequencies characterising the critical point undergoes a change of 1 from a local minimum (zero imaginary frequencies) to a first order transition-state structure (one imaginary frequency). Both the single change in bond topology and the ± 1 change in critical point character gives this kind of reaction its description as a "first-order motion" $R_{to}^{c}1$ reaction. The changes are analogous for the second step of the reaction in Figure 6.13 with

$TS_{21:\mu22,23} \rightarrow LM_{21:22}$, again a single change in bond topology and the ±1 change in critical point character. The full first-order motions $R_{to}^c 1$ reactions graph for the distal structures is that given earlier in Figure 6.5. The graph in Figure 6.5 and that shown as a PES in Figure 6.10 have a one-to-one correspondence – they are *isomorphic*. This demonstrates the direct relationship between a bond-topology graph (comprising species and $R_{to}^c 1$ reactions) and the PES.

Higher order motions are also conceivable. For the example of free-base porphyrin, there are two bonders thus "second-order motions" are also possible. Second-order motions $R_{to}^c 1$ reactions are where both hydrogen atom bonders each undergo a single bond-topology change. Correspondingly, the character of the critical point changes by ±2 imaginary frequencies. An example for the distal structures is *trans* species **3a** changing to the second-order transition-state structure species **11a** and then onto *trans'* species **3b** via $R_{to}^c 1$ reactions as shown in Figure 6.14.

3a *trans*	$\rightleftharpoons$	**11a** *trans ⇌ trans'*	$\rightleftharpoons$	**3b** *trans'*
		cis ⇌ cis^{opp}		
$LM_{21:23}$	$\rightleftharpoons$	$2S_{\mu21,24:\mu22,23}$	$\rightleftharpoons$	$LM_{24:22}$
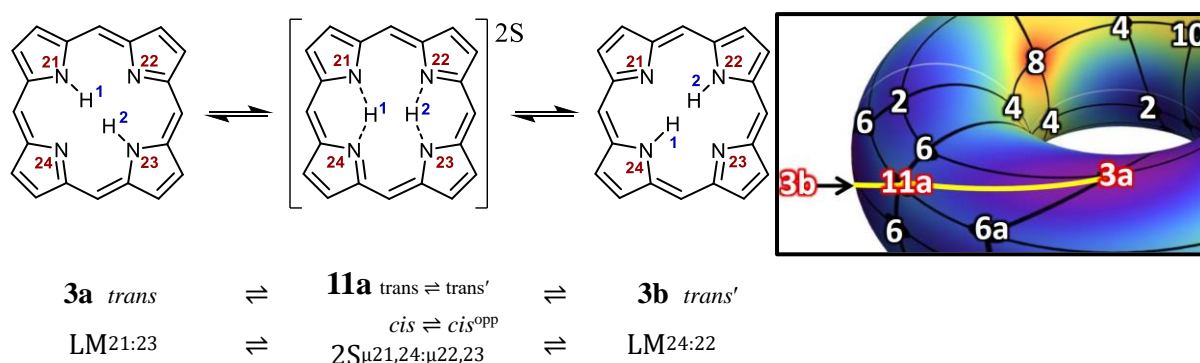
Figure 6.14 $R_{to}^c 1$ reaction of *trans* species **3a** to *trans'* species **3b** *via* the intermediate second-order transition-state structure **11a** representative of second-order motions on the corresponding graph. Each step of this $R_{to}^c 1$ reaction involves a single bond-topology change for both hydrogen bonders. The compact notation for the species makes the bond-topology changes obvious. The inset shows a portion of the toroidal mapping of the distal-structures PES from Figure 6.10 with these three species highlighted in red. The species on the PES are graph vertices and the yellow "diagonal" lines as graph edges, here specifically indicating second order motions $R_{to}^c 1$ reactions.

Again, the compact notation makes the topology changes for the reaction easy to track and characterise. On the toroidal mapping of the distal-structures PES from Figure 6.14, second-order motions are indicated by the yellow "diagonal" edges joining the species indicated.

Both first- and second-order motions can be combined into a single graph representing all possible relationships corresponding to allowed reaction pathways. In Appendix H, Figure H.2 is shown the *combined* first- and second-order motions graph. Although the layout of this graph is a 2D projection, the toroidal form is obvious as would be expected for the distal structures.

The graphing software provided in **E_File_12** makes exploring the first- and second-order motions relationships between the bond-topology species easy. This software also generates of an extensive range of Graph Theory related data as detailed in Appendix H.

6.6.3 *Small chemical variations can induce dramatic changes in polytope energy-ordering.*

Chemical changes to the basic nature of the macrocycle, as well as substituent effects, will modulate the resulting potential-energy landscape from that of free-base porphyrin. In principle, *any* of the 28 structural polytope genera could provide stable isomers. Table 6.3 lists 12 structures that have been identified or considered, as isomers of porphyrin derivatives, with details provided in **E_File_14**. Note that the highest-energy genus for free-base porphyrin, **28**, constitutes the global minimum isomer when the inner H is substituted by alkali metals. Once the genus is assigned to a compound, all possible unimolecular reaction mechanisms can then be read off its graph. These results stress the importance of being able to catalogue all possible structures that could possibly form isomers.

As described earlier in Section 6.5, species and their corresponding genera and families can be generated based upon the condition of non-dissociation (Equations 3a and 4a) or allowing for dissociation (Equations 3b and 4b), the former being used for the discussion of tautomerism in free-base porphyrin.

From the perspective of structural possibilities for porphyrin and porphyrin-like compounds, small chemical changes to the basic nature of the macrocycle, as well as substituent effects, will change the energy landscape from that of free-base porphyrin. In principle, *any* of the 28 topological polytope genera could provide stable isomer-analogues instead of genus **3**. Some pertinent examples are listed below with the substituted porphyrin/porphyrinoid compounds identified by the free-base porphyrin genus number hyphenated to a letter, *e.g.*, **2-b**. This is not to be confused with free-base porphyrin *species*

which are unhyphenated. Only the macrocycle and minimal details of the inner substituent is shown for clarity. Full calculation details are included along with the molecular coordinates in **E_File_14**

Table 6.3 Some isomers identified or proposed pertaining to the corresponding genera of free-base porphyrin tautomerisation using the $\mathcal{S}_4\mathcal{B}_2$ partitioning. Peripheral substituents of the macrocycle are omitted for simplicity and the reader is directed to the original references for full structures. Where indicated, some structures are discussed further in the following discussion with DFT calculated coordinates in the **E_File_14**.

#	Genus	A_1	A_2	Ref.	#	Genus	A_1	A_2	Ref.
1	(structure)	H	H	**1-b**, **1-c** see *a*	12	(structure)	Tl	Tl	[235], see *a*
2	(structure)	H	H	2-b, 2-c [236-238], see *a*	13	(structure)	Mg	PCl₃	[239]
3	(structure)	H Alk	H H	ubiquitous common	14	(structure)	Me H	Cd(II) Re(CO)₃	[240-241], see *a* [242]
4	(structure)	H	CHR	[243]	23	(structure)	Re(CO)₃ Tc(CO)₃ Re(CO)₃ Tl	Re(CO)₃ Tc(CO)₃ Tc(CO)₃ Tl	[242] [242] [242] [244]
6	(structure)	H in a calixaphyrin analogue, [152] see *a*	BF₂	[152]	24	(structure)	Ph Me Ph Alk, Ar Me Me Me Alk Ph Me Alk	Co Co, Ni, Cu Cu, Zn Mn, Fe, Cu Hg Ir Ni Rh Ru Co, Zn Si(IV), Al, Mg	[245] [246] [247] [248] [249] [250] [251] [252] [253] [254] [255], see *a*
8	(structure)	PdCl₂	CHR	[243]					
11	(structure)	BF₂ B PdCl₂ Rh(CO)₂	BF₂ B CHR Rh(CO)₂	[256] [148] [257] [258-259]	28	(structure)	Li Na K Ca Ag(I) Hg(I)	Li Na K Na Ag(I) Hg(I)	[260-262] [262] [262] [263] [264-265] [266]

a: See following discussion and **E_File_14** for details,

Genus **1**

Calculations predict that the carba-oxa-bacteriochlorin **1-b,** and carba-oxa-porphyrintetrone **1-c**, favour the *gem* isomer as most stable relative to *cis* or *trans* forms:

| 2-b : ΔG = 39 kJ mol^{-1} | 1-b : ΔG = [0] | 18 π electron circuit |

| 2-c : ΔG = 26 kJ mol^{-1} | 1-c : ΔG = [0] | 18 π electron circuit |
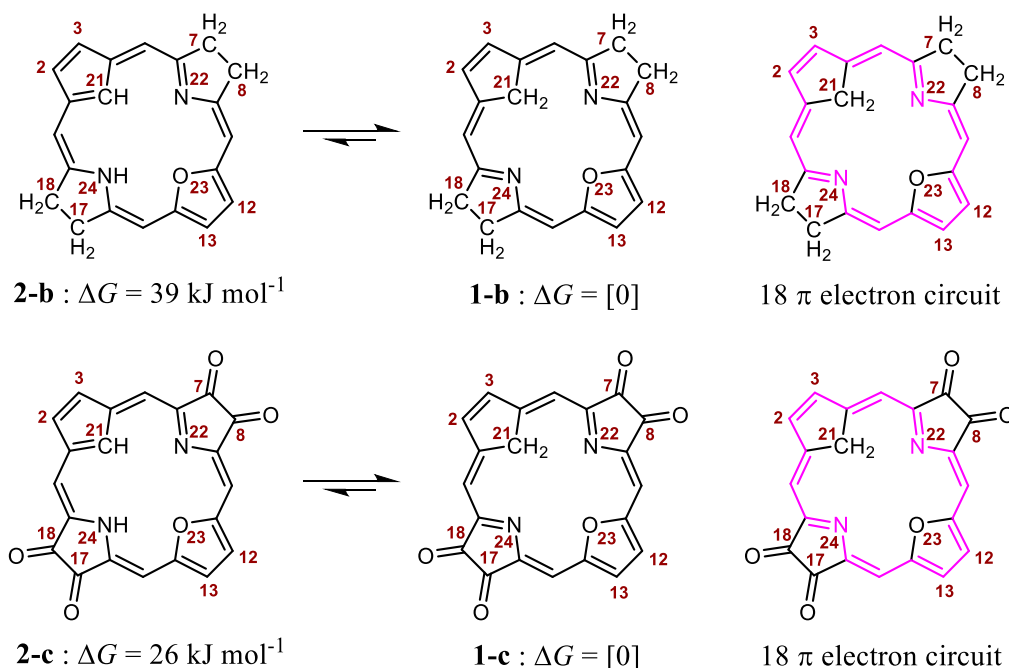
Figure 6.15 The stabilisation of species **1-b** and **1-c** is achieved by (i) to removal of the zwitterionic character evident in **1** by C and O substitution of atoms 21 and 23, respectively, and (ii) modification of β-positions 7, 8, 17, and 18 to restrict the 18-electron π-system exclusively onto the inner N atoms 22 and 24 and outer β-positions 2, 3, 12, and 13 as highlighted.

Genus **2**

Cis isomers of porphyrins may be stabilised by macrocycle distortion and/or intermolecular interactions;[267-269] the relative energy of asymmetric *cis* tautomers GS$^{21|22}$ *etc.* can modulate TS structures like **6** to control reaction directionality for the equilibration of GS$^{21|23}$ and GS$^{22|24}$.[270] Substitution of opposing N atoms with C and O produces *cis* carba-oxa-porphyrins;[236-238] calculations predict these to be more stable than the *gem* tautomer **1-a**:

| 2-a : ΔG = [0] | 1-a : ΔG = 22 kJ mol^{-1} |
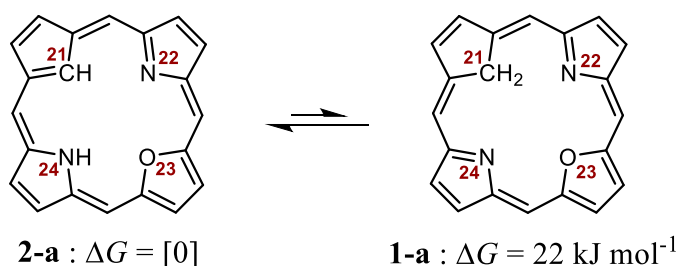
Genus **6**

Although the fully conjugated π-system is bisected, compounds like the BF$_2$-calixaphyrin complex[152] **6-a**, have the same bonding topology as the lowest energy transition structure of free-base porphyrin, **6**.

6-a   **6**
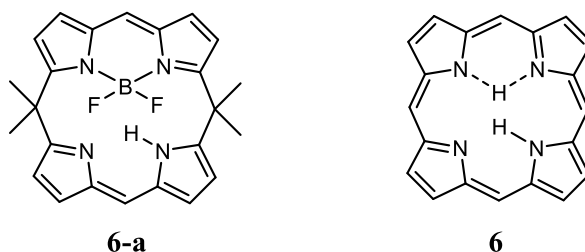
Genus **11**

The high-energy distal genus **11** of free-base porphyrin, involving both bonders bridging two bonding sites, becomes[142, 144-145, 148] the ground-state of the B(F)OB(F) inner-bound porphyrin (see Figure 6.2c) **11-a**.

11-a   **11**
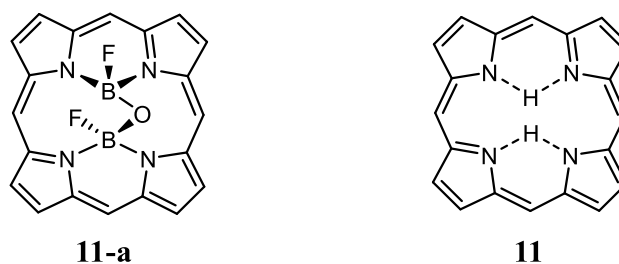
Genus **12**

Reported dithallium(I) complex[235] **12-a,** upon optimisation, collapsed to **28-a**:

12-a   **28-a**

M = Tl(H$_2$O)$_n$ n = 0, 1 or 2
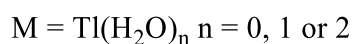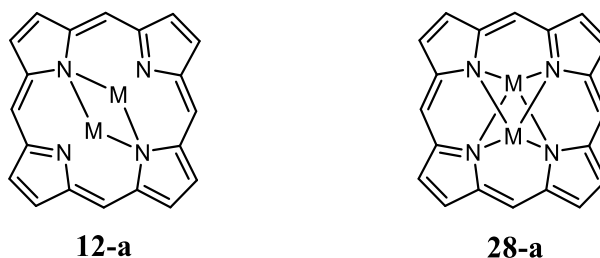
<u>Genera **13** and **14**</u>

The reported Mg/P(III) complex[239] **13-a** and reasonable variants **13-b** and **13-c** were not stable upon optimisation changing to, for example, **14-a**:
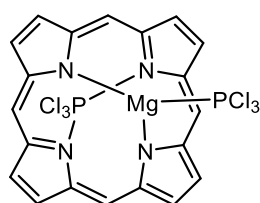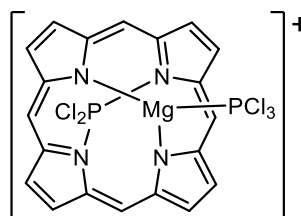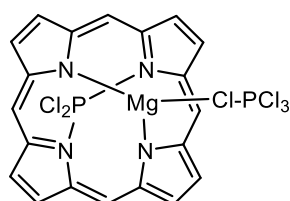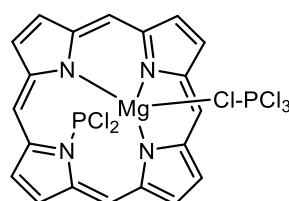
In contrast, modelling of the Cd(II) complex[240-241] **14-b**, and Re(I) complex **14-c**, confirmed the reported structures:
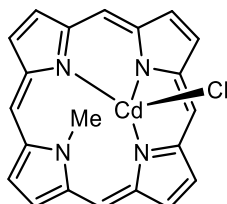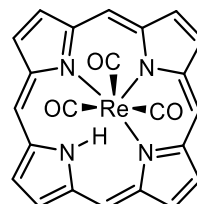
Whilst the close spatial relationship between **14** and **24** makes differentiating between them difficult, the complex **14-c** is an unequivocal example of genus **14**.

<u>Genus **23**.</u>

The di-Re(I) complex[242] **23-a** was modelled, confirming the reported structure:
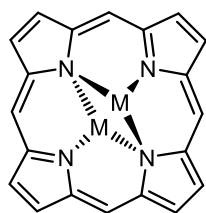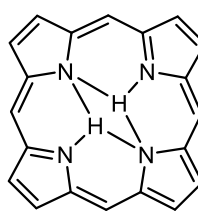
23-a : M = Re(CO)$_3$ **23**

Genus **24**.

The Si(IV) complex[255] **24-b** was modelled, confirming the reported structure:

24-b
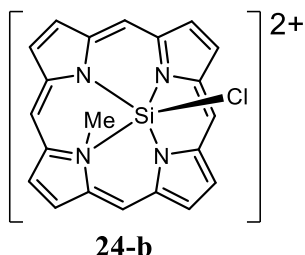
Genus **28**.

The highest-energy structure identified for free-base porphyrin is the macrocyclic ring with an internally trapped "di-hydron", **28**. This has the maximum-possible symmetry, D_{4h}, of all free-base porphyrin tautomerisation species. If H is replaced with Li as **28-a**, then this structure becomes the potential-energy *global minimum* of the new molecular system.[260]

28-a   **28**
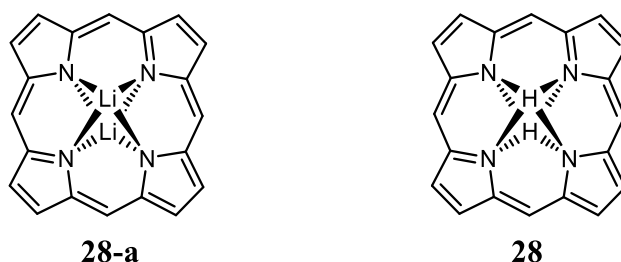
This is an extreme example of what is possible.

6.7   Inclusion of additional chemical constraints where appropriate

A useful extension of the developed formalism is the inclusion of additional chemical constraints such as the maximum number of bonds that any bonding site or bonder can form – this needing to be carefully considered depending upon the nature of the chemical system. Adding such additional restrictions simply reduces the number of possible families, genera, and species from contention. In practice, for larger m and/or n, very many fewer possibilities may need to be considered (the bond-topology graphs become *truncated*). An example of this would be the consideration of doubly protonated free-base porphyrin. This is described by the $\mathcal{S}_4\mathcal{B}_4$ class as now four hydrons are available to bond to the four nitrogen atom

bonding sites.[271]   Figure 6.16 shows four hypothetical example structures for doubly protonated free-base porphyrin dication, labelled **H4p-I** through to **H4p-IV**.
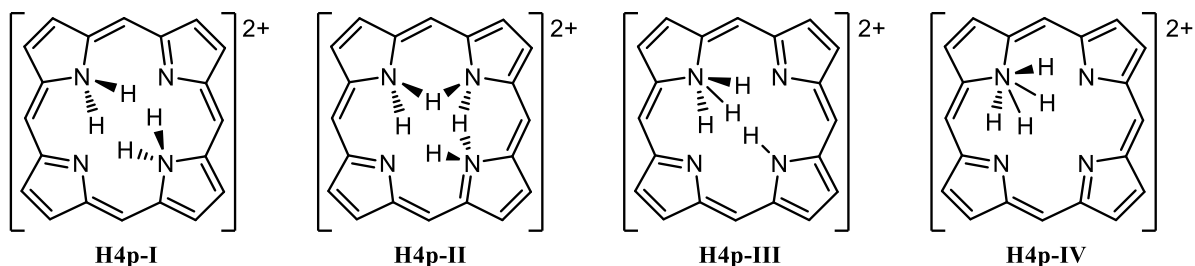
Figure 6.16 Some examples of doubly protonated porphyrin dicationic species. For structures **H4p-I** and **H4p-II**, the maximum coordination number of the nitrogen atoms is 4. For structures **H4p-III** and **H4p-IV**, the maximum coordination number of the nitrogen atoms is 5 and 6, respectively. These latter two structures are unlikely to represent physically realistic molecular entities due to both the intrinsic limitations on valence for nitrogen and for steric constraints imposed by the macrocycle cavity. The latter two structures can be eliminated from the analysis by limiting to 2 the maximum number of $\mathcal{S} - \mathcal{B}$ bonds a nitrogen bonding site can accommodate.

In structures **H4p-I** and **H4p-II**, the maximum coordination number of the nitrogen atoms is four, whilst for **H4p-III** and **H4p-IV**, the maximum coordination number of the nitrogen atoms is five and six, respectively. Based on maximum valence and steric constraints of the macrocycle cavity, for a $\mathcal{S}_4\mathcal{B}_4$ partitioning of this system, a reasonable additional constraint would be to limit to two the maximum number of $\mathcal{S} - \mathcal{B}$ bonds a nitrogen bonding site can accommodate. As such, nonclassical structures (see Section A.32) like **H4p-III** and **H4p-IV** would be excluded from the analysis. Without this additional constraint, there are 494 possible genera as listed in Appendix G Table G.3. With the additional constraint, only 64 genera are now possible, as listed in Appendix G.6, Table G.5. The software provided in **E_File_11** (see Appendix G for details) allows for restriction in the bonding-site coordination, and results for the number of families, genera, and species obtained limiting $\mathcal{S}_m\mathcal{B}_n$ to at most $k = 1$ or 2 bonds to any site are provided in Appendix G Tables G.4 through Table G.9 for most of *m*, *n* ≤ 8.

6.8   Types of chemical systems to which the $\mathcal{S}_m\mathcal{B}_n$ partitioning approach to the Polytope Formalism of constitutional isomerism applies

The $\mathcal{S}_m\mathcal{B}_n$ partitioning approach is intended to apply in a very wide range of chemical applications of practical interest with some examples indicated in Table 6.4.

For coordination centres $\mathcal{S}_1\mathcal{B}_n$ which can be written as ML_n, the only possible bond topology changes are dissociative ones. If bond topology changes are not allowed, the only possible general form of isomerism available here is stereoisomerism for which the Polytope Formalism of stereoisomerism applies.

Isomerism in amino acids like NH_2CH_2COOH is an example of $\mathcal{S}_3\mathcal{B}_3$ partitioning with bonding sites N, O, and O, with three H bonders. H-transfer reactions, to which this type of designation generally applies, in manifold variations, are the subject of a recent *Journal of Physical Chemistry B* special issue.[272]

Table 6.4 Some examples of constitutional isomerism partitionable as $\mathcal{S}_m\mathcal{B}_n$.

Chemical system	$\mathcal{S}_m\mathcal{B}_n$	Bonding Sites ($\mathcal{S}$)	Bonders ($\mathcal{B}$)
coordination centre ML_n	$\mathcal{S}_1\mathcal{B}_n$	M	L
macrocycle, *e.g.* porphyrin	$\mathcal{S}_4\mathcal{B}_2$	4 N atoms	2 inner H atoms
metal cluster compounds, *e.g.*, $Co_2(CO)_8$	$\mathcal{S}_2\mathcal{B}_8$	2 Co	8 CO
amino acid intramolecular hydron transfer	$\mathcal{S}_3\mathcal{B}_3$	N, O, and O	3 H
enzyme – substrate	$\mathcal{S}_m\mathcal{B}_n$	on enzyme	on substrate
Ligand – protein	$\mathcal{S}_m\mathcal{B}_n$	on protein	on ligand
cell-membrane ion channels	$\mathcal{S}_m\mathcal{B}_n$	sites on the protein	ion(s) and water
ionophore, *e.g.* crown ethers, polyethers, and polyureas	$\mathcal{S}_m\mathcal{B}_n$	what the ions bond to	the ions
metal-organic frameworks	$\mathcal{S}_m\mathcal{B}_n$	metal or organic framework component	guests
ionic liquid	$\mathcal{S}_m\mathcal{B}_n$	larger ions that form an effective local structure in the viscous fluid	smaller ions that move around in this structure
organic rearrangements, *e.g.* olefinic carbocations	$\mathcal{S}_m\mathcal{B}_n$	the parts that stay the same	the parts that rearrange
olefin metathesis	$\mathcal{S}_2\mathcal{B}_2$	diagonal C atoms from each C=C	other diagonal C atoms

A related type of example, focused upon in this work, is the $\mathcal{S}_4\mathcal{B}_2$ partitioning pertaining to tautomerism in free-base porphyrin[267-270, 273-282] (structures **2**, **3**, **6**, *etc.*, see Figure 6.3 and Figure 6.4). As outlined above, here the four $\mathcal{S}$ are the four nitrogen bonding sites on the macrocycle and the two $\mathcal{B}$ are the two inner hydrogen atoms, considered to be "mobile" amidst the external nitrogen environment. Simple variations of this basic motif including having bonders that are internally connected by, *e.g.*, replacing[142, 144-145, 148] the two

H atoms with the B(F)OB(F) group that presents two boron atoms as the mobile[7] bonders (Figure 6.2c). The strepsisomerisation "bond-walk" mechanism described in Chapter 4, Section 4.5.3 with animation in **E_File_7** is a direct application of the principles described here.

The $\mathcal{S}_m\mathcal{B}_n$ partitioning also applies to isomerism involving host-guest interactions. This includes substrate – enzymes and ligand-protein interactions where the definition of a "bond" may be required to include hydrogen bonding, van der Waals interactions and donor – acceptor interactions. Another example is that of either cation-bonding or anion-bonding ionophores. In each case, $\mathcal{S}$ constitute sites on the larger, bond-topologically static structure whilst $\mathcal{B}$ constitute the ion(s). Examples include crown ethers, polyethers, and anion-binding polyureas. A related scenario is the bonding of neutral molecules inside metal-organic frameworks (MOFs) in which the metals or organic-framework component provide the sites $\mathcal{S}$ and the host molecules are the mobile bonders $\mathcal{B}$. In all these host-guest systems, dissociation of the guest may be equally relevant as bound structures.

Ionic liquids also provide examples. These often comprise a high viscosity and, therefore, pseudo-stationary network of large anions or cations that form bonding sites, $\mathcal{S}$, to which more mobile cationic or anionic bonders, $\mathcal{B}$, bond, respectively. All possible constitutional isomers in each "unit cell" of the ionic liquid are therefore also specified by this Polytope Formalism of constitutional isomerism approach.

As a final example of the adaptability of the of the $\mathcal{S}_m\mathcal{B}_n$ partitioning approach, consider the constitutional isomerism relationship between Me–OCH$_2$–H and Me–CH$_2$O–H. By recognising that the inner CH$_2$–O group is comprised of bonders CH$_2$ and O, the sites $\mathcal{S}$ are Me and H and this is an example of $\mathcal{S}_2\mathcal{B}_2$. The $\mathcal{S}_m\mathcal{B}_n$ partitioning approach for the Polytope Formalism of constitutional isomerism can be applied to many analogous applications pertinent to modern industrial chemistry, *e.g.*, the intramolecular rearrangements of olefinic carbocations during petroleum cracking.

6.9   Conclusions.

The intrinsic Combinatorics and Graph Theory basis[29] of the Polytope Formalism of stereochemistry is adapted for application to configurational isomerism and provides a

unifying conceptual framework for the two main branches of isomerism. The formalism not only enumerates isomers (potential-energy local minimum structures) but also all possible concerted unimolecular bond topotropic rearrangements (R_{to}^c1), these being the constitutional isomerism analogue of R_{st}^c1 reactions as featured in the Polytope Formalism of stereoisomerism. The size of the problem scales superexponentially as $\mathcal{O}(2^{N^2})$. To provide tractability, many problems of practical interest to chemists can be recast by focussing only on atoms of interest through a scheme, introduced here, where the chemical system is partitioned as $\mathcal{S}_m\mathcal{B}_n$.

$\mathcal{S}_m\mathcal{B}_n$ provides a class descriptor and allows for simplification, expression of the core ideas in terms of simple chemical concepts, and the development of widely applicable conceptual and computational tools. Critically, this $\mathcal{S}_m\mathcal{B}_n$ approach retains all the R_{to}^c1 mechanistic information of the full approach. The resulting catalogue of bond-topology structures, associated graph, and assigned energies compactly encode the full PES of the active components of the chemical system. Software tools provided allow for the general application of the $\mathcal{S}_m\mathcal{B}_n$ partitioning to widely ranging chemical systems.

The worked example of inner-hydrogen tautomerism of free-base porphyrin, complying to the $\mathcal{S}_4\mathcal{B}_2$ class under the conditions of bonding site topology D_{4h} symmetry and non-dissociation, is explored in detail. This reveals 225 distinct bond-topology species grouped into 28 genera which are further grouped into 10 families. All generic structures were modelled and characterised by quantum chemical calculations. For a subset of structures described as distal structures the full PES was calculated by DFT using a traditional reaction coordinate scanning approach. The resulting surface which, due to periodicity of the coordinates, maps onto a torus with the location of the critical points (bond-topology species) and the allowed R_{to}^c1 reaction pathways between them are shown to form the same graph as that generated through the Combinatorics and Graph Theory approach of the formalism supporting the premise that PES and graph have a one-to-one mapping. The graph theoretic approach underlying the Polytope Formalism of constitutional isomerism thus provides an alternative approach for generating and compactly encoding a PES.

For the example of free-base porphyrin, a most interesting feature is that small substitutions to the porphyrin could lead to dramatic modulation of both the relative magnitudes and character of the potential-energy critical points. As a result, 12 of the 28

genera can be associated with the isomers of chemically perturbed porphyrins, implying also that many other genera will take on significant roles as transition states, *etc*. This shows that compounds can be made displaying isomers and isomerisation pathways that are very different to those of (common) close analogues. This stresses the importance of having available a complete catalogue of what types of structures might be feasible, and what $R_{to}^c 1$ mechanisms might be operative.

6.10   Methods

All geometry optimisations and vibrational normal mode analyses were performed using Gaussian-16[153]. All structures were characterised using vibrational frequency analyses with these being used within the harmonic-oscillator model to correct for Gibbs free energies.

The overall shape of the "distal" structure potential-energy surface (PES) studied was calculated using Density Functional Theory (DFT) at the B3LYP-D3(BJ)/6-31G(d) level,[154, 157, 283] in the gas phase, using an adaptive coordinate system defined through the Generalized Internal Coordinate (GIC) options for the "Opt" keyword in the Gaussian16 software package.[153] Details of the adaptive coordinate system are given in Appendix I.

All cross-bonded structures were manually performed at the B3LYP-D3(BJ)/6-31G(d) level,[154, 157, 283] in the gas phase.

Further geometry optimisation and vibrational mode analysis on all distal structures and select cross-bonded structure stationary points were carried out at the double-hybrid B2PLYP-D3(BJ)/cc-pVDZ level,[283-285] using implicit $CHCl_3$ solvation within the polarised continuum model[159] with standard parameters, to calculate free energies that are better representative of experimental conditions.

Free-base porphyrin analogues were generally optimised at the B3LYP-D3(BJ)/6-31G(d) level in the gas phase. For third and higher row elements, appropriate SDD pseudopotentials and basis sets were utilised.

7   A new design strategy for drug-target molecular conception based on internal modification of polyaromatic ligands

7.1   Summary

A new strategy based on "internal modification" of polyaromatic structures and henceforth given the name "pontation", involves the replacement of a ring-fusion aromatic C–C bond with a C–X–C linkage to give an aromatic bridged annulene, where X = O, NH, CH_2, CF_2, *etc.* The bridging group X can protrude from either face of the ring thus giving rise to two potential stereoisomers. Basic properties of pontated model and drug compounds with X = O (O-pontation) are calculated using Density-Functional Theory (DFT). A large scope for pontation is revealed through searches of the Chemical Abstract Service (CAS) and RCSB.org[30] databases with many medicinal chemistry relevant examples found.

Three examples of O-pontation applied to the existing drug compounds propranolol, the synthetic cannabinoid MDMB FUBINACA, and a quinazolinonepyrrolodihydropyrrolone Pim-1 kinase inhibitor are studied in detail with ligand-host docking calculations performed, these indicating improved binding in for two of the three examples as compared to the parent polyaromatic ligand.

General pharmacophore-relevant effects of pontation are explored by DFT calculations on a series of pontated naphthalenes, showing predictable and useful variations in properties including pronounced potential for van der Waals interactions. The pontation strategy can also be applied to non-aromatic compounds like azulene and isoindole, giving stabilised aromatic annulenes.

As pontation is a radically different type of structure modification to traditional approaches, it has important ramifications for drug patents.

226

7.2   Introduction

Understanding stereoisomerism is a critical aspect of medicinal chemistry, for both scaffold[286] and fragment-based[287-288] strategies. In Chapter 4, a new fundamental type of stereochemistry, called akamptisomerism, was identified. Identification of oxygen-bridged oxabicyclo[*m.n*.1] compounds as akamptisomers inspired a new principle for drug-target molecular design. This principle, based on "internal modification" of polyaromatic structures and named "pontation", involves the replacement of a ring-fusion aromatic C–C bond with a C–X–C linkage to give a bridged annulene. The bridging X group protrudes from one face of the resulting annulene and can be situated on either side giving, pairs of aromatic stereoisomers. Such stereoisomer pairs have similar steric properties as that of the parent structure, yet different chemical properties whilst retaining overall aromaticity. Importantly, the reduced symmetry of each stereoisomer has the potential for improved ligand-host interactions by exploiting the asymmetry inherent to most host binding sites. Herein, focus is mainly on X = O with resultant pontated stereoisomer pairs also being akamptisomeric, though X = NH, CH_2, CF_2 are equally viable simple variations on the design principle.

The potential of the pontation design principle for medicinal chemistry is explored through three polyaromatic drug examples considered in detail.   Drug-design software featuring ligand docking-scores and binding free-energy calculations are used to rate the parent drugs against their pontated variants for host binding.   These predict improved binding in two of the three examples. This principle can be applied to millions of known molecules, including 34% of the ligands of the current ligand-host complexes in the RCSB.org database. In this sense, pontation can be considered as a "fine-tuning" design concept that could be applied to known polyaromatic drug candidates seeking improved properties.

To evaluate the pontation design principle in general terms, DFT calculations on a range of model systems is used to demonstrate energy and property relationships between both the pontated variants and their parent system. The effect of pontation on some pharmacologically relevant properties are shown to be favourable. Further, pontation of the non-aromatic azulene and isoindole systems is shown to lead to stabilised aromatic variants opening up new drug-design prospects for structural motifs that are otherwise problematic.

This design principle also has important implications for drug patents. As pontation is unrelated to other chemical substitution strategies typically covered in drug patents, it promises scope for patent extensions and renewals.

7.3   Pontation as a design strategy for generating new aromatic molecules

The inspiration for this work comes from akamptisomeric molecules like 1,6-epoxy[10]annulene **2**[183-188] for which derivatives include **4** and **5** and others that have been patented as β-androgenic blockers,[289-290] and the closely related diastereoisomers **7** and **8** that are model compounds[191] for graphene oxide. These molecules can be conceptually envisaged as variants of the parent compounds **1**, **3**, and **6**, respectively, through an operation that is called "pontation", meaning "bridged", and is represented in Figure 7.1 by a special arrow symbol. This is the first introduction of this arrow into the chemical literature.

The pontation design strategy involves replacement of a C–C bond that fuses two aromatic rings with a bent C–X–C linkage and is referred to as X-pontation. Most of the worked examples in this Chapter have X = O and are thus O-pontated. As oxygen is a bi-coordinate centre, these O-pontated molecules are also akamptisomeric.

An essential feature of pontation is that, in general, the X group can be either above or below the plane of the parent molecule, hence generating two stereoisomers. For the pontation of naphthalene **1**, the two molecules so designed are identical, and are labelled **2** and **2'** instead of being assigned a different number. For the other two examples, the pontated molecules are distinct. Specifically, **4** and **5** are enantiomeric akamptisomers, whereas **7** and **8** are diastereomeric akamptisomers.

In terms of the chemistry, it is noted that the parent compounds **1**, **3**, **6**, and **9** are all polyaromatic molecules with their pontated forms being annulenes with $4n + 2$ π electrons, stressing the key feature that pontation preserves aromaticity. Hence, as a design concept, pontated systems resemble their parent molecules. Notwithstanding, perturbations to the chemistry will result and depend upon both the nature of the parent system and the pontating X group. This may be advantageous or otherwise depending upon the intended application. For example, O-pontation of naphthalene **1** gives the mildly acid labile[192] **2** whereas CH_2-pontation gives an analogue that is far stabler, both thermally and to acid attack.

(a)

naphthalene   pontated analogue

$\Delta G = 0$

1   with X = O: **2**   **2'**

side view

top view

(b)

1-naphthoic acid   pontated enantiomeric akamptisomers

COOH   COOH   COOH

$\Delta G = 0$

3   **4**   **5**

(c)

anthracene   doubly pontated diastereomeric akamptisomers

$\Delta G = 57$

6   **7**   **8**

(d)   β-naphthylalanine   pontated diastereomeric akamptisomers

HOOC NH$_2$ HOOC NH$_2$ HOOC NH$_2$

$\Delta G = 0$

9   **10**   **11**
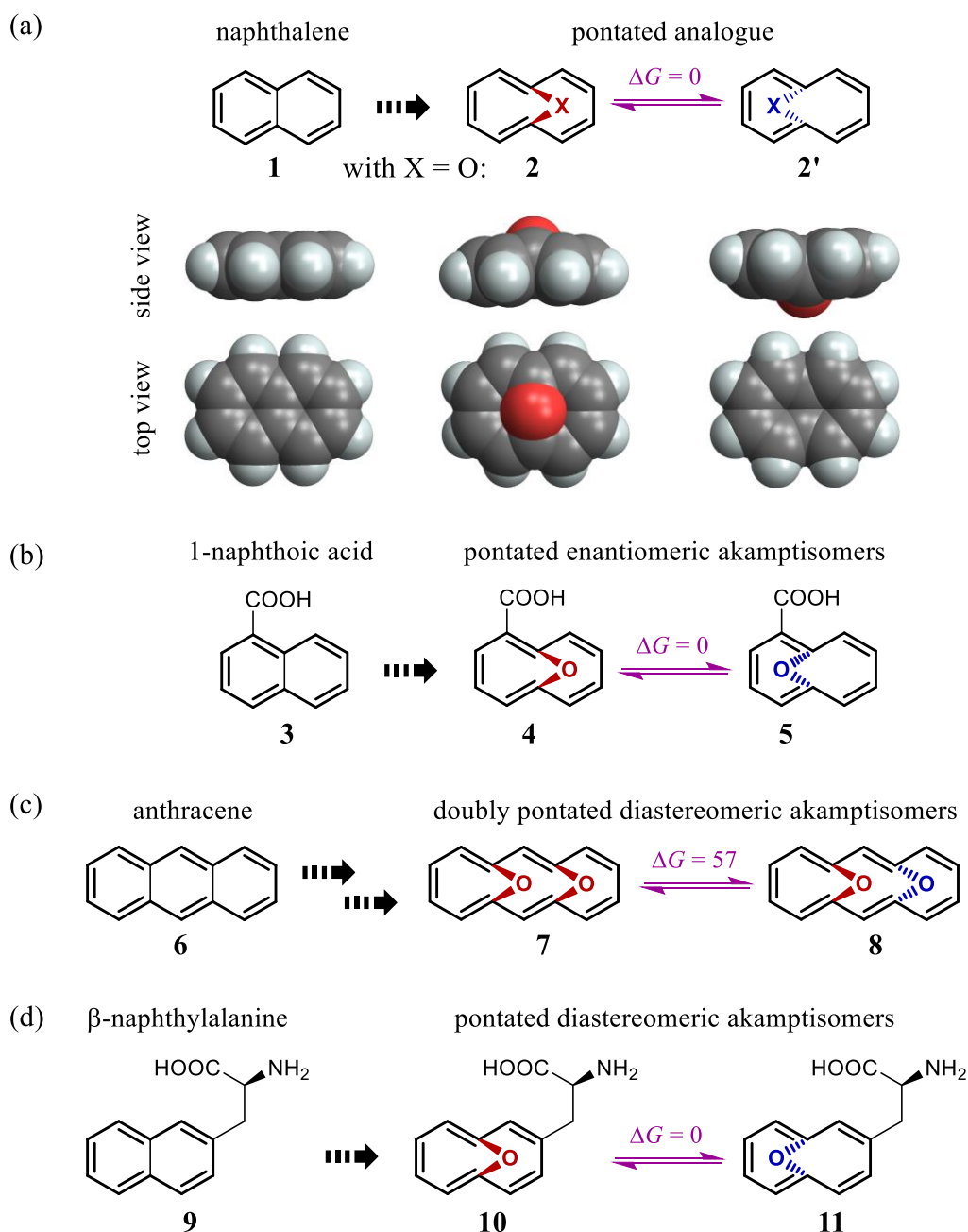
Figure 7.1 Demonstration of the effect of pontation on sample polyaromatic molecules. The pontation arrow is introduced here for the first time to the chemistry literature. (a) Pontation by X (X-pontation) on naphthalene to give bridged [10]annulenes. For X = O these are akamptisomeric analogues **2** and **2′**. (b) O-pontation on **3** gives the enantiomeric akamptisomers **4** and **5.** (c) Double O-pontation on anthracene **6** can produce diastereomeric akamptisomers, two of which are shown. (d) The chiral naphthylalanine **9**, upon O-pontation gives a pair of diastereomeric akamptisomers **10** and **11**. Calculated free-energy differences are indicated, in kJ mol^{-1}.

Pontation produces aromatic annulenes that are no longer intrinsically symmetric about the aromatic plane with the X group protruding from one face of the annulene. This is

highlighted in Figure 7.1 using 3D renderings of density-functional theory (DFT) models for naphthalene **2** and its O-pontated variants **2** and **2'**. As these images show, the protrusion of the pontated oxygen atom out of the molecular plane is small, as is the induced deviation from planarity. The expectation that the binding of a pontated ligand to the same site in a host as the original ligand is therefore a reasonable one. Nevertheless, the structural and electronic perturbations brought about through pontation could significantly alter intermolecular interactions that determine drug-host binding. Given that the ligand binding sites in most host molecules are intrinsically assymetric, one would expect the two pontated stereoisomers to differentially interact with the binding site, this manifesting as *selectivity* for one stereoisomer over the other. From a drug-design standpoint, pontation thus offers a means of fine-tuning existing polyaromatic drugs or drug candidates.

Previous strategies for fine-tuning drug properties have largely relied upon ring-atom substitution or ring-periphery substitution. Pontation contrasts and complements those strategies by acting as a type of 'internal modification' where the degree of property changes can range from small to arbitrarily large depending upon the nature of X. For medicinal chemistry, the pontation design concept provides a functionally dense modification alternative or addition to traditional approaches.

7.4   Basic properties of pontated polyaromatic compounds

In focussing on X = O, important questions pertaining to a pair of pontated systems concerns the relative energies of each isomer and the rate at which they are likely to interconvert at room temperature. To address these questions, DFT calculations are performed (see Methods Section 7.12). The calculated Gibbs free energies and rate constants are listed in Table 7.1 and on Figure 7.1. Small energy differences are predicted between the O-pontated stereoisomer pairs as in all cases the features that differentiate the isomers concern only long-range interactions, meaning that both isomers are likely to be of similar synthetic feasibility. Also, the barriers for the concerted thermal unimolecular isomerisation (akamptisomerisation by and $R_{st}^c 1$ mechanism – see Sections 2.4.2 and 2.4.20) are high in these bicyclo[4.4.1] compounds indicating they are thermally stable with respect to this mechanism.

Table 7.1 Properties of akamptisomeric O-pontated compounds and $R_{st}^c 1$ akamptisomerisation reactions calculated using B3LYP-D3(BJ)/6-31G*.[a]

Parent polyaromatic molecules[b]	Ring system[b]	Pair	Solvent[c]	ΔG / kJ mol^{-1}	θ_{ts} / °	$\Delta G^{\ddagger}$ / kJ mol^{-1}	τ_{298}
naphthalene	[4.4.1]	2, 2′	CHCl$_3$	[0]	180	222	10^{19} y
1-naphthoic acid	[4.4.1]	**4, 5**	CHCl$_3$	[0]	180	222	10^{19} y
anthracene	[4.4.1]	**6, 7**	CHCl$_3$	57	156	80, 24[d]	2 s, 1 ns [d]
2-naphthylalanine	[4.4.1]	10, 11	water	[0]	177	222	10^{19} y
propranolol	[4.4.1]	13, 14	water	[0]	173	218	10^{19} y
MDMB-FUBINACA	[4.3.1]	16, 17	water	8	148	352	10^{30} y
quinazolinone inhibitor	[4.4.1]	19, 20	water	−17	169	268	10^{25} y

a: ΔG is the free energy change for conversion of the first-listed molecule to the second-listed one; $\Delta G^{\ddagger}$ is the activation energy for the concerted unimolecular thermal akamptisomerisation process that could interconvert the isomers, with θ_{ts} being the bond angle around the key oxygen atom at the transition state, and τ_{298} the estimated thermal lifetime of the high-energy energy isomer at 298 K.

b: Molecules from which the akamptisomers can be envisaged, either by pontation or other chemical modifications, and the most apt bicyclo[*m.n.*1] ring system that depicts them.

c: Calculations performed using the PCM dielectric continuum model to describe the solvent.

d: First number is for the (αα)-isomer → (αβ)-isomer. The second is the reverse reaction.

7.5   Scope for pontation

In the context of medicinal chemistry, pontation represents a fundamentally new chemical design stategy. As such, it has not been exploited in existing patent specifications, except for just two patents involving non-akamptisomeric methano-bridged derivatives.[289-290] Nevertheless, pontation can, in principle, be applied to a vast number of polyaromatic compounds. As at March 2019, the Chemical Abstracts Service (CAS) database lists 2,445,953 referenced naphthalene derivatives alone to which pontation could be applied. This process would therefore result in the designing of 4,891,906 new molecules per X group. These pontated molecules could have similar chemical properties to the original but different intermolecular interactions, as controlled by the nature of X. The listed molecules include naphthylalanine derivatives, of which 28,860 are flagged in the databases as being "biologically relevant". Compound **9** (2-naphthylalanine), is an unnatural amino acid[291] of importance in, *e.g.*, in the *-relix* series of GnRHR antagonists. O-pontation leads to diastereomeric akamptisomers **10** and **11**, immediately presenting new synthetic design options.

More broadly, to focus on the likely relevance of this to drug design, it is noted that as of March 2019, 34% of the 110,701 ligand-host structures listed in the RCSB.org database,

feature a polyaromatic group. Many, though not all of these ring systems will be suitable for pontation. This is discussed further in Section 7.9.

7.6   Examples of O-pontation applied to existing drug compounds

To quantitatively investigate the effect of pontation on ligand binding, three example systems are considered with ligand-target docking calculations performed. These are (Figure 7.2): propranolol[292] **12** binding to the turkey ß1 adrenergic receptor, utilising a structure with bound antagonist cyanopindolol[293] (2YCY)[294], MDMB-FUBINACA **15** docked into the CB1 cannabinoid receptor 1-G protein[295], utilising[294] 6N4B, and the quinazolinonepyrrolodihydro-pyrrolone **18**, an inhibitor of human Pim-1 kinase[296], as reported in structure[294] 6MT0. Details of the target[297-300] and ligand preparation,[301-302] receptor grid preparation,[303] docking-score calculation,[303] binding free-energy calculation[297, 300, 303-304] and ligand interaction diagram generation[298] are described in detail in the Appendix H. This work was done in collaboration with Prof David Hibbs and Dr Jonathon Du of The School of Pharmacy at The University of Sydney. Results are summarised in Table 7.2 and Figure 7.2, Figure 7.3, and Figure 7.4.

Table 7.2   Docking simulations[d] of ligands into protein targets and their O-pontated analogues envisaged by pontation.

protein site	ligand	variant		docking score[a]	ΔG[b] / kJ mol^{-1}
2YCY Site A	native	-		−8.94	−291.2
	propranolol	original	**12**	−9.39	−282.0
		O-pontated	**13**[c]	−8.49	−293.3
		O-pontated	**14**	−8.95	−284.5
2YCY Site B	native	-		−9.02	−307.5
	propranolol	original	**12**	−8.77	−288.7
		O-pontated	**13**	−8.90	−284.5
		O-pontated	**14**	−8.95	−293.7
6N4B	MDMB-FUBINACA	original	**15**	−10.92	−349.4
		O-pontated	**16**	−10.75	−310.5
		O-pontated	**17**	−11.01	−342.3
6MT0	quinazolinone	original	**18**	−7.64	−277.4
		O-pontated	**19**	−4.97	−195.4
		O-pontated	**20**	−4.92	−195.8

a: From Glide V8.1;[303] more-negative scores imply better docking.
b: From PrimeMM-GBSA calculations.[304]
c: Lowest energy of all configurations generated, but the 3rd best docking score (others −8.81 and −8.63).
d: This work was done in collaboration with Prof David Hibbs and Dr Jonathon Du of The School of Pharmacy at The University of Sydney.

7.6.1   *Propranolol and its two pontated variants in 2YCY*

The first ligand-host interaction considered is the β-blocker propranolol binding to two binding sites of the turkey ß1 adrenergic receptor[293], varying structure 2YCY[294]. The native ligand, cyanopindolol, can bind to two similar protein sites, Site A and Site B (see Figure 7.2), with docking scores of −8.9 and −9.0, respectively, as well as calculated binding free energies of −291.2 and −307.5 kJ mol^{-1} (Table 7.2). As cyanopindolol is a complex heteroaromatic ligand with multiple types of protein interactions, also considered is the pontation-amenable ligand, propranolol **12**, instead. Its calculated binding scores to Sites A and B are similar, −9.4 and −8.8, as are its calculated binding free energies −282.0 and −288.7 kJ mol^{-1}, respectively. Its O-pontated variants **13** and **14** are predicted to have slightly poorer docking scores for Site A and slightly better scores for Site B, but better binding energies for Site A and one poorer and one better binding energy for Site B. Overall, the calculated changes in docking score and binding energy are sufficiently small to indicate the overall similarity of **12** – **14**, but, quantitatively, the predicted enhancement of ΔG for **13** over **14** for Site A of 11.3 kJ mol^{-1} implies that the O-pontated variants will show hundred-fold increased binding compared to the original ligand **12**. Hence, this result supports the premise of pontated derivatives displaying enhanced properties.
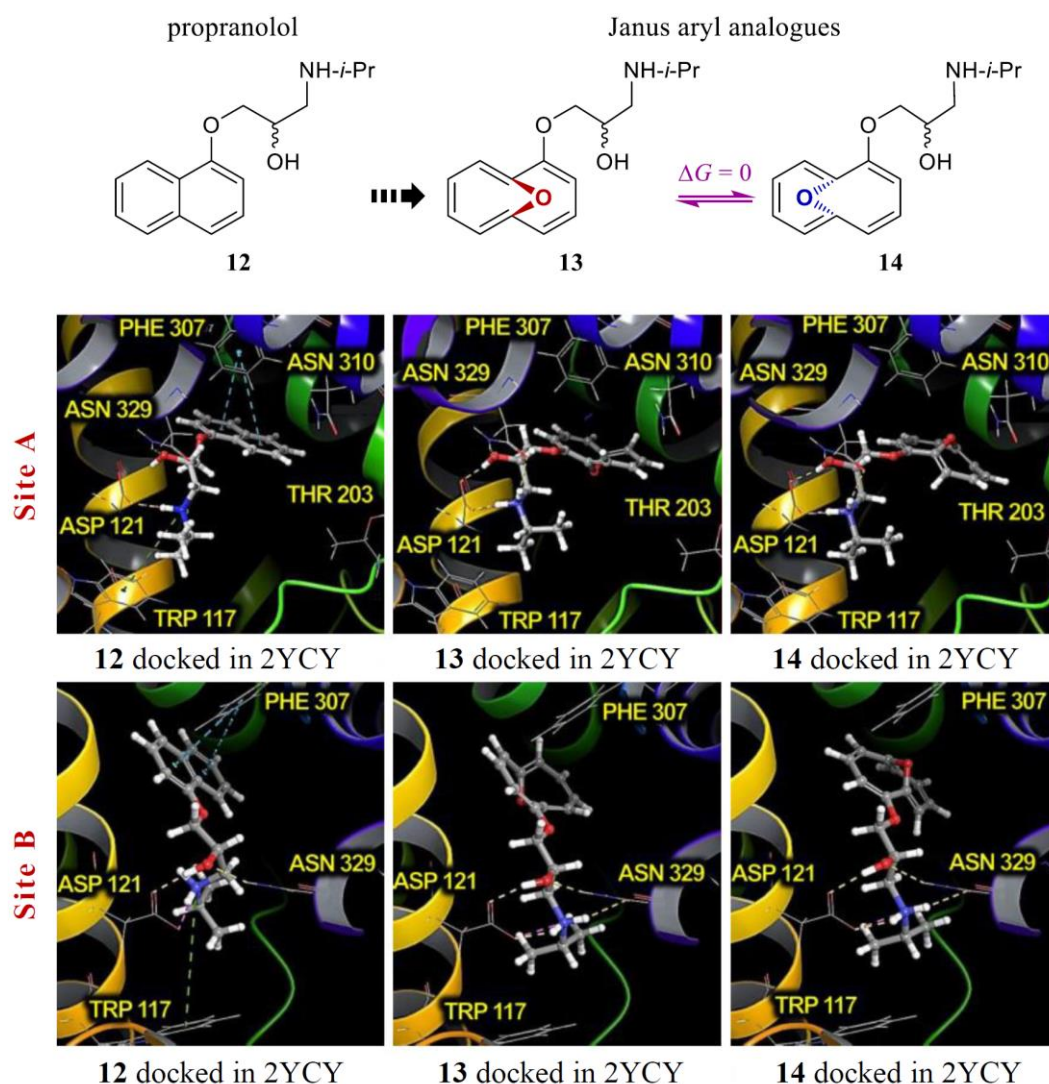
Figure 7.2 β1-adrenergic antagonist, propranolol **12**, and its O-pontated akamptisomeric variants **13** and **14** docked into Site A and Site B of the adrenergic receptor in 2YCY. See Table 7.2 for calculated binding energies. Calculated free-energy difference Δ*G* in water is indicated, in kJ mol^{-1}, see Table 7.1. Atoms and bonds marked in red are above the quasi-plane of the pontated ring system whilst those in blue are below it. This work was done in collaboration with Prof David Hibbs and Dr Jonathon Du of The School of Pharmacy at The University of Sydney

7.6.2 *MDMB-FUBINACA and its two pontated variants in 6N4B*

Results for the ligand-host interactions of MDMB-FUBINACA **15**, and its O-pontated variants **16** and **17**, docked into cannabinoid receptor 1-G protein[295], utilising structure[294] 6N4B show a similar pattern to the previous example with one variant (**17**) calculated to bind more strongly than the other (**16**). This can be rationalised as arising from adverse steric

234

interactions for **16** that disrupt the van der Waals interactions and an impaired lipophilic interaction.

Both the docking score and calculated binding free energy for variant **16** is slightly less favourable than for the original **15**. For variant **17**, the docking score is significantly better than the original **15** but its calculated binding free energy is slightly less.
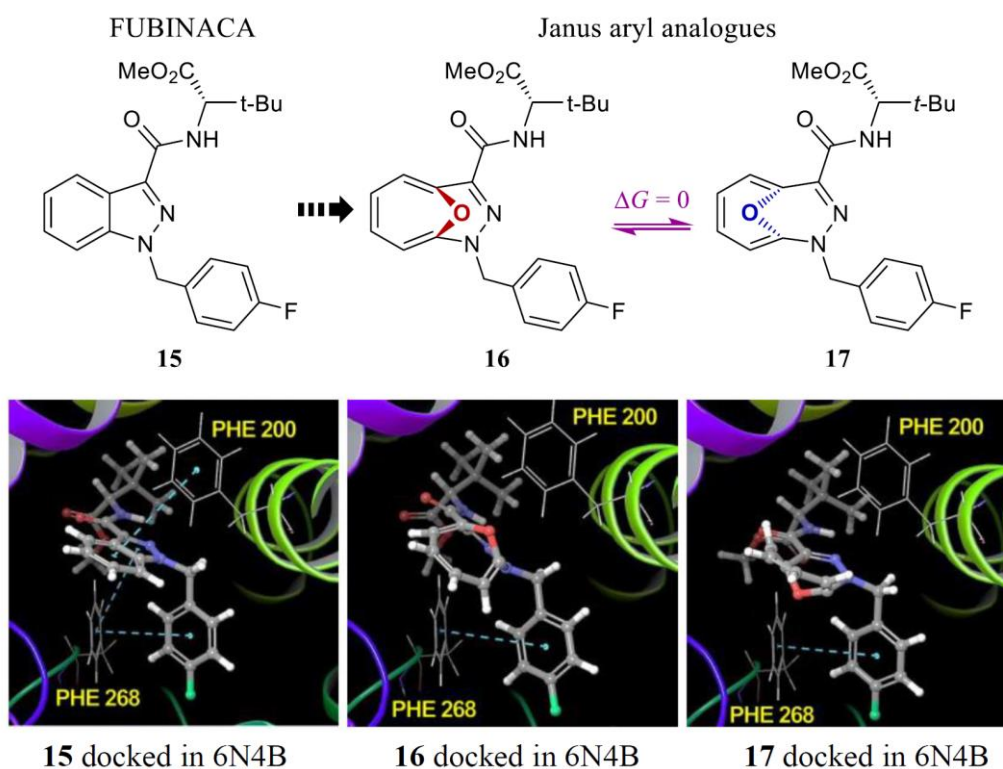
Figure 7.3 Cannabinoid agonist, MDMB-FUBINACA **15**, and its O-pontated akamptisomeric variants **16** and **17** docked into the binding site of the CB1 receptor in 6N4B. See Table 7.2 for calculated binding energies. Calculated free-energy differences ΔG in water are indicated, in kJ mol^{-1}, see Table 7.1. Atoms and bonds marked in red are above the quasi-plane of the pontated ring system whilst those in blue are below it. This work was done in collaboration with Prof David Hibbs and Dr Jonathon Du of The School of Pharmacy at The University of Sydney

7.6.3 A quinazolinonepyrrolodihydropyrrolone Pim-1 kinase inhibitor and its two pontated variants in 6MT0

A different scenario is presented by the final ligand-host interaction, a quinazolinonepyrrolodihydropyrrolone inhibitor **18**, binding to human Pim-1 kinase[296], as reported in structure[294] 6MT0. All methods predict that the two O-pontated variants, **23** and **20,** bind poorly compared to the native ligand, thus providing a counterexample.
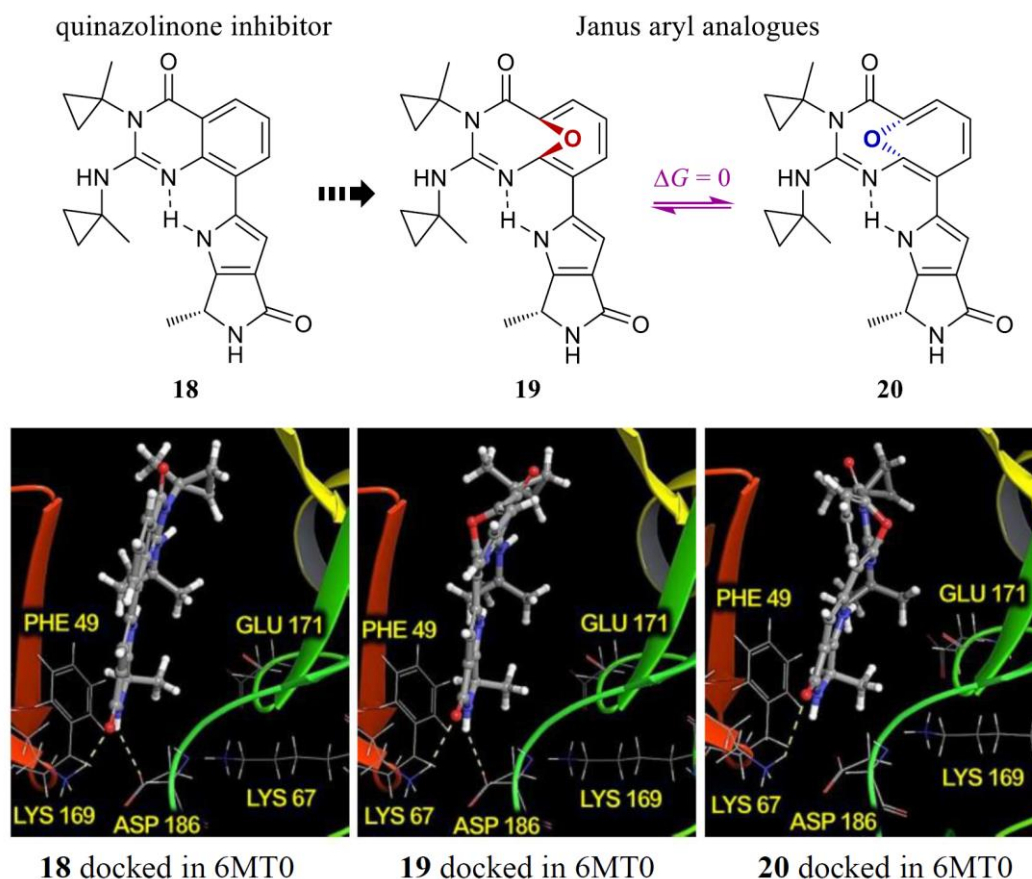
Figure 7.4   The quinazolinonepyrrolodihydropyrrolone Pim-1 kinase inhibitor **18**, and its O-pontated akamptisomeric variants **19** and **20** docked into Pim-1 from 6MT0. See Table 7.2 for calculated binding energies. Calculated free-energy difference Δ*G* in water is indicated, in kJ mol^{-1}, see Table 7.1. Atoms and bonds marked in red are above the quasi-plane of the pontated ring system whilst those in blue are below it. This work was done in collaboration with Prof David Hibbs and Dr Jonathon Du of The School of Pharmacy at The University of Sydney

7.7   General pharmacophore-relevant effects of pontation to existing drug compounds

Pontation alters both geometry and electronic structure as compared to the parent compound. Both effects will have consequences for ligand-host interactions and chemical stability, both from the standpoint of synthesis and *in vivo*. To investigate these effects, DFT calculations were performed for the series of compounds, shown in Figure 7.5, produced by X-pontation of naphthalene, where X = O, NH, CH$_2$, and CF$_2$. Full details are found in the Methods Section 7.12. From the space-filling renderings of the compounds in Figure 7.5, it is clear that the shape changes increase with the increasing size of X. Understandably, the

most dramatic changes are present on the side of the ring where X group protrudes. Additionally, the pitch of the C–H bonds relative to the parent **1**, also exhibits an increasing perturbation with consequences for would-be peripheral substituents to the outer ring.
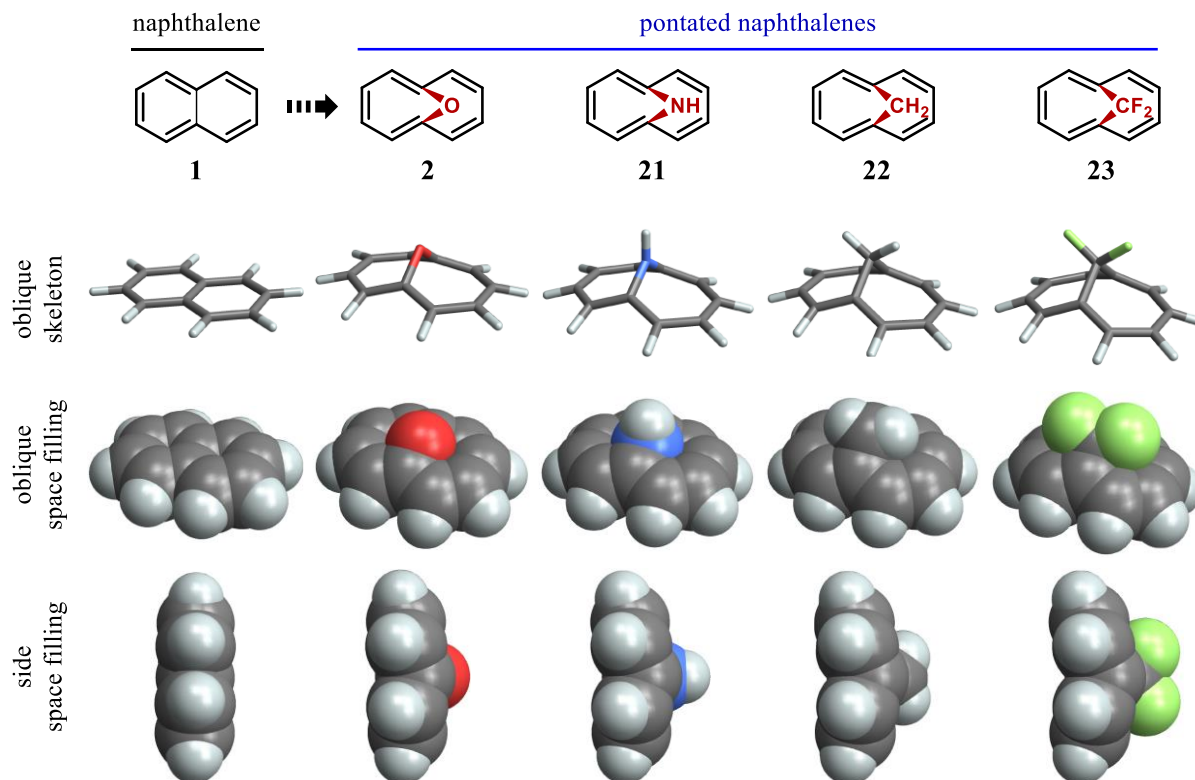
Figure 7.5 Naphthalene **5** and its pontated variants **2**, **21**, **22**, and **23** with X = O, NH, CH$_2$, and CF$_2$, respectively. Skeleton and space-filling renderings of DFT models show the shape relationships. The perturbation to the shape increases from left to right with O-bridge producing the least and the CF$_2$-bridge the most. Atoms and bonds marked in red are above the quasi-plane of the pontated system.

For the series **2**, **21** – **23** in Figure 7.5, the effects of pontation are quantitatively examined by DFT modelling these and the reference compounds **1** and **24** with the results summarised in Table 7.3.
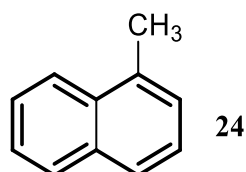
Table 7.3 Modelled surface areas^{a,c}, volumes^{a,c}, highest occupied molecular orbital (HOMO) levels, and trace of the polarisabilities^{c} α for two reference compounds **1** and **24** (shaded rows), and pontated naphthalenes **2**, **21** – **23**.

Compound	Bridge	Surface area^{a} / Å^2	Volume^{a} / Å^3	HOMO^{b} / eV	Tr[α]^{b} / Å^3
1	none	190 [1]^c	185 [1]^c	−6.09	214 [1]^c
24	none	207 (1.09)^c	206 (1.11)^c	−5.97	306 (1.43)^c
2	O	197 (1.04)^c	195 (1.05)^c	−6.13	312 (1.46)^c
21	NH	202 (1.06)^c	200 (1.08)^c	−5.91	315 (1.46)^c
22	CH$_2$	203 (1.07)^c	205 (1.11)^c	−5.94	313 (1.46)^c
23	CF$_2$	211 (1.11)^c	215 (1.16)^c	−6.22	308 (1.44)^c

a: Calculated from the solvent accessible surface using a 0.1Å solvent sphere radius on B3LYP-D3(BJ)/cc-pVTZ gas-phase geometries. *b*: B3LYP-D3(BJ)/cc-pVTZ gas-phase. *c*: Proportion relative to that for **1**.

The calculated surface areas and volumes of the pontated compounds are very similar to the reference parent compound **1** demonstrating the degree of their isosteric character with X = O producing the least surface area and volume changes, and X = CF$_2$ the most.

A major route for catabolism of exogenous compounds *in vivo* is oxidative biotransformation by cytochrome P450 isozymes.[305] The highest occupied molecular orbital (HOMO) level is a direct measure of general oxidisability of compounds. The DFT calculated HOMO levels indicate that **2** and **23** are less oxidisable than the parent naphthalene **1**, with **22** roughly equivalent to the reference compound **24**. That **2** is *less* oxidisable than naphthalene contrasts with the effect typical of oxygen substituents at the periphery of aromatic systems. This can be rationalised as the lone pairs of the oxygen bridge being unable to effectively conjugate with the π-system and the high electronegativity of oxygen inductively lowering the π molecular orbitals. The effect of X on the HOMO levels displays a clear trend indicating that inductive effects dominate.

The polarisability of compounds is a measure of their capacity for van der Waals interaction.[306] The trace of the DFT polarisabilities α are listed and compared to the reference compounds **1** and **24**. These polarisabilities for the pontated compounds are largely the same as for **24** but substantially larger than that for the parent naphthalene **1**. Pontation is thus predicted to significantly enhance van der Waals interactions with consequences for improved drug binding, all other things being equal.

An additional way of measuring pontated molecules' capacity for intramolecular van der Waals interaction is through the modelling of two component systems. DFT calculations of the naphthalene homodimer, naphthalene – pontated-naphthalene heterodimers, and

pontated-naphthalene homodimers were carried out with the results shown in Figure 7.6. Full details are in the Methods Section 7.12.
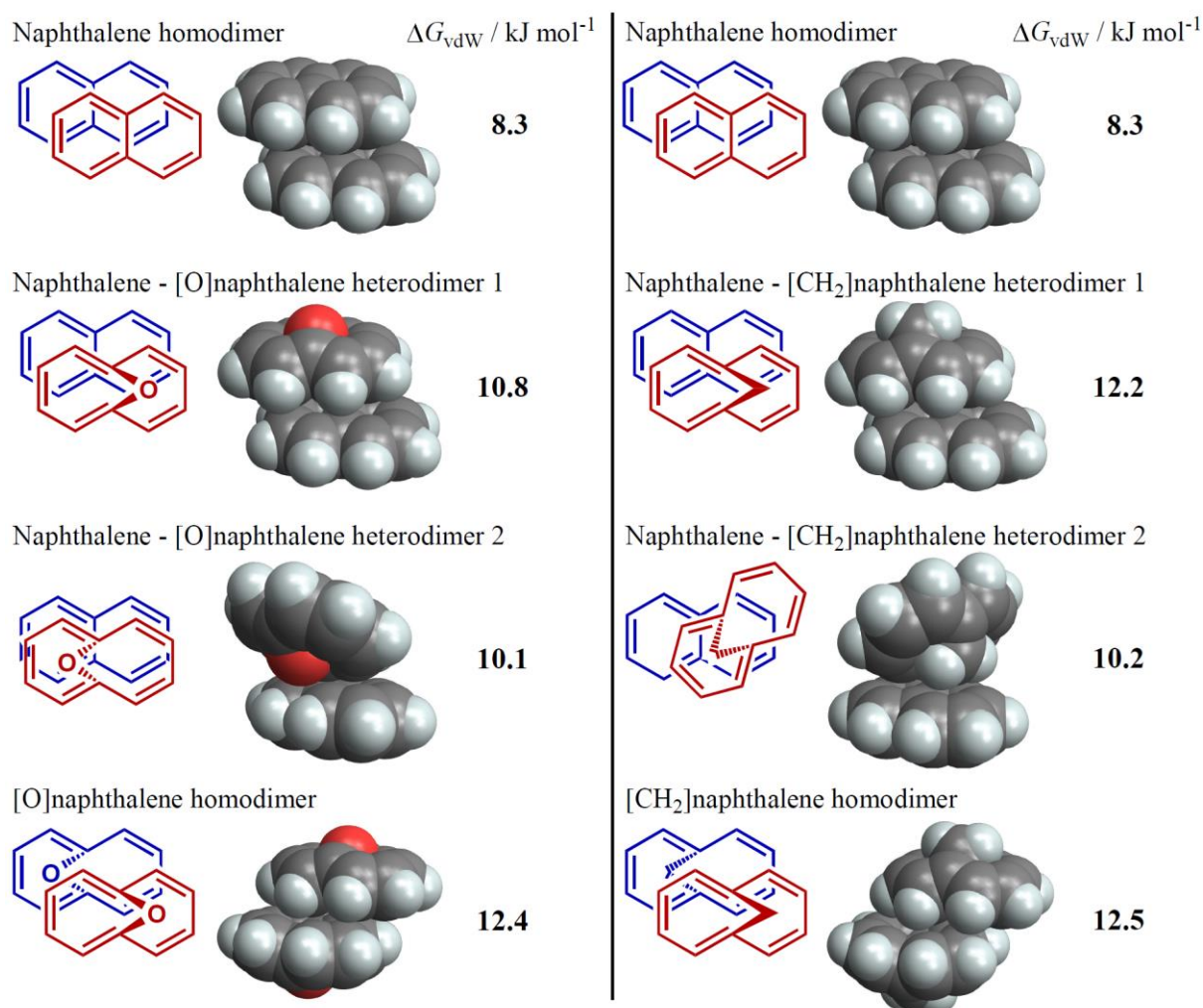
Figure 7.6 DFT calculated intramolecular van der Waals energies for naphthalene, O-pontated naphthalene, and CH$_2$-pontated naphthalene homo- and heterodimers. Energies were calculated at the B3LYP-D3(BJ)/cc-pVTZ level. The pontated molecules show enhanced intramolecular interactions relative to the parent naphthalene.

For each dimer modelled, the extra van der Waals intramolecular interaction arising from a single pontation equates to approximately 2 kJ mol^{-1} and is largely additive although orientation effects are also evident. This finding is consistent with the polarisability results and supports the expectation for improved ligand-host binding of pontated drug molecules, all other things equal.

7.8 Application of pontation to other conjugated systems

As noted, the primary consequence of pontation is the formation on an annulene. This is typically a [10]annulene where the bridge group X helps to maintain effective p-orbital overlap allowing the 10e π-system to be aromatic. A significant consequence of this is that the aromatic pathway has been isolated to the outer-ring periphery obviating any cross-ring interactions that arise where there would otherwise be a direct bridgehead C–C bond.

This property can be utilised to convert non-aromatic compounds like azulene **25**, shown in Figure 7.7, into the bridged aromatic [10]annulene **26**. Whilst minimally perturbing the shape of the molecule, the change to the electronic structure is profound with significant stabilisation gained.

In a similar way, the highly reactive isoindole **27**, upon pontation is predicted to become the more stable **28**. Currently, there are no isoindole-based drugs with pontation potentially offering a way to access stable isoindole-like pharmacophores.
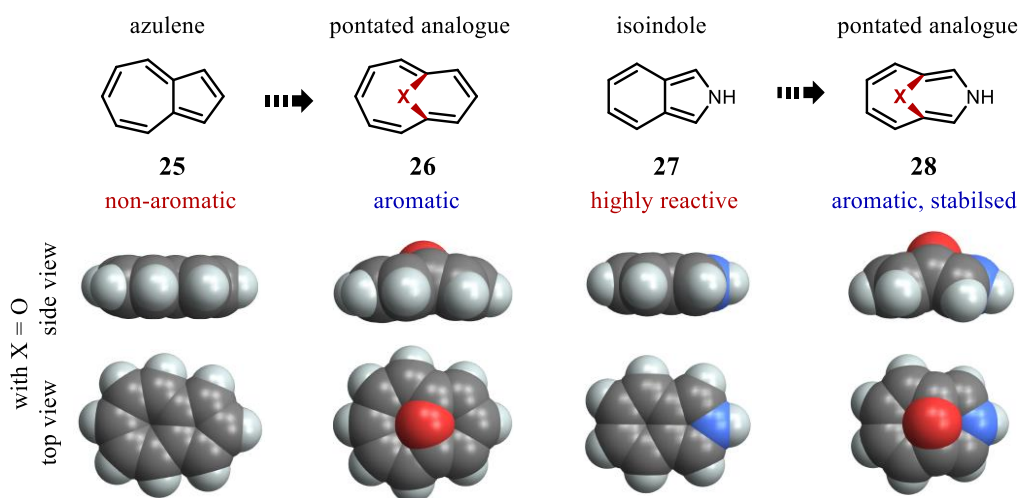
Figure 7.7 Predicted effects of pontation of azulene **25** and isoindole **26** to give **27** and **28**, respectively. Azulene is non-aromatic but the pontated form, shown here with X = O, is expected to be aromatic and akamptisomeric. Isoindole is a highly reactive species that is predicted to become stabilised and aromatic upon pontation (shown here with X = O giving akamptisomers). Atoms and bonds marked in red are above the quasi-plane of the pontated system.

With X = O for **26** and **28**, there is also the possibility of R_{st}^c1 akamptisomerisation. These are shown in Table 7.4 with those for O-pontated naphthalene **2**, given for reference. The most striking result is that the bicyclo[5.3.1] structure of **26** is predicted to be fluxional. This can be seen as a consequence of the larger side of the **26** providing more space for the oxygen atoms to swing to the other face of the molecule.

Table 7.4 DFT calculated properties of akamptisomers and $R^c_{st}1$ akamptisomerisation reactions calculated using B3LYP-D3(BJ)/6-31G*.[a]

Parent polyaromatic molecules[b]	Ring system[b]	Pair	Solvent[c]	ΔG / kJ mol^{-1}	θ_{ts} / °	$\Delta G^{\ddagger}$ / kJ mol^{-1}	τ_{298}
naphthalene	[4.4.1]	**2, 2′**	CHCl$_3$	[0]	180	222	10^{19} y
azulene	[5.3.1]	26, 26′	CHCl$_3$	[0]	135	63	2 ms
isoindole	[4.3.1]	28, 28′	CHCl$_3$	[0]	147	300	∞

a: ΔG is the free energy change for conversion of the first-listed molecule to the second-listed one; $\Delta G^{\ddagger}$ is the activation energy for the concerted unimolecular thermal akamptisomerisation process that could interconvert the isomers, with θ_{ts} being the bond angle around the key oxygen atom at the transition state, and τ_{298} the estimated thermal lifetime of the high-energy energy isomer at 298 K.

b: Molecules from which the akamptisomers can be envisaged, either by pontation or other chemical modifications, and the most apt bicyclo[*m.n*.1] ring system that depicts them.

c: Calculations performed using the PCM dielectric continuum model to describe the solvent.

7.9   Discussion

Whilst pontation can, in principle, be applied to any polyaromatic system, in practice, certain systems are more amenable when factors such as ease of synthesis and stability against acid or nucleophilic attack are considered.

A primary consideration is the resulting strain that arises in pontated molecules. The principal factor for this is the size of the resultant bicyclo system. Of the systems examined in this Chapter, the bicyclo[5.3.1] compound **26**, formally derived from azulene **25**, is the least strained and allows for better p-orbital overlap in the π-system. The numerous bicyclo[4.4.1] compounds are also reasonably accommodating of the bridging group and, indeed, many compounds conforming to this type were synthesised by Vogel, *et al.*[184, 186, 190, 307-314] and provide much of the early literature regarding synthesis and chemical stability. Smaller bicyclo[4.3.1] systems do, however, exhibit pronounced strain with sub-optimal p-orbital overlap about the bridgehead atoms.

A second consideration is the susceptibility of peripheral ring atoms to nucleophilic attack as this could lead to undesirable hydrolysis *in vivo*. Electronegative elements adjacent to the bridgehead positions are expected to be undesirable as the polarisation of a bridgehead–heteroatom bond would be compounded by the strain at the bridgehead position.

Of the polyaromatic ligands present in complexes in the RCSB.org database, a large proportion of these are indole derivatives. Pontation of indoles is expected to lead to particularly unstable compounds as it embodies each of these undesirable factors.

The nature of the pontating group is also highly significant. Whilst oxygen is sterically the easiest to accommodate and has a stabilising effect on the molecular orbital, the O-pontated compound **2** is, nevertheless, mildly acid labile, rearranging to 1-naphthol. The CH$_2$ analogue **22**, is far more resistant to attack by acid despite its frontier molecular orbitals being higher in energy as compared to both **2** and naphthalene **1**. The CF$_2$ analogue **23**, whilst having the stability of **22** and the orbital stabilisation of **2** does exhibit the largest structural perturbation from planarity of the pontated naphthalenes examined here. When implementing a pontation drug-design strategy, each of these effects will need to be considered.

7.10 Conclusions

In summary, a new concept for drug design is presented based on the principle of pontation of fused-ring bonds in aromatic molecules, making new isomeric pairs with predictable shape and chemical-property changes that could be exploited in medicinal chemistry. Three examples pertinent to modern drug design were considered, with ligand-host docking calculations predicting that improved binding should occur in two cases. This highlights the potential of pontation as a tool for drug discovery. It presents quite a different strategy to drug discovery approaches that have been widely utilised to date, concentrating on internal modification of ligand polyaromatic ring structure rather than either ring-atom substitution or peripheral-atom substitution. This chemical design process can be applied to millions of known compounds, many of which have biological or medicinal relevance.

DFT modelling of various pontated naphthalenes reveals that relative to the parent naphthalene, there are predictable and useful of changes of chemical properties that are important for medicinal chemistry

As a design strategy, pontation provides a radically different, yet complementary, approach to traditional strategies. An additional consequence of this approach is that existing patents do not cover this type of drug variation. As such, pontation also presents opportunities for patent renewal and patent extensions.

7.11   Afterword – proof-of-principle has been established

To test the hypothesis expounded above, Crossley and Reekie have synthesised enantiomeric CH_2-pontated variants of a potent synthetic cannabinoid agonist. The resolved enantiomers both showed tighter binding to both CB1 and CB2 isoforms as compared to the parent polyaromatic compound. Importantly, the different enantiomers showed reverse selectivity for the two receptors. Further, *in vitro* absorption, distribution, metabolism, and excretion tests showed a ten-fold increase in solubility with other measures being similar to those of the parent compound. Structures and further details of the testing results are not provided here because of patent sensitivity. This work provides in-principle support for the medicinal chemistry design concept expounded in this Chapter.

7.12   Methods

All geometry optimisations, vibrational normal mode analyses, and molecular size measurements were carried out using Gaussian-16[153]. All structures were characterised using vibrational frequency analyses. All space-filling models use each element's standard van der Waals radius.

Gibbs free energies for O-pontated compounds local minima and transition-state structures were calculated by applying the B3LYP density functional[154] and the 6-31G* basis set,[157] with dispersion interactions treated empirically using Grimme's D3(BJ) method,[283] and implicit solvation corrections using the polarised continuum model (PCM).[159] Gibbs free energies were corrected for using the harmonic-oscillator model. Maximum errors[156] in this approach are of the order of 20 kJ mol^{-1}, much smaller than the range of results considered. The calculations are thus expected to yield a realistic qualitative description of isomerisation reactions.

The calculated properties in Table 7.3 were performed at a higher level of theory, namely B3LYP-D3(BJ)/cc-pVTZ in the gas phase. The molecular surface areas and volumes were calculated using a solvent-sphere radius of 0.1Å. The van der Waals intramolecular interaction energies were performed at the B3LYP-D3(BJ)/cc-pVTZ level of theory with basis-set superposition error[315] corrected for using the counterpoise approach.[316]

8   Conclusions and outlook

This Thesis dealt with the development of the Polytope Formalism and its application to isomerism and associated isomerisation. The Polytope Formalism's mathematical underpinnings, based upon the sub-disciplines of Combinatorics and Graph Theory, allows for a rigorous and systematic treatment of the subject matter and provides a unifying conceptual framework for both main branches of isomerism: stereoisomerism and constitutional isomerism.

This Thesis focusses on the conceptual frameworks and relationships of isomerism; the necessary formal definitions, nomenclature, and representations that have impacts reaching into unexpected areas such as patent specifications; the requisite controlled and precise vocabulary that facilitates nuanced communication; and the digital/computational formalisms that underpin the chemistry software and database tools that empower chemists to perform much of their work.

Chapter 1 presented the current state of isomerism and the deficiencies and challenges addressed in this work.

Chapter 2 examined problems with the existing isomerism-related terminology and introduced new terms required for the concise and precise description and discussion of the challenges identified in Chapter 1 and addressed in this work.

Chapter 3 gave a detailed overview of the polytopal-rearrangement model of stereoisomerisation at a single "stereocentre". Previously, this was focussed on inorganic higher coordination numbers and on describing unexpected fluxional behaviour. In this work, it was shown that this model could be generalised and given sufficient mathematical rigour for the new purposes of classifying all stereoisomerism and describing all allowed unimolecular reaction pathways for which the $R_{st}^c 1$ symbol was introduced. This complete and general approach was named the Polytope Formalism of stereoisomerism. A key finding was the comprehensive description of the scope of structural and dynamical phenomena available to the ML_2 stereocentre embedded within a larger molecular framework. Key to this was the inclusion of all possible non-dissociative vibrational motions and the embedding of the stereocentre in its external environment. For ML_2 and ML_3, sets of polytopes, generally useful

for classifying isomerism, were developed of different sizes, each of which was closed under the operation of all included $R^c_{st}1$ reactions. Concise sets were developed in which all polytopes related single $R^c_{st}1$ reactions have different symmetries, allowing unique assignment of any real embedded stereocentre structure to a set member. These sets are expected to be able to be descriptive of most, if not all, ML_2 and ML_3 isomerism and isomerisation, but expanded sets are possible should the need arise. This classification scheme is named the Polytope Formalism of stereoisomerism. It is constructed within the rigours of Graph Theory and properly accounts for singularities in both local and global coordinate-system definitions.

Chapter 4 describes the experimental demonstration of bond-angle reflection (BAR) stereoisomerism involving an ML_2 stereocentre. This unrecognised type of stereoisomerism was named akamptisomerism and is the final fundamental form of stereoisomerism. In this work, akamptisomerism was demonstrated in four stereoisomeric *transoid* B(F)–O–B(F) quinoxalinoporphyrin complexes (B_2OF_2pqx). The stereoisomerisation of the B_2OF_2pqx isomers was shown to follow a strict diastereomeric interconversion pattern and to exclusively operate by a stereospecific BAR $R^c_{st}1$ mechanism passing through a linear B–O–B geometry.

The unprecedented nature of the demonstrated stereoisomerism arising from purely BAR-related geometric differences demanded new terminology, 2D depiction, and nomenclatural stereodescriptors. IUPAC recommended wedge-hash notation for stereo-structures failed to adequately represent the B_2OF_2pqx isomers, with non-standard usage therefore introduced. In the absence of existing and adequate IUPAC stereodescriptors, the akamptisomerism-specific stereodescriptors *parvo* and *amplo* were introduced and shown to be compatible with existing stereodescriptors.

The manifestation of akamptisomerism was only possible due to the embedding constraints imposed by the porphyrin macrocycle upon the B(F)OB(F) group, fortuitously placing the stereoisomerisation kinetics within a "goldilocks zone" where both stereoisomer isolability and observable dynamical behaviour were simultaneously present.

Chapter 5 directly addressed the unrecognised nature of akamptisomeric systems. The structural diversity and failing of existing stereochemical practice regarding molecules with an ML_2 stereogenic unit was explored in detail. Oxygen-bridged oxabicyclo[*m.n*.1] systems were identified as a simple but large class of akamptisomeric compounds based upon

the embedding principle. A restrictive, and hence limited, search of the Chemical Abstracts Service (CAS) database revealed over 440 000 compounds that are akamptisomeric, many of which have relevance to materials science, natural products, and medicinal chemistry. These were examined in detail revealing that whilst most were amply described within existing IUPAC recommendations and by modern software, approximately five percent were identified as being poorly characterised and poorly represented digitally. Some of these problems were identified as arising from deficiencies in the IUPAC stereochemistry prescriptions that are biased towards common smaller bicyclo compounds. Oxygen-bridged oxabicyclo[$m.n.1$] systems of medium size, including the B_2OF_2pqx stereoisomers examined in Chapter 4, are identified as exhibiting an increased scope of stereoisomerism exclusively associated with akamptisomerism. The structural requirements needed to exhibit this increased scope is presented and the *parvo* and *amplo* stereodescriptors shown to be generally applicable.

Inclusion of akamptisomerism-related details to the 2D depiction, systematic naming, and digital representation thus enables a comprehensive stereoisomeric treatment with important implications for patent specifications and machine-led chemical research.

Taking a bigger perspective, the whole current *ad hoc* approach to stereochemistry for arbitrary ML_n stereocentres leads to a call for the current IUPAC stereochemical recommendations to be systematised. This supports the assertion that the Polytope Formalism of stereoisomerism provides a common and systematic conceptual framework for terminology, nomenclature, and representation.

Chapter 6 described how, at a more abstract level, the same Combinatorics and Graph Theory underpinnings of the Polytope Formalism of stereoisomerism as described in Chapter 3 could be directly adapted as a Polytope Formalism of constitutional isomerism. This was possible because whereas stereoisomerism deals with all possible *spatial permutations*, constitutional isomerism deals with all possible *bond topology permutations*. As with stereoisomerism, the Polytope Formalism of constitutional isomerism not only defines all isomers but also all possible allowed unimolecular rearrangement reactions for which the $R_{to}^c 1$ symbol was introduced in direct analogy to $R_{st}^c 1$. Given that the same mathematical approach equally describes both stereoisomerism and constitutional isomerism, the Polytope Formalism provides a unifying conceptual framework for all isomerism.

Whilst mathematically rigorous and complete, the Polytope Formalism of constitutional isomerism scales superexponentially of the order $\mathcal{O}(2^{N^2})$ as the number of atoms *N*. To provide tractability for many problems of practical interest, a simplifying approach was described where a chemical system can be partitioned into a small number of active components called bonders $\mathcal{B}$ and bonding sites $\mathcal{S}$. This is called the $\mathcal{S}_m\mathcal{B}_n$ class. This partitioning approach was described in detail through the worked example of inner hydrogen tautomerisation of free-base porphyrin representing the D_{4h} site symmetric $\mathcal{S}_4\mathcal{B}_2$ class.

Taxonomic terminology closely analogous to that used in the Polytope Formalism of stereoisomerism was used to classify different bond topology structures. For free-base porphyrin, 225 distinct bond topology species were found grouped into 28 distinct genera, further grouped into 10 families. High-level quantum chemical calculations were used to characterise each of the 28 generic structures. The energies and character of the generic structures were then combined with graphs of the species and species relationships to create compact representations of the potential-energy surface (PES). It was shown that small modifications to the porphyrin system can have a dramatic reorganisation of the PES. Further, the different generic structures were shown to provide structural prototypes for other porphyrin/porphyrinoid complexes. This demonstrated the broader utility of this specific $\mathcal{S}_4\mathcal{B}_2$ partitioning approach to the Polytope Formalism of constitutional isomerism to porphyrin chemistry beyond just free-base porphyrin.

A diversity of other molecular systems was described for which the $\mathcal{S}_m\mathcal{B}_n$ partitioning approach to the Polytope Formalism of constitutional isomerism can be applied indicating that it could be broadly applied to many practical problems of interest to chemists. To facilitate such work, numerous software tools are provided for generating species, genera and families for arbitrary *m*, $n \leq 8$ and software for generating graphs describing the allowed reaction pathways between species.

In Chapter 7 the focus is turned toward a practical implementation of some of the findings in this work. A new concept for drug design is presented where the fusion bond of a polyaromatic group is replaced by a bridging group to give a new aromatic annulene structure. This design transformation was called "pontation" with the resulting molecular group being isosteric to the parent polyaromatic but now featuring two distinct faces of the group. If the bridging group is an oxygen atom, the resulting oxabicyclo annulenes are

akamptisomeric and correspond to a subset of structures as described in Chapter 5. In principle, the bridge group could be NH, CH_2, CF_2 or similar.

Pontation on existing polyaromatic drugs were modelled and docking calculations performed to assess their binding to drug targets relative to their parent compounds. Two of the three systems studied in detail indicated improved binding. A survey of databases indicated that pontation could be applied to a significant fraction of medicinal chemistry relevant compounds.

DFT calculations indicate that pontated molecules exhibit increased van der Waals intramolecular interactions. The nature of the bridging group also offers scope for additional intramolecular interactions.

Pontation can also be applied to non-aromatic compounds like azulenes and isoindoles to give aromatic annulenes that are stabilised relative to the parent compounds, thus offering additional pharmacophores for drug discovery research.

Pontation can be described as a type of "internal substitution" that is fundamentally different to traditional approaches where ring atoms or peripheral atoms are substituted for. As such, the concept of pontation does not explicitly exist in current drug patent specifications. This provides medicinal chemistry with new approaches for patent extension or child patents.

8.1   Outlook

The findings of this Thesis are of fundamental importance to the entire fields of isomerism and isomerisation in their broadest sense. In providing a refined look at a seemingly well-established landscape, it will allow a fresh perspective and hitherto unrecognised possibilities to be identified and exploited.

This new way of looking at isomerism will mandate a change in the way that it is taught in undergraduate settings. Whilst the discovery of akamptisomerism "shook the tree" with respect to stereoisomerism, it has nonetheless been taught as yet another *ad hoc* extension to the subject. The Polytope Formalism framework presented in this Thesis provides a systematic approach. The appeal of any systematic approach is that is guarantees

completeness of treatment. Regarding isomerism, the Polytope Formalism is expected to illuminate chemical space beyond what is currently known and define all that is possible.

The most immediate further developments are:

1.   The mathematical proofs of the stereoisomerism treatment of ML_n for arbitrary *n*.

2.   The merging of the treatments of the Polytope Formalisms of stereoisomerism and constitutional isomerism so that both may be *simultaneously* addressed. The common conceptual and mathematical underpinnings of these Polytope Formalisms are designed to facilitate this next step.

More broadly, through application of the concepts expounded in this Thesis we expect development in the following areas:

1.   The empowering of machine-led approaches to chemistry. This includes high-throughput virtual screening of drug candidates, rational drug design, and functional material research. The conformational analyses performed for biochemistry can be viewed as a special case of the Polytope Formalism of stereoisomerism.

2.   The Polytope Formalism framework provides a way of evaluating existing and suggesting new chemical mechanistic processes. As the formalisms define strictly allowed elementary processes, they are also capable of ruling out unphysical processes. Applications are very broad including synthetic reaction mechanisms, enzymatic mechanisms, and synthetic catalysis.

3.   The Graph Theory features of the Polytope Formalisms provide a means of compactly and concisely describing potential energy surfaces. As much physical chemistry is focussed on such things, the formalisms are expected to prove a powerful tool.

4.   Following on from the Polytope Formalisms' capacity to describe potential energy surfaces, Graph Theory tools can be brought to bear to solve problems involving complex webs of interacting processes.

5.   A further corollary is that the rational design of catalysts could be greatly empowered through the application of the formalisms' principles.

6.   Molecules featuring *transoid* B(F)OB(F)-macrocycle complexes provide an invertible electrical dipole. A patent for a fully scalable (mesoscopic to individual molecular)

ferroelectric material based upon this has already been filed by the author Peter Canfield.

Additionally, as the work of this Thesis often deals with pedagogical details, very many shortcomings in IUPAC's definitions have been exposed. With so many governmental regulatory and legal definitions directly or indirectly reliant upon these, it is important that IUPAC reassess and updates its "Gold Book" terminology.

Further projects include the establishment of Wikipedia pages and the formatting of this Thesis into book form for both students and researchers.

As a final "blue sky" speculation, it should be noted that whilst all discussion in the Thesis has been firmly planted in the realm of chemistry, there are obvious instances where a Polytope Formalism could enter the demesne of nuclear physics. As a simple example, the displacement vector for the bond-stretch vibration of a diatomic operates in *both* directions: one leading to dissociation, the other, if permitted, leads to fusion of the nuclei. Whether the Polytope Formalism has any utility in such areas remains to be seen.

8.2   Afterword

This entire philosophical discourse pivoted around the chance observation of a small, but persistent, contamination issue when attempting to chirally resolve the B(F)OB(F)-quinoxalinoporphyrin stereoisomers in 2006. Recognising the extraordinary confluence of factors that needed to be in place leading to that, and thereafter all the fortuitous further developments, is both astounding and humbling. To describe this as "remarkable" hardly does it justice.

References

1.   Wolfram Research, I. *Mathematica*, 11.3; Wolfram Research, Inc.: Champaign, Illinois, 2018.

2.   Favre, H. A., Powell, Warren H, *Nomenclature of Organic Chemistry*; The Royal Society of Chemistry, 2014, p P001-1568.

3.   Connelly, N. G.; Damhus, T.; Hartshorn, R.; Hutton, A. T., *IUPAC Red Book*; RSC Publishing: London, 2005.

4.   McNaught, A. D.; Wilkinson, A., *IUPAC Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*; Blackwell: Oxford, 1997.

5.   Kuhn, T. S., *The structure of scientific revolutions*; University of Chicago Press: Chicago, 1962.

6.   Balaban, A. T., *Chemical applications of graph theory*. Academic Press: London, 1976.

7.   Canfield, P. J.; Blake, I. M.; Cai, Z.-L.; Luck, I. J.; Krausz, E.; Kobayashi, R.; Reimers, J. R.; Crossley, M. J., A new fundamental type of conformational isomerism. *Nat. Chem.* **2018**, *10*, 615-624.

8.   Le Bel, J.-A., Sur les relations qui existent entre les formules atomiques des corps organiques et le pouvoir rotatoire de leurs dissolutions. *Bulletin de la Société Chimique de Paris* **1874**, *22* 337–347.

9.   Van't Hoff, J. H., Sur les formules de structure dans l'espace. *Archives Néerlandaises des Sciences Exactes et Naturelles* **1874**, *9*, 445-454.

10.   Werner, A., Beitrag zur Konstitution anorganischer Verbindungen. *Zeitschrift für anorganische Chemie* **1893**, *3*, 267-330.

11.   Pasteur, L., On the relations that can exist between crystalline form, and chemical composition, and the sense of rotary polarization. *Annales de Chimie et de Physique* **1848**, *24*, 442-449.

12.   Roscoe, H. E.; Schorlemmer, C., *Treatise On Chemistry; Vol III Organic Chemistry*; Macmillan: London, 1890.

13.   Eiloart, A., Stereochemical models of organic molecules. *Am. Chem. J.* **1891**, *13*, 559-64.

14.   Sachse, H., Ueber die geometrischen Isomerien der Hexamethylenderivate. *Berichte der deutschen chemischen Gesellschaft* **1890**, *23*, 1363-1370.

15.   Mohr, E., Die Baeyersche Spannungstheorie und die Struktur des Diamanten. *Journal für Praktische Chemie* **1918**, *98*, 315-353.

16.   King, H., The possibility of a new instance of optical activity without an asymmetric carbon atom. *Proc. Chem. Soc. London* **1914**, *30*, 249-251.

17.   Cain, J. C.; Micklethwait, F. M. G., Studies in the diphenyl series. Part VI. The configuration of diphenyl and its derivatives. *J. Chem. Soc. Trans.* **1914**, *105*, 1437-1441.

18.   le Bel, J. A., Derivatives of ammonium chloride. *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences* **1890**, *110*, 144-147.

19.   Horner, L.; Winkler, H.; Rapp, A.; Mentrup, A.; Hoffmann, H.; Beck, P., Phosphororganische verbindungen optisch aktive tertiäre phosphine aus optisch aktiven quartären phosphoniumsalzen. *Tetrahedron Lett.* **1961**, *2*, 161-166.

20.   Brois, S. J., Aziridines. XII. Isolation of a stable nitrogen pyramid. *J. Am. Chem. Soc.* **1968**, *90*, 508-9.

21.   Gutowsky, H. S.; McCall, D. W.; Slichter, C. P., Nuclear Magnetic Resonance Multiplets in Liquids. *J. Chem. Phys.* **1953**, *21*, 279-292.

22.   Boger, D. L.; Miyazaki, S.; Kim, S. H.; Wu, J. H.; Castle, S. L.; Loiseleur, O.; Jin, Q., Total Synthesis of the Vancomycin Aglycon. *J. Am. Chem. Soc.* **1999**, *121*, 10004-10011.

23.   Nicolaou, K. C.; Harrison, S. T., Total Synthesis of Abyssomicin C, Atrop-abyssomicin C, and Abyssomicin D: Implications for Natural Origins of Atrop-abyssomicin C. *J. Am. Chem. Soc.* **2007**, *129*, 429-440.

24.   Witzig, R. M.; Faseke, V. C.; Haussinger, D.; Sparr, C., Atroposelective synthesis of tetra-ortho-substituted biaryls by catalyst-controlled non-canonical polyketide cyclizations. *Nat. Catal.* **2019**, *2*, 925-930.

25.   Reisberg, S. H.; Gao, Y.; Baran, P. S.; Walker, A. S.; Helfrich, E. J. N.; Clardy, J., Total synthesis reveals atypical atropisomerism in a small-molecule natural product, tryptorubin A. *Science* **2020**, *367*, 458-463.

26.   Muetterties, E. L.; Storr, A. T., Topological analysis of polytopal rearrangements. Sufficient conditions for closure. *J. Am. Chem. Soc.* **1969**, *91*, 3098-9.

27.   Muetterties, E. L., Polytopal form and isomerism. *Tetrahedron* **1974**, *30*, 1595-604.

28.   Muetterties, E. L.; Guggenberger, L. J., Idealized polytopal forms. Description of real molecules referenced to idealized polygons or polyhedra in geometric reaction path form. *J. Am. Chem. Soc.* **1974**, *96*, 1748-56.

29.   Muetterties, E. L., Topological representation of stereoisomerism. I. Polytopal rearrangements. *J. Am. Chem. Soc.* **1969**, *91*, 1636-43.

30.   Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235-242.

31.   Lavoisier, A., *Elements of Chemistry* William Creech: Edinburgh, 1790.

32.   IUPAC What we do. https://iupac.org/what-we-do/ (accessed 12 October 2021).

33.   Blackman, A., *Chemistry*, Fourth edition. ed.; Wiley: Milton, Qld, 2019.

34.   Fischer, E.; Scheibler, H., Zur Kenntnis der Waldenschen Umkehrung. II. *Berichte der deutschen chemischen Gesellschaft* **1908**, *41*, 889-893.

35.   Walden, P., Sur l'inversion optique des composés organiques1. *J. Chim. Phys.* **1911**, *9*, 160-197.

36.   Fischer, E., Nachtrag zu der Abhandlung: Waldensche Umkehrung und Substitutionsvorgang. *Justus Liebigs Annalen der Chemie* **1912**, *386*, 374-386.

37.   Berry, R. S., Correlation of Rates of Intramolecular Tunneling Processes, with Application to Some Group V Compounds. *J. Chem. Phys.* **1960**, *32*, 933-938.

38.   Muetterties, E. L., Topological representation of stereoisomerism. II. The five-atom family. *J. Am. Chem. Soc.* **1969**, *91*, 4115-22.

39.   Meakin, P.; Guggenberger, L. J.; Jesson, J. P.; Gerlach, D. H.; Tebbe, F. N.; Peet, W. G.; Muetterties, E. L., Structure and stereochemical nonrigidity of six-coordinate complexes. *J. Am. Chem. Soc.* **1970**, *92*, 3482-4.

40.   Muetterties, E. L.; Wiersema, R. J.; Hawthorne, M. F., Detection of polytopal isomers in the solution state. I. Eight-atom family. *J. Am. Chem. Soc.* **1973**, *95*, 7520-2.

41.   Guggenberger, L. J.; Muetterties, E. L., Reaction path analysis. 2. The nine-atom family. *J. Am. Chem. Soc.* **1976**, *98*, 7221-5.

42.   Hoffmann, R.; Beier, B. F.; Muetterties, E. L.; Rossi, A. R., Seven-coordination. A molecular orbital exploration of structure, stereochemistry, and reaction dynamics. *Inorg. Chem.* **1977**, *16*, 511-22.

43.   Muetterties, E. L., Stereochemically nonrigid structures. *Acc. Chem. Res.* **1970**, *3*, 266-273.

44.   Ugi, I.; Marquarding, D.; Klusacek, H.; Gillespie, P.; Ramirez, F., Berry pseudorotation and turnstile rotation. *Acc. Chem. Res.* **1971**, *4*, 288-296.

45. Bartell, L. S.; Plato, V., Gillespie-Nyholm aspects of force fields. I. Points-on-a-sphere and extended Hueckel molecular orbital analyses of trigonal bipyramids. *J. Am. Chem. Soc.* **1973**, *95*, 3097-3104.

46. Strich, A.; Veillard, A., Electronic structure of phosphorus pentafluoride and polytopal rearrangement in phosphoranes. *J. Am. Chem. Soc.* **1973**, *95*, 5574-5581.

47. Eisenhut, M.; Mitchell, H. L.; Traficante, D. D.; Kaufman, R. J.; Deutch, J. M.; Whitesides, G. M., Pseudorotation in XPF4. *J. Am. Chem. Soc.* **1974**, *96*, 5385-5397.

48. Westheimer, F. H., Essay 10 - The Polytopal Rearrangement at Phosphorus. In *Organic Chemistry: A Series of Monographs*, de Mayo, P., Ed. Academic Press: 1980; Vol. 42, pp 229-271.

49. Whitesides, G. M.; Eisenhut, M.; Bunting, W. M., Pseudorotation in arylbis(4,4'-dimethyl-2,2'-biphenylylene)phosphoranes. *J. Am. Chem. Soc.* **1974**, *96*, 5398-5407.

50. Mesch, K. A.; Quin, L. D., Syn to anti isomerization and degradation of phosphines of the 7-phosphanorbornene system by action of methanol. *Tetrahedron Lett.* **1980**, *21*, 4791-4794.

51. Mislow, K., Role of pseudorotation in the stereochemistry of nucleophilic displacement reactions. *Acc. Chem. Res.* **1970**, *3*, 321-331.

52. Minkin, V. I., Stereochemical and mechanistic consequences of the polarity rule and polytopic rearrangements in three-coordinate structures of Group VI elements. *Zh. Org. Khim.* **1977**, *13*, 1129-1137.

53. Kharabaev, N. N.; Minkin, V. I., Theoretical Modeling of the Stereoisomerization of Tetracoordinated Be(II) Bis(chelate) Complexes. *Russian Journal of Coordination Chemistry* **2003**, *29*, 525-528.

54. Minkin, V. I.; Nivorozhkin, L. E.; Korobov, M. S., Stereodynamics and degenerate ligand exchange in solutions of tetracoordinate chelate complexes of nontransition metals. *Russ. Chem. Rev.* **1994**, *63*, 289-311.

55. Minkin, V. I., Theoretical study of mechanisms of polytopical rearrangements of molecules in sulfurans and their analogs. *Zh. Org. Khim.* **1975**, *11*, 1993-2001.

56. Minkin, V. I., Theoretical study of the mechanisms of polytopic rearrangements of sulfurane molecules and their analogs. *Kvant. Khimiya* **1975**, 179.

57. Mauksch, M.; Schleyer, P. v. R., Effective Monkey Saddle Points and Berry and Lever Mechanisms in the Topomerization of SF4 and Related Tetracoordinated AX4 Species. *Inorg. Chem.* **2001**, *40*, 1756-1769.

58. Wakatsuki, Y.; Yamazaki, H., Formation of dicyanocobaltacyclopentanes from a cobalt(acrylonitrile) complex and isomerization of the cis-isomers via a polytopal rearrangement. *J. Chem. Soc., Chem. Commun.* **1980**, 1270-1271.

59. Farrugia, L. J., Novel metal framework rearrangements in mercury-bridged cluster complexes: X-ray crystal structures of[Hg{Fe2M(μ3-COMe)(CO)7(η-C5H5)}2](M □Co and Rh). *J. Chem. Soc., Chem. Commun.* **1987**, 147-149.

60. Tatsumi, K.; Nakamura, A.; Komiya, S.; Yamamoto, A.; Yamamoto, T., An associative mechanism for reductive elimination of d8 NiR2(PR3)2. *J. Am. Chem. Soc.* **1984**, *106*, 8181-8188.

61. Minkin, V. I.; Pichko, V. A.; Abubikerova, I. A.; Simkin, B. Y., An MNDO study of mechanisms of rapid intramolecular enantiomerization of bis-chelate Hg(II) complexes. *J. Mol. Struct.: THEOCHEM* **1991**, *227*, 295-304.

62. Rzepa, H. S.; Cass, M. E., A Computational Study of the Nondissociative Mechanisms that Interchange Apical and Equatorial Atoms in Square Pyramidal Molecules. *Inorg. Chem.* **2006**, *45*, 3958-3963.

63.	Osborn, J. A.; Shapley, J. R., Rapid intramolecular rearrangements in pentacoordinate transition metal compounds. Rearrangement mechanism of some fluxional iridium(I) complexes. *J. Am. Chem. Soc.* **1970**, *92*, 6976-6978.

64.	Thorn, D. L.; Hoffmann, R., The olefin insertion reaction. *J. Am. Chem. Soc.* **1978**, *100*, 2079-2090.

65.	Tatsumi, K.; Hoffmann, R.; Yamamoto, A.; Stille, J. K., Reductive Elimination of d8-Organotransition Metal Complexes. *Bull. Chem. Soc. Jpn.* **1981**, *54*, 1857-1867.

66.	César, V.; Bellemin-Laponnaz, S.; Gade, Lutz H., Cationic and Neutral Rhodium(I) Oxazolinylcarbene Complexes. *Eur. J. Inorg. Chem.* **2004**, *2004*, 3436-3444.

67.	Daul, C.; Niketic, S.; Rauzy, C.; Schläpfer, C.-W., Polytopal Rearrangement of [Ni(acac)2(py)]: A New Square Pyramid⇄Trigonal Bipyramid Twist Mechanism. *Chemistry – A European Journal* **2004**, *10*, 721-727.

68.	Khan, R. K. M.; Zhugralin, A. R.; Torker, S.; O'Brien, R. V.; Lombardi, P. J.; Hoveyda, A. H., Synthesis, Isolation, Characterization, and Reactivity of High-Energy Stereogenic-at-Ru Carbenes: Stereochemical Inversion through Olefin Metathesis and Other Pathways. *J. Am. Chem. Soc.* **2012**, *134*, 12438-12441.

69.	Torker, S.; Khan, R. K. M.; Hoveyda, A. H., The Influence of Anionic Ligands on Stereoisomerism of Ru Carbenes and Their Importance to Efficiency and Selectivity of Catalytic Olefin Metathesis Reactions. *J. Am. Chem. Soc.* **2014**, *136*, 3439-3455.

70.	Asatryan, R.; Ruckenstein, E.; Hachmann, J., Revisiting the polytopal rearrangements in penta-coordinate d7-metallocomplexes: modified Berry pseudorotation, octahedral switch, and butterfly isomerization. *Chemical Science* **2017**, *8*, 5512-5525.

71.	Minyaev, R. M.; Minkin, V. I., Theoretical study of the mechanism of associative nucleophilic substitution in compounds of five-coordinated nontransition elements of groups V and VI. *Journal of Structural Chemistry* **1979**, *20*, 715-725.

72.	De Keijzer, A. E. H.; Koole, L. H.; Buck, H. M., Pseudorotation in pentacoordinated phosphorus compounds. The influence of the conformational transmission effect on the barriers to pseudorotation in cyclic alkoxyphosphoranes. *J. Am. Chem. Soc.* **1988**, *110*, 5995-6001.

73.	Couzijn, E. P. A.; Slootweg, J. C.; Ehlers, A. W.; Lammertsma, K., Stereomutation of Pentavalent Compounds: Validating the Berry Pseudorotation, Redressing Ugi's Turnstile Rotation, and Revealing the Two- and Three-Arm Turnstiles. *J. Am. Chem. Soc.* **2010**, *132*, 18127-18140.

74.	Vancea, L.; Bennett, M. J.; Jones, C. E.; Smith, R. A.; Graham, W. A. G., Stereochemically nonrigid six-coordinate metal carbonyl complexes. 1. Polytopal rearrangement and x-ray structure of tetracarbonylbis(trimethylsilyl)iron. *Inorg. Chem.* **1977**, *16*, 897-902.

75.	Kreiter, C. G.; ÖZkar, S., Gehinderte Ligandenbewegungen in übergangsmetall-komplexen: XIV. Das dynamische vërhalten von η-dienchrom(0) - und molybdän(0)-komplexen. *J. Organomet. Chem.* **1978**, *152*, C13-C18.

76.	Kampe, C. E.; Boag, N. M.; Knobler, C. B.; Kaesz, H. D., Synthesis and characterization of edge-double-bridged Ru3(.mu.-H,.mu.-X)(CO)10 (X = Cl, Br, I) and face-triple-bridged Ru3(.mu.-H,.mu.3-I)(CO)9. Crystal and molecular structure of Ru3(.mu.-H,.mu.-Br)(CO)10, Ru3(.mu.-I)2(CO)10 and Ru3(.mu.-H,.mu.3-I)(CO)9. *Inorg. Chem.* **1984**, *23*, 1390-1397.

77.	Lee, C. Y.; Wang, Y.; Liu, C. S., Stereochemical nonrigidity in six-coordinate monochelate complexes via polytopal rearrangement. *Inorg. Chem.* **1991**, *30*, 3893-3899.

78.	Soubra, C.; Oishi, Y.; Albright, T. A.; Fujimoto, H., Intramolecular Rearrangements in Six-Coordinate Ruthenium and Iron Dihydrides. *Inorg. Chem.* **2001**, *40*, 620-627.

79.   Rzepa, H. S.; Cass, M. E., In Search of the Bailar and Rây−Dutt Twist Mechanisms That Racemize Chiral Trischelates: A Computational Study of ScIII, TiIV, CoIII, ZnII, GaIII, and GeIV Complexes of a Ligand Analogue of Acetylacetonate. *Inorg. Chem.* **2007**, *46*, 8024-8031.

80.   Trindle, C.; Datta, S. N.; Bouman, T. D., Nonrigid molecule effects on the energy levels of XeF6. *Int. J. Quantum Chem.* **1977**, *11*, 627-664.

81.   Hawthorne, S. L.; Fay, R. C., Stereochemically rigid seven-coordinate complexes. Synthesis, structure, and dynamic stereochemistry of chlorotris(N,N-dialkylmonothiocarbamato)titanium(IV) complexes. *J. Am. Chem. Soc.* **1979**, *101*, 5268-5277.

82.   Dedieu, A., Theoretical study of the olefin insertion step in the chlorotris(triphenylphosphine)rhodium(I)-catalyzed hydrogenation of olefins. *Inorg. Chem.* **1981**, *20*, 2803-2813.

83.   Fay, R. C.; Weir, J. R.; Bruder, A. H., Nuclear magnetic resonance studies of stereochemical rearrangements in pentagonal-bipyramidal (.eta.5-cyclopentadienyl)tris(N,N-dimethyldithiocarbamato)titanium(IV), -zirconium(IV), and -hafnium(IV). *Inorg. Chem.* **1984**, *23*, 1079-1089.

84.   Burdett, J. K.; Hoffmann, R.; Fay, R. C., Eight-coordination. *Inorg. Chem.* **1978**, *17*, 2553-2568.

85.   Fay, R. C.; Howie, J. K., Stereochemistry and stereochemical rearrangements of eight-coordinate tetrakis chelates. 1. Group 4B .beta.-diketonates. *J. Am. Chem. Soc.* **1979**, *101*, 1115-1122.

86.   Weir, J. R.; Fay, R. C., Stereochemistry and metal-centered rearrangements of eight-coordinate niobium(V) and tantalum(V) dithiocarbamates and monothiocarbamates. *Inorg. Chem.* **1986**, *25*, 2969-2976.

87.   Zou, W.; Tao, Y.; Kraka, E., Describing Polytopal Rearrangements of Fluxional Molecules with Curvilinear Coordinates Derived from Normal Vibrational Modes: A Conceptual Extension of Cremer–Pople Puckering Coordinates. *J. Chem. Theory Comput.* **2020**, *16*, 3162-3193.

88.   Mingos, D. M. P.; McGrady, J. E.; Rohl, A. L., Moments of inertia in cluster and coordination compounds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; pp 1-54.

89.   Moss, G. P., Basic terminology of stereochemistry. *Pure Appl. Chem.* **1996**, *68*, 2193-2222.

90.   Casanova, D.; Cirera, J.; Llunell, M.; Alemany, P.; Avnir, D.; Alvarez, S., Minimal Distortion Pathways in Polyhedral Rearrangements. *J. Am. Chem. Soc.* **2004**, *126*, 1755-1763.

91.   Brecher, J., Graphical representation of stereochemical configuration (IUPAC Recommendations 2006). *Pure Appl. Chem.* **2009**, *78*, 1897-1970.

92.   Hong, M., Syntheses and properties of β-(acetoxyl)ethylsilanes. *Xiamen da xue xue bao Zi ran ke xue ban = Journal of Xiamen university (natural science).* **1989**, *28*, 618-622.

93.   Redfield, D. A.; Nelson, J. H., Mechanism of cis-trans isomerization for square planar complexes of the type ML2X2. *J. Am. Chem. Soc.* **1974**, *96*, 6219-6220.

94.   Stohrer, W. D.; Hoffmann, R., Bond-stretch isomerism and polytopal rearrangements in (CH)5+, (CH)5-, and (CH)4CO. *J. Am. Chem. Soc.* **1972**, *94*, 1661-1668.

95.   Bürgi, H., Molecular Geometry. Von R. J. Gillespie. Van Nostrand Reinhold Comp., London 1972. 1. Aufl., IX, 228 S., zahlr. Abb. u. Tab., geb. ca. DM 38. WILEY-VCH Verlag GmbH: Weinheim, 1974; Vol. 86, pp 565-565.

96.   Bürgi, H.-B., Stereochemistry of Reaction Paths as Determined from Crystal Structure Data—A Relationship between Structure and Energy. *Angew. Chem., Int. Ed.* **1975**, *14*, 460-473.

97.   Bürgi, H. B.; Dunitz, J. D., From crystal statics to chemical dynamics. *Acc. Chem. Res.* **1983**, *16*, 153-161.

98.   Jahn, H. A.; Teller, E.; Donnan, F. G., Stability of polyatomic molecules in degenerate electronic states - I—Orbital degeneracy. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences* **1937**, *161*, 220-235.

99.   Reimers, J. R.; Cai, Z.-L.; Bilić, A.; Hush, N. S., The appropriateness of density-functional theory for calculations of molecular electronic properties. *Ann. N.Y. Acad. Sci.* **2003**, *1006*, 235-51.

100.   Fukuzumi, S.; Ohkubo, K.; E, W.; Ou, Z.; Shao, J.; Kadish, K. M.; Hutchison, J. A.; Ghiggino, K. P.; Sintic, P. J.; Crossley, M. J., Metal-Centered Photoinduced Electron Transfer Reduction of a Gold(III) Porphyrin Cation Linked with a Zinc Porphyrin to Produce a Long-Lived Charge-Separated State in Nonpolar Solvents. *J. Am. Chem. Soc.* **2003**, *125*, 14984-14985.

101.   Reimers, J. R.; Hush, N. S.; Crossley, M. J., Inter-porphyrin coupling: how strong should it be for molecular electronics applications? *J. Porphyrins Phthalocyanines* **2002**, *06*, 795-805.

102.   Crossley, M. J.; Sintic, P. J.; Walton, R.; Reimers, J. R., Synthesis and physical properties of biquinoxalinyl bridged bis-porphyrins: models for aspects of Photosynthetic Reaction Centres. *Org. Biomol. Chem.* **2003**, *1*, 2777-2787.

103.   Crossley, M. J.; Try, A. C.; Walton, R., Synthesis of accurate distance models of the primary donor - primary acceptor pair of bacterial photosynthetic reaction centres. *Tetrahedron Lett.* **1996**, *37*, 6807-6810.

104.   Reimers, J. R.; Bilić, A.; Cai, Z.-L.; Dahlbom, M.; Lambropoulos, N. A.; Solomon, G. C.; Crossley, M. J.; Hush, N. S., Molecular Electronics: From Basic Chemical Principles to Photosynthesis to Steady-State Through-Molecule Conductivity to Computer Architectures. *Aust. J. Chem.* **2004**, *57*, 1133-1138.

105.   Lee, S.-H.; Blake, I. M.; Larsen, A. G.; McDonald, J. A.; Ohkubo, K.; Fukuzumi, S.; Reimers, J. R.; Crossley, M. J., Synthetically tuneable biomimetic artificial photosynthetic reaction centres that closely resemble the natural system in purple bacteria. *Chemical Science* **2016**, *7*, 6534-6550.

106.   Hutchison, J. A.; Sintic, P. J.; Brotherhood, P. R.; Scholes, C.; Blake, I. M.; Ghiggino, K. P.; Crossley, M. J., Control of Photoinduced Charge Transfer Lifetimes in Porphyrin Arrays by Ligand Addition. *J. Phys. Chem. C* **2009**, *113*, 11796-11804.

107.   Hutchison, J. A.; Sintic, P. J.; Crossley, M. J.; Nagamura, T.; Ghiggino, K. P., The photophysics of selectively metallated arrays of quinoxaline-fused tetraarylporphyrins. *Phys. Chem. Chem. Phys.* **2009**, *11*, 3478-3489.

108.   Yeow, E. K. L.; Sintic, P. J.; Cabral, N. M.; Reek, J. N. H.; Crossley, M. J.; Ghiggino, K. P., Photoinduced energy and electron transfer in bis-porphyrins with quinoxaline Tröger's base and biquinoxalinyl spacers. *Phys. Chem. Chem. Phys.* **2000**, *2*, 4281-4291.

109.   Schulze, T. F., et al., Efficiency Enhancement of Organic and Thin-Film Silicon Solar Cells with Photochemical Upconversion. *J. Phys. Chem. C* **2012**, *116*, 22794-22801.

110.   Crossley, M. J.; Sintic, P. J.; Hutchison, J. A.; Ghiggino, K. P., Chemical models for aspects of the photosynthetic reaction centre: synthesis and photophysical properties of tris- and tetrakis-porphyrins that resemble the arrangement of chromophores in the natural system. *Org. Biomol. Chem.* **2005**, *3*, 852-865.

111.   Lu, T. X.; Reimers, J. R.; Crossley, M. J.; Hush, N. S., Rigid Fused Oligoporphyrins as Potential Versatile Molecular Wires. 1. Geometry and Connectivity of 1,4,5,8-Tetraazaanthracene-Bridged Systems. *J. Phys. Chem.* **1994**, *98*, 11878-11884.

112.   Ohkubo, K.; Sintic, P. J.; Tkachenko, N. V.; Lemmetyinen, H.; E, W.; Ou, Z.; Shao, J.; Kadish, K. M.; Crossley, M. J.; Fukuzumi, S., Photoinduced electron-transfer dynamics

and long-lived CS states of donor–acceptor linked dyads and a triad containing a gold porphyrin in nonpolar solvents. *Chem. Phys.* **2006**, *326*, 3-14.

113. E, W., et al., Control of the Orbital Delocalization and Implications for Molecular Rectification in the Radical Anions of Porphyrins with Coplanar 90° and 180° β,β'-Fused Extensions. *J. Phys. Chem. A* **2008**, *112*, 556-570.

114. Reimers, J. R.; Lü, T. X.; Crossley, M. J.; Hush, N. S., Molecular electronic properties of fused rigid porphyrin-oligomer molecular wires. *Chem. Phys. Lett.* **1996**, *256*, 353-359.

115. Crossley, M. J.; Johnston, L. A., Laterally-extended porphyrin systems incorporating a switchable unit. *Chemical Communications* **2002**, 1122-1123.

116. Kadish, K. M., et al., Quinoxalino[2,3-b']porphyrins Behave as π-Expanded Porphyrins upon One-Electron Reduction: Broad Control of the Degree of Delocalization through Substitution at the Macrocycle Periphery. *J. Phys. Chem. B* **2007**, *111*, 8762-8774.

117. Kuroda, S.-i.; Azumi, R.; Matsumoto, M.; King, L. G.; Crossley, M. J., Electron spin resonance of Cu-porphyrin of dimer-type in Langmuir-Blodgett films. *Thin Solid Films* **1997**, *295*, 92-94.

118. Crossley, M. J.; Prashar, J. K., Thiophene-Appended Porphyrin Systems. *Tetrahedron Lett.* **1997**, *38*, 6751-6754.

119. Ou, Z.; E, W.; Shao, J.; Burn, P. L.; Sheehan, C. S.; Walton, R.; Kadish, K. M.; Crossley, M. J., Electrochemical and spectroelectrochemical properties of building blocks for molecular arrays: Reactions of quinoxalino[2,3-b]porphyrins containing metal(II) ions. *J. Porphyrins Phthalocyanines* **2005**, *9*, 142-151.

120. Crossley, M. J.; Govenlock, L. J.; Prashar, J. K., Synthesis of porphyrin-2,3,12,13- and 2,3,7,8-tetraones: building blocks for the synthesis of extended porphyrin arrays. *J. Chem. Soc. Chem. Commun.* **1995**, 2379-80.

121. Armstrong, R. S.; Foran, G. J.; Hough, W. A.; D'Alessandro, D. M.; Lay, P. A.; Crossley, M. J., Spectroelectrochemical evidence for communication within a laterally-bridged dimanganese(iii) bis-porphyrin. *Dalton Trans.* **2006**, 4805-4813.

122. Zhu, W.; Sintic, M.; Ou, Z.; Sintic, P. J.; McDonald, J. A.; Brotherhood, P. R.; Crossley, M. J.; Kadish, K. M., Electrochemistry and Spectroelectrochemistry of β,β'-Fused Quinoxalinoporphyrins and Related Extended Bis-porphyrins with Co(III), Co(II), and Co(I) Central Metal Ions. *Inorg. Chem.* **2010**, *49*, 1027-1038.

123. Ohkubo, K.; Garcia, R.; Sintic, P. J.; Khoury, T.; Crossley, M. J.; Kadish, K. M.; Fukuzumi, S., Change in the Site of Electron-Transfer Reduction of a Zinc–Quinoxalinoporphyrin/Gold–Quinoxalinoporphyrin Dyad by Binding of Scandium Ions and the Resulting Remarkable Elongation of the Charge-Shifted-State Lifetime. *Chemistry – A European Journal* **2009**, *15*, 10493-10503.

124. Ou, Z.; Kadish, K. M.; E, W.; Shao, J.; Sintic, P. J.; Ohkubo, K.; Fukuzumi, S.; Crossley, M. J., Substituent Effects on the Site of Electron Transfer during the First Reduction for Gold(III) Porphyrins. *Inorg. Chem.* **2004**, *43*, 2078-2086.

125. Ou, Z.; Zhu, W.; Fang, Y.; Sintic, P. J.; Khoury, T.; Crossley, M. J.; Kadish, K. M., Unusual Multi-Step Sequential AuIII/AuII Processes of Gold(III) Quinoxalinoporphyrins in Acidic Non-Aqueous Media. *Inorg. Chem.* **2011**, *50*, 12802-12809.

126. Ou, Z.; Khoury, T.; Fang, Y.; Zhu, W.; Sintic, P. J.; Crossley, M. J.; Kadish, K. M., Gold(III) Porphyrins Containing Two, Three, or Four β,β'-Fused Quinoxalines. Synthesis, Electrochemistry, and Effect of Structure and Acidity on Electroreduction Mechanism. *Inorg. Chem.* **2013**, *52*, 2474-2483.

127. Sendt, K.; Johnston, L. A.; Hough, W. A.; Crossley, M. J.; Hush, N. S.; Reimers, J. R., Switchable electronic coupling in oligoporphyrin molecular wires examined through the

measurement and assignment of electronic absorption spectra. *J. Am. Chem. Soc.* **2002**, *124*, 9299-309.

128.   Kashiwagi, Y.; Ohkubo, K.; McDonald, J. A.; Blake, I. M.; Crossley, M. J.; Araki, Y.; Ito, O.; Imahori, H.; Fukuzumi, S., Long-Lived Charge-Separated State Produced by Photoinduced Electron Transfer in a Zinc Imidazoporphyrin-C60 Dyad. *Org. Lett.* **2003**, *5*, 2719-2721.

129.   Lee, S.-H.; Larsen, A. G.; Ohkubo, K.; Cai, Z.-L.; Reimers, J. R.; Fukuzumi, S.; Crossley, M. J., Long-lived long-distance photochemically induced spin-polarized charge separation in β,β′-pyrrolic fused ferrocene-porphyrin-fullerene systems. *Chemical Science* **2012**, *3*, 257-269.

130.   Cheng, Y. Y.; Khoury, T.; Clady, R. G. C. R.; Tayebjee, M. J. Y.; Ekins-Daukes, N. J.; Crossley, M. J.; Schmidt, T. W., On the efficiency limit of triplet–triplet annihilation for photochemical upconversion. *Phys. Chem. Chem. Phys.* **2010**, *12*, 66-71.

131.   Larsen, J.; Brüggemann, B.; Khoury, T.; Sly, J.; Crossley, M. J.; Sundström, V.; Åkesson, E., Structural Induced Control of Energy Transfer within Zn(II)−Porphyrin Dendrimers. *J. Phys. Chem. A* **2007**, *111*, 10589-10597.

132.   Gallaher, J. K.; Wright, K. M.; Frazer, L.; MacQueen, R. W.; Crossley, M. J.; Castellano, F. N.; Schmidt, T. W., High efficiency deep red to yellow photochemical upconversion under solar irradiance. *Energy & Environmental Science* **2021**, *14*, 5541-5551.

133.   Yeow, E. K. L.; Ghiggino, K. P.; Reek, J. N. H.; Crossley, M. J.; Bosnian, A. W.; Scheming, A. P. H. J.; Meijer, E. W., The dynamics of electronic energy transfer in novel multiporphyrin functionalized dendrimers: A time-resolved fluorescence anisotropy study. *J. Phys. Chem. B* **2000**, *104*, 2596-2606.

134.   Hasobe, T.; Kashiwagi, Y.; Absalom, M. A.; Sly, J.; Hosomizu, K.; Crossley, M. J.; Imahori, H.; Kamat, P. V.; Fukuzumi, S., Supramolecular Photovoltaic Cells Using Porphyrin Dendrimers and Fullerene. *Advanced Materials* **2004**, *16*, 975-979.

135.   Larsen, J.; Brüggemann, B.; Polívka, T.; Sundström, V.; Åkesson, E.; Sly, J.; Crossley, M. J., Energy Transfer within Zn-Porphyrin Dendrimers: Study of the Singlet−Singlet Annihilation Kinetics. *J. Phys. Chem. A* **2005**, *109*, 10654-10662.

136.   Dastoor, P. C., et al., Understanding and Improving Solid-State Polymer/C60-Fullerene Bulk-Heterojunction Solar Cells Using Ternary Porphyrin Blends. *J. Phys. Chem. C* **2007**, *111*, 15415-15426.

137.   Larsen, J.; Andersson, J.; Polívka, T.; Sly, J.; Crossley, M. J.; Sundström, V.; Åkesson, E., Energy transfer and conformational dynamics in Zn–porphyrin dendrimers. *Chem. Phys. Lett.* **2005**, *403*, 205-210.

138.   Fukuzumi, S.; Hasobe, T.; Ohkubo, K.; Crossley, M. J.; Kamat, P. V.; Imahori, H., π-Complex formation in electron-transfer reactions of porphyrins. *J. Porphyrins Phthalocyanines* **2004**, *08*, 191-200.

139.   Hasobe, T.; Kamat, P. V.; Absalom, M. A.; Kashiwagi, Y.; Sly, J.; Crossley, M. J.; Hosomizu, K.; Imahori, H.; Fukuzumi, S., Supramolecular Photovoltaic Cells Based on Composite Molecular Nanoclusters: Dendritic Porphyrin and C60, Porphyrin Dimer and C60, and Porphyrin−C60 Dyad. *J. Phys. Chem. B* **2004**, *108*, 12865-12872.

140.   Curiel, D.; Ohkubo, K.; Reimers, J. R.; Fukuzumi, S.; Crossley, M. J., Photoinduced electron transfer in a β,β′-pyrrolic fused ferrocene–(zinc porphyrin)–fullerene. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5260-5266.

141.   Fukuzumi, S.; Ohkubo, K.; Saito, K.; Kashiwagi, Y.; Khoury, T.; Crossley, M. J., Multiple photosynthetic reaction centers composed of supramolecular assemblies of a zinc porphyrin dendrimer with pyridylnaphthalenediimide. *J. Porphyrins Phthalocyanines* **2011**, *15*, 1292-1298.

142. Belcher, W. J.; Boyd, P. D. W.; Brothers, P. J.; Liddell, M. J.; Rickard, C. E. F., New Coordination Mode for the Porphyrin Ligand in the Boron Porphyrin Complex B_2OF_2(TTP). *J. Am. Chem. Soc.* **1994**, *116*, 8416-8417.

143. Belcher, W. J.; Hodgson, M. C.; Sumida, K.; Torvisco, A.; Ruhlandt-Senge, K.; Ware, D. C.; Boyd, P. D. W.; Brothers, P. J., Porphyrin complexes containing coordinated BOB groups: synthesis, chemical reactivity and the structure of [BOB(tpClpp)]$^{2+}$. *Dalton Trans.* **2008**, 1602–1614.

144. Brothers, P. J., Boron complexes of porphyrins and related polypyrrole ligands: unexpected chemistry for both boron and the porphyrin. *Chem. Commun.* **2008**, 2090–2102.

145. Brothers, P. J., Boron Complexes of Pyrrolyl Ligands. *Inorg. Chem.* **2011**, *50*, 12374-12386.

146. Albrett, A. M.; Conradie, J.; Boyd, P. D. W.; Clark, G. R.; Ghosh, A.; Brothers, P. J., Corrole as a binucleating ligand: preparation, molecular structure and density functional theory study of diboron corroles. *J. Am. Chem. Soc.* **2008**, *130*, 2888-2889.

147. Xu, N.; Ono, T.; Morita, Y.; Komatsu, T.; Hisaeda, Y., Rectangular Holes in Porphyrin Isomers Act As Mono- and Binucleating Ligands: Stereochemistry of Mono- and Diboron Porphycenes and Their Protonation Behaviors. *Inorg. Chem.* **2021**, *60*, 574-583.

148. Brothers, P.; Tay, A. C. Y.; Frogley, B.; Ware, D. C.; Conradie, J.; Ghosh, A., Tetrahedral Pegs in Square Holes: Stereochemistry of Diboron Porphyrazines and Phthalocyanines. *Angew. Chemie Int. Ed.* **2018**, *58*, 3057-3061.

149. Tay, A. Synthesis of Boron Porphyrinoids. The University of Auckland, Auckland, New Zealand, 2017.

150. Moss, G. P., Nomenclature of tetrapyrroles. *Pure Appl. Chem.* **1987**, *59*, 779-832.

151. Goerigk, L.; Sharma, R., The INV24 test set: how well do quantum-chemical methods describe inversion and racemization barriers? *Can. J. Chem.* **2016**, *94*, 1133-1143.

152. Tay, A. C. Y.; Frogley, B. J.; Ware, D. C.; Brothers, P. J., Boron calixphyrin complexes: exploring the coordination chemistry of a BODIPY/porphyrin hybrid. *Dalton Trans.* **2018**, *47*, 3388-3399.

153. Frisch, M. J., et al. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

154. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648-52.

155. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J., Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623.

156. Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S., A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184-32215.

157. Hehre, W. J.; Ditchfield, R.; Pople, J. A., Self-consistent molecular orbital methods. XII. Further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules *J. Chem. Phys.* **1972**, *56*, 2257-61.

158. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

159. Tomasi, J.; Mennucci, B.; Cammi, R., Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999-3093.

160. Floris, F. M.; Tomasi, J.; Pascual Ahuir, J. L., Dispersion and repulsion contributions to the solvation energy: Refinements to a simple computational model in the continuum approximation. *J. Computat. Chem.* **1991**, *12*, 784-791.

161.   Stewart, J. J. P., Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *J. Mol. Model.* **2013**, *19*, 1-32.

162.   Hratchian, H. P.; Schlegel, H. B., Accurate reaction paths using a Hessian based predictor–corrector integrator. *J. Chem. Phys.* **2004**, *120*, 9918-9924.

163.   Hratchian, H. P.; Schlegel, H. B., Using Hessian Updating To Increase the Efficiency of a Hessian Based Predictor-Corrector Reaction Path Following Method. *J. Chem. Theory Comput.* **2005**, *1*, 61-69.

164.   Wolinski, K.; Hinton, J. F.; Pulay, P., Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations. *J. Am. Chem. Soc.* **1990**, *112*, 8251-8260.

165.   Yanai, T.; Tew, D. P.; Handy, N. C., A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51-57.

166.   Kobayashi, R.; Amos, R. D., The application of CAM-B3LYP to the charge-transfer band problem of the zincbacteriochlorin–bacteriochlorin complex. *Chem. Phys. Lett.* **2006**, *420*, 106–109.

167.   Chai, J.-D.; Head-Gordon, M., Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615-6620.

168.   Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J., Towards a systematic molecular orbital theory for excited states. *J. Phys. Chem.* **1992**, *96*, 135.

169.   Reimers, J. R.; Krausz, E., An analytical data inversion method for Magnetic Circular Dichroism spectra dominated by the "B-term". *Phys. Chem. Chem. Phys.* **2014**, *16*, 2315-2322

170.   Favre, H. A.; Powell, W. H., *Nomenclature of Organic Chemistry: IUPAC Recommendations and Preferred Names 2013. IUPAC–RSC. ISBN* . Royal Society of Chemistry- IUPAC Chemical Nomenclature and Structure Representation Division: London, 2013.

171.   Chemical Abstracts Service. https://www.cas.org/.

172.   Albrett, A. M.; Conradie, J.; Ghosh, A.; Brothers, P. J., DFT survey of monoboron and diboron corroles: regio- and stereochemical preferences for a constrained, low-symmetry macrocycle. *Dalton Trans.* **2008**, 4464-4473.

173.   Crossley, M. J.; Burn, P. L., Rigid, laterally-bridged bis-porphyrin systems. *J. Chem. Soc. Chem. Commun.* **1987**, 39.

174.   von Bruchhausen, F.; Oberembt, H.; Feldhaus, A., Über das Oxy-acanthin und Berbamin. III. *Justus Liebigs Ann. Chem.* **1933**, *507*, 144-159.

175.   Gordon, A. J.; Gallagher, J. P., Stereochemistry of dicoordinated oxygen. I. NMR studies of doubly ortho-bridged diphenyl ethers and the possibility of oxygen inversion. *Tetrahedron Lett.* **1970**, *11*, 2541-2544.

176.   Alvarez, E.; Díaz, M. T.; Pérez, R.; Ravelo, J. L.; Regueiro, A.; Vera, J. A.; Zurita, D.; Martín, J. D., Simple Designs for the Construction of Complex Trans-Fused Polyether Toxin Frameworks. A Linear Strategy Based on Entropically Favored Oxirane Ring Enlargement in Epoxycycloalkenes Followed by Carbon-Carbon or Carbon-Oxygen Bond-Forming Cyclizations. *J. Org. Chem.* **1994**, *59*, 2848-2876.

177.   Zárraga, M.; Alvarez, E.; Ravelo, J. L.; Rodriguez, V.; Rodriguez, M. L.; Martín, J. D., Approaches to the synthesis of the tetrahydropyran subunits of marine trans-fused polyether toxins. *Tetrahedron Lett.* **1990**, *31*, 1633-1636.

178.   Davin, L. B.; Lewis, N. G., Lignin primary structures and dirigent sites. *Curr. Opin. Biotechnol.* **2005**, *16*, 407-415.

179.   Gupta, R. B.; Kaloustian, M. K.; Franck, R. W.; Blount, J. F., Stereochemical studies on the atomic inversion of divalent oxygen: the use of dinaphtho[*a,j*]anthracenes to restrict rotational motion. *J. Am. Chem. Soc.* **1991**, *113*, 359-361.

180.   Simoni, D.; Roberti, M.; Rondanin, R.; Kozikowski, A. P., Enantioselective deprotonation of the 8-oxabicyclo[3.2.1]octan-3-one: synthesis of 8-oxa-norcocaines and 8-oxa-pseudonorcocaines. *Tetrahedron Lett.* **1999**, *40*, 4425-4428.

181.   Kozikowski, A. P.; Simoni, D.; Roberti, M.; Rondanin, R.; Wang, S.; Du, P.; Johnson, K. M., Synthesis of 8-Oxa analogues of norcocaine endowed with interesting cocaine-like activity. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1831-1836.

182.   Kainz, A.; Eiden, F., Investigations into pyran derivatives. 129. 8-Oxabicyclo[3.2.1]octanes: synthesis and structure of the four diastereomeric 8-oxacocaines. *Arch. Pharm. (Weinheim, Ger.)* **1990**, *323*, 191-4.

183.   Sondheimer, F.; Shani, A., 1,6-Oxido[10]annulene and 1-Benzoxepin. *J. Am. Chem. Soc.* **1964**, *86*, 3168-3169.

184.   Vogel, E.; Biskup, M.; Pretzer, W.; Böll, W. A., Cyclodecapentaenes with a 1,6-Heteroatomic Bridge. *Angew. Chem., Int. Ed.* **1964**, *3*, 642-643.

185.   Cope, A. C.; Anderson, B. C., Proximity Effects. VI. Stereospecific Syntheses of cis- and trans-1,4-Cyclooctanediols. *J. Am. Chem. Soc.* **1957**, *79*, 3892-3895.

186.   Vogel, E.; Böll, W. A., Substitution of 1,6-Methanocyclodecapentaene. *Angew. Chem., Int. Ed.* **1964**, *3*, 642-642.

187.   Anastassiou, A. G.; Kasmai, H. S.; Badri, R., The Question of Bicycloconjugation in Naphtho-Fused 9-Heterobicyclo[4.2.1]nona-2,4,7-trienes as Examined by 13C-NMR Spectroscopy. *Angew. Chem., Int. Ed.* **1980**, *19*, 639-640.

188.   Kuffner, U.; Schlögl, K., Optisch aktive derivate des 1,6-methano[10]anulens darstellung, chiroptische eigenschaften und optische reinheit. *Tetrahedron Lett.* **1971**, *12*, 1773-1776.

189.   McCrae, W.; Freid, J. H.; Edwards, J. A. 13,14-seco-1,2,3,4-tetrahydro-phenanthrenes. 3639462, 1972.

190.   Vogel, E.; Tuckmantel, W.; Schögl, K.; Widhalm, M.; Kraka, E.; Cremer, D., Zur konfigurativen stabilität syn/anti-isomerer überbrückter [14]annulene mit anthracen-perimeter. *Tetrahedron Lett.* **1984**, *25*, 4925-4928.

191.   Kim, S.; Jung, H. J.; Kim, J. C.; Lee, K.-S.; Park, S. S.; Dravid, V. P.; He, K.; Jeong, H. Y., In Situ Observation of Resistive Switching in an Asymmetric Graphene Oxide Bilayer Structure. *ACS Nano* **2018**, *12*, 7335-7342.

192.   Shani, A.; Sondheimer, F., Unsaturated macrocyclic compounds. LI. 1,6-Oxido[10]annulene. *J. Am. Chem. Soc.* **1967**, *89*, 6310-6317.

193.   Canfield, P. J. A molecule for an electro-active material. 2019/169444, 2019.

194.   Wu, J. A divalent compound for treating ovarian cancer. 105481876 A, 2016.

195.   Shafizadeh, F.; Bhadane, N. R., Badgerin, a new germacranolide from Artemisia arbuscula subspecies arbuscula. *J. Org. Chem.* **1972**, *37*, 274-277.

196.   Fraga, B. M.; Terrero, D.; Cabrera, I.; Reina, M., Studies on the sesquiterpene lactones from Laurus novocanariensis lead to the clarification of the structures of 1-epi-tatridin B and its epimer tatridin B. *Phytochemistry* **2018**, *153*, 48-52.

197.   Jakupovic, J.; Baruah, R. N.; Bohlmann, F.; King, R. M.; Robinson, H., New Types of Germacranolides from Austroliabum candidum1. *Planta Med.* **1986**, *52*, 204-206.

198.   Marco, J. A.; Sanz-cervera, J. F.; Carda, M.; Lex, J., Oxygenated germacranes from Santolina chamaecyparissus. *Phytochemistry* **1993**, *34*, 1549-1559.

199.   Appendino, G.; Aviello, G.; Ballero, M.; Borrelli, F.; Fattorusso, E.; Petrucci, F.; Santelia, F. U.; Taglialatela-Scafati, O., Cytotoxic germacrane sesquiterpenes from the aerial parts of Santolina insularis. *J. Nat. Prod.* **2005**, *68*, 853-857.

200. Valeria D'Auria, M.; Paloma, L. G.; Minale, L.; Riccio, R.; Debitus, C., Jereisterol A and B : Two 3β-methoxy-secosteroids from the pacific sponge Jereicopsis graphidiophora. *Tetrahedron Lett.* **1991**, *32*, 2149-2152.

201. Hu, Z., et al., Fusicoccane-Derived Diterpenoids from Alternaria brassicicola: Investigation of the Structure–Stability Relationship and Discovery of an IKKβ Inhibitor. *Org. Lett.* **2018**, *20*, 5198-5202.

202. Nakanishi, S.; Matsuda, Y., Protein kinase inhibitors of microbial origin and their application. *Cell Sci.* **1992**, *8*, 609-17.

203. Ljungman, M., Targeting the DNA Damage Response in Cancer. *Chem. Rev. (Washington, DC, U. S.)* **2009**, *109*, 2929-2950.

204. Singh, J.; Hudkins, R., L.; Mallamo, J. P.; Underiner, T., L.; Tripathy, R. Bridged indenopyrrolocarbazoles 99/62523, 1999, 1999.

205. Sato, B.; Nakajima, H.; Hori, Y.; Hino, M.; Hashimoto, S.; Terano, H., A new antimitotic substance, FR182877. II. The mechanism of action. *J. Antibiot.* **2000**, *53*, 204-206.

206. Yoshimura, S.; Sato, B.; Kinoshita, T.; Takase, S.; Terano, H., A new antimitotic substance, FR182877. III. Structure determination. *J. Antibiot.* **2000**, *53*, 615-622.

207. Buey, R. M., et al., Cyclostreptin binds covalently to microtubule pores and lumenal taxoid binding sites. *Nat. Chem. Bio.* **2007**, *3*, 117.

208. ChemOffice. http://www.cambridgesoft.com/software/overview.aspx.

209. MarvinSketch 5.11.4. chemaxon.com.

210. ChemSketch version 2018.2.1. https://www.acdlabs.com/resources/freeware/chemsketch/.

211. Muller, P., Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994). *Pure Appl. Chem.* **1994**, *66*, 1077-1184.

212. Ashenhurst, J. Naming Bicyclic Compounds – Fused, Bridged, and Spiro. https://www.masterorganicchemistry.com/2014/08/14/bridged-bicyclic-compounds-and-how-to-name-them/.

213. Mak, J. Y. W.; Pouwer, R. H.; Williams, C. M., Natural Products with Anti-Bredt and Bridgehead Double Bonds. *Angew. Chem., Int. Ed.* **2014**, *53*, 13664-13688.

214. Kutateladze, A. G.; Krenske, E. H.; Williams, C. M., Reassignments and Corroborations of Oxo-Bridged Natural Products Directed by OSE and DU8+ NMR Computation. *Angew. Chem., Int. Ed.* **2019**, *58*, 7107-7112.

215. Alder, R. W.; East, S. P., In/Out Isomerism. *Chem. Rev.* **1996**, *96*, 2097-2112.

216. OpenSMILES. http://opensmiles.org/.

217. The IUPAC International Chemical Identifier (InChI^TM). https://iupac.org/who-we-are/divisions/division-details/inchi/.

218. Segler, M. H. S.; Preuss, M.; Waller, M. P., Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **2018**, *555*, 604-610.

219. Green, C. P.; Engkvist, O.; Pairaudeau, G., The convergence of artificial intelligence and chemistry for improved drug discovery. *Future Med. Chem.* **2018**, *10*, 2573-2576.

220. Wilson, E. B.; Decius, J. C.; Cross, P. C., *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*; McGraw-Hill Book Company: New York, 1955.

221. Reimers, J. R., A practical method for the use of curvilinear coordinates in calculations of normal-mode projected displacements and Duschinsky rotation matrices for large molecules. *J. Chem. Phys.* **2001**, *115*, 9103-9.

222. Darrow, J. J., The Patentability of Enantiomers: Implications for the Pharmaceutical Industry. *Stanford Technology Law Review* **2007**, *2*.

223. Development of new stereoisomeric drugs. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/development-new-stereoisomeric-drugs.

224. Investigation of chiral active substances. https://www.ema.europa.eu/en/investigation-chiral-active-substances#current-effective-version-section).

225. Nicolaou, K. C.; Hale, C. R. H.; Nilewski, C.; Ioannidou, H. A., Constructing molecular complexity and diversity: Total synthesis of natural products of biological and medicinal importance. *Chem. Soc. Rev.* **2012**, *41*, 5185-5238.

226. Frobenius, G., Ueber die Congruenz nach einem aus zwei endlichen Gruppen gebildeten Doppelmodul. *Journal für die Reine und Angewandte Mathematik* **1886**, *101*, 273-299.

227. Alspach, B.; Aronoff, S., Enumeration of structural isomers in alicyclic hydrocarbons and in porphyrins. *Can. J. Chem.* **1977**, *55*, 2773-2777.

228. Balaban, A. T., Applications of graph theory in chemistry. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 334-343.

229. Trinajstić, N.; Klein, D. J.; Randić, M., On some solved and unsolved problems of chemical graph theory. *Int. J. Quantum Chem.* **1986**, *30*, 699-742.

230. Slanina, Z., An Interplay Between The Phenomenon Of Chemical Isomerism And Symmetry Requirements: A Perennial Source Of Stimuli For Molecular-Structure Concepts, As Well As For Algebraic And Computational Chemistry Dedicated to the 50th anniversary of the publishing of Pólya's Hauptsatz. In *Symmetry*, Hargittai, I., Ed. Pergamon: 1986; pp 585-616.

231. Faulon, J.-L.; Bender, A., *Handbook of Chemoinformatics Algorithms*; Chapman and Hall/CRC: London, 2010.

232. Faulon, J.-L.; Churchwell, C. J.; Visco, D. P., The Signature Molecular Descriptor. 2. Enumerating Molecules from Their Extended Valence Sequences. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 721-734.

233. Visco, D. P., Jr.; Faulon, J.-L. M.; Roe, D. C. Enumerating molecules. https://doi.org/10.2172/918764.

234. Balaban, A. T., Chemical graphs. XXV. Isomers of the porphyrins. Exercise using Polya's counting theorem. *Rev. Roum. Chim.* **1975**, *20*, 227-232.

235. Ndoyom, V.; Fusaro, L.; Dorcet, V.; Boitrel, B.; Le Gac, S., Sunlight-Driven Formation and Dissociation of a Dynamic Mixed-Valence Thallium(III)/Thallium(I) Porphyrin Complex. *Angew. Chem., Int. Ed.* **2015**, *54*, 3806-3811.

236. Liu, D.; Ferrence, G. M.; Lash, T. D., Oxybenziporphyrins, Oxypyriporphyrins, Benzocarbaporphyrins, and Their 23-Oxa and 23-Thia Analogues: Synthesis, Spectroscopic Characterization, Metalation, and Structural Characterization of a Palladium(II) Organometallic Derivative. *J. Org. Chem.* **2004**, *69*, 6079-6093.

237. Jain, P.; Ferrence, G. M.; Lash, T. D., Preparation of Furan and Thiophene-Derived Fulvene Dialdehydes: Synthesis and Structural Characterization of a 22-Oxa-21-carbaporphyrin and a Related Palladium(II) Organometallic Complex. *J. Org. Chem.* **2010**, *75*, 6563-6573.

238. Smolczyk, T. J.; Lash, T. D., Alphabet soup within a porphyrinoid cavity: synthesis of heterocarbaporphyrins with CNNO, CNOO, CNSO and CNSeO Cores from an oxacarbatripyrrin. *Chem. Commun.* **2018**, *54*, 9003-9006.

239. Abedian, N.; Dehghani, H., Novel molecular complexation between meso-tetraarylporphyrinato magnesium(II) and phosphorus(III) chloride. *Inorg. Chem. Commun.* **2013**, *36*, 77-80.

240. Berezin, D. B.; Shukhto, O. V.; Reshetyan, M. S., Metal exchange reactions in cadmium complexes with porphyrins of various structures. *Russ. J. Gen. Chem.* **2010**, *80*, 518-526.

241.   Berezin, D. B.; Lazareva, N. V.; Shukhto, O. V.; Kumeev, R. S.; Semeikin, A. S., Metal exchange in cadmium porphyrins with spatially shielded coordination centers. *Russ. J. Inorg. Chem.* **2013**, *58*, 1052-1057.

242.   Tsutsui, M.; Hrung, C. P.; Ostfeld, D.; Srivastava, T. S.; Cullen, D. L.; Meyer, E. F., Unusual metalloporphyrins. XXIII. Unusual metalloporphyrin complexes of rhenium and technetium. *J. Am. Chem. Soc.* **1975**, *97*, 3952-3965.

243.   Johnson, A. W., Condensation of ethyl diazoacetate with cobalt porphyrins. *J. Chem. Soc. Perkin Trans. I* **1975**, 2076-2085.

244.   Lai, J.-J.; Khademi, S.; Meyer, E. F.; Cullen, D. L.; Smith, K. M., Bis-thallium(I) porphyrin complexes. *J. Porphyrins Phthalocyanines* **2001**, *05*, 621-627.

245.   Callot, H. J.; Cromer, R.; Louati, A.; Gross, M., Synthesis of N,N′-phenyleneporphyrines. Models for cytochrome P-450–1-aminobenzotriazole inactivation products. *J. Chem. Soc., Chem. Commun.* **1986**, 767-769.

246.   Guldi, D. M.; Neta, P.; Hambright, P.; Rahimi, R., Ring reduction of [N-methyltetrakis(4-sulfonatophenyl)porphinato]cobalt(II), -nickel(II), and -copper(II) and subsequent methyl group migration. Reversible reaction between methyl radicals and Ni(II)TSPP. *Inorg. Chem.* **1992**, *31*, 4849-4853.

247.   Sparks, L. D.; Chamberlain, J. R.; Hsu, P.; Ondrias, M. R.; Swanson, B. A.; Ortiz de Montellano, P. R.; Shelnutt, J. A., Raman spectroscopic characterization of isomers of copper and zinc N-phenylprotoporphyrin IX dimethyl esters. *Inorg. Chem.* **1993**, *32*, 3153-3161.

248.   Kuila, D.; Kopelove, A. B.; Lavallee, D. K., Effect of N-substituents and axial ligands on reduction potentials of N-substituted metalloporphyrins. *Inorg. Chem.* **1985**, *24*, 1443-1446.

249.   Wang, M.-C.; Sue, L.-S.; Liau, B.-C.; Ko, B.-T.; Elango, S.; Chen, J.-H., Mercury Complexes of meso-Tetra-(p-cyanophenyl)porphyrin and N-methylporphyrin: meso-Tetra(p-cyanophenyl)porphyrinatomercury(II) and Chloro(N-methyl-meso-tetraphenylporphyrinato)mercury(II). *Inorg. Chem.* **2001**, *40*, 6064-6068.

250.   Ogoshi, H.; Setsune, J.-I.; Yoshida, Z.-I., Syntheses and reactions of iridium complexes of octaethylporphyrin. *J. Organomet. Chem.* **1978**, *159*, 317-328.

251.   Latos-Grazynski, L., Proton nuclear magnetic resonance characterization of chloro(N-methyl-5,10,15,20-tetraphenylporphyrinato)nickel(II) and chloro(N-methyloctaethylporphyrinato)nickel(II) complexes. *Inorg. Chem.* **1985**, *24*, 1681-1686.

252.   Thu, H.-Y.; Tong, G. S.-M.; Huang, J.-S.; Chan, S. L.-F.; Deng, Q.-H.; Che, C.-M., Highly Selective Metal Catalysts for Intermolecular Carbenoid Insertion into Primary C-H Bonds and Enantioselective C-C Bond Formation. *Angew. Chem., Int. Ed.* **2008**, *47*, 9747-9751.

253.   Seyler, J. W.; Leidner, C. R., Redox and reaction chemistry of (octaethylporphyrinato) ruthenium-phenyl complexes. *Inorg. Chem.* **1990**, *29*, 3636-3641.

254.   Berezin, D. B.; Txao, V. T.; Azorina, A. A.; Shukhto, O. V.; Guseinov, S. S.; Berezina, N. M., Synthesis, chemical stability, and electrocatalytic properties of zinc(II) and cobalt(II) complexes of N-methyltetraphenylporphine. *Russ. J. Inorg. Chem.* **2015**, *60*, 1267-1274.

255.   Tsukahara, K.; Ito, K.; Matsui, M.; Sudo, M.; Yoshiaki, O., Photophysical properties of N-alkylporphyrins and their metalloporphyrins linked with quinolinium ion in solution. *Chem. Phys. Lett.* **1997**, *281*, 261-266.

256.   Novikova, N. I.; Lo, A. S. V.; Gordon, K. C.; Brothers, P. J.; Simpson, M. C., Diboron Porphyrins: The Raman Signature of the In-Plane Tetragonal Elongation of the Macrocycle. *J. Phys. Chem. A* **2018**, *122*, 5121-5131.

257.   Bartczak, T. J., The structure of cis-dichloro(25-ethoxycarbonyl-2,3,7,8,12,13,17,18-octaethyl-23,24-methano-23H,24H-porphyrin-N21,N22,H)palladium(II) chloroform solvate, [PdCl2(C40H50N4O2)].CHCl3. *Acta Crystallogr. Sect. C* **1985**, *41*, 865-869.

258.   Srinivasan, A.; Toganoh, M.; Niino, T.; Osuka, A.; Furuta, H., Synthesis of N-Confused Tetraphenylporphyrin Rhodium Complexes Having Versatile Metal Oxidation States. *Inorg. Chem.* **2008**, *47*, 11305-11313.

259.   Takenaka, A.; Sasada, Y.; Omura, T.; Ogoshi, H.; Yoshida, Z.-I., Crystal structure of μ-octaethylporphinato-bis[dicarbonylrhodium(I)]. *J. Chem. Soc., Chem. Commun.* **1973**, 792-793.

260.   Dawson, D. Y.; Arnold, J., Ionic versus Covalent Bonding in Dilithium Porphyrins: X-ray Structure of Dilithium Tetraphenylporphyrin Bis(etherate). *J. Porphyrins Phthalocyanines* **1997**, *1*, 121-124.

261.   Brand, H.; Capriotti, J. A.; Arnold, J., New Lithium Porphyrin Derivatives: Synthesis of Li2(P)(Et2O)2 (P = TTP, TBPP) and Solution Structure of Li2(TTP)(Et2O)2 by ^{7}Li and ^{15}N NMR. *Inorg. Chem.* **1994**, *33*, 4334-4337.

262.   Arnold, J.; Dawson, D. Y.; Hoffman, C. G., Synthesis and characterization of lithium, sodium, and potassium porphyrin complexes. X-ray crystal structures of Li2(C6H12O2)2TMPP, Na2(THF)4OEP, and K2(pyridine)4OEP. *J. Am. Chem. Soc.* **1993**, *115*, 2707-2713.

263.   Bonomo, L.; Lehaire, M.-L.; Solari, E.; Scopelliti, R.; Floriani, C., The First Crystalline Calcium Porphyrin and Tetrakis(tert-butylphenyl)porphyrinato Calcium(II): Its Synthesis, Structure, and Binding Properties Towards Alkali and Alkaline Earth Metal Salts. *Angew. Chem., Int. Ed.* **2001**, *40*, 771-774.

264.   Dang, T. T.; Durot, S.; Monnereau, L.; Heitz, V.; Barbieri, A.; Ventura, B., Highlight on the solution processes occurring on silver(I)-assembling porphyrins in the presence of an excess of silver salt. *Dalton Trans.* **2017**, *46*, 9375-9381.

265.   Harrach, G.; Valicsek, Z.; Horváth, O., Water-soluble silver(II) and gold(III) porphyrins: The effect of structural distortion on the photophysical and photochemical behavior. *Inorg. Chem. Commun.* **2011**, *14*, 1756-1761.

266.   Valicsek, Z.; Lendvay, G.; Horváth, O., Equilibrium, photophysical, photochemical and quantum chemical examination of anionic mercury(I) porphyrins. *J. Porphyrins Phthalocyanines* **2009**, *13*, 910-926.

267.   Thomas, K. E.; McCormick, L. J.; Vazquez-Lima, H.; Ghosh, A., Stabilization and Structure of the cis Tautomer of a Free-Base Porphyrin. *Angew. Chem., Int. Ed.* **2017**, *56*, 10088-10092.

268.   Thomassen, I. K.; McCormick, L. J.; Ghosh, A., Molecular Structure of a Free-Base β-Octaiodo-meso-tetraarylporphyrin. A Rational Route to cis Porphyrin Tautomers? *Cryst. Growth Des.* **2018**, *18*, 4257-4259.

269.   Thomas, K. E.; Slebodnick, C.; Ghosh, A., Facile Supramolecular Engineering of Porphyrin cis Tautomers: The Case of β-Octabromo-meso-tetraarylporphyrins. *ACS Omega* **2020**, *5*, 8893-8901.

270.   Crossley, M. J.; Field, L. D.; Harding, M. M.; Sternhell, S., Kinetics of tautomerism in 2-substituted 5,10,15,20-tetraphenylporphyrins: directionality of proton transfer between the inner nitrogens. *J. Am. Chem. Soc.* **1987**, *109*, 2335.

271.   Aronoff, S.; Weast, C. A., Spectra of porphyrins and their acid salts. *J. Org. Chem.* **1941**, *06*, 550-557.

272.   Hammes-Schiffer, S., Virtual Issue on Proton Transfer. *J. Phys. Chem. B* **2021**, *125*, 3725-3726.

273.   Storm, C. B.; Teklu, Y., Nitrogen-hydrogen tautomerism in porphyrines and chlorines. *J. Am. Chem. Soc.* **1972**, *94*, 1745-1747.

274.   Crossley, M. J.; Harding, M. M.; Sternhell, S., Tautomerism in 2-substituted 5,10,15,20-tetraphenylporphyrins. *J. Am. Chem. Soc.* **1986**, *108*, 3608-13.

275.   Merz, K. M.; Reynolds, C. H., Tautomerism in free base porphyrins: The porphyrin potential energy surface. *J. Chem. Soc., Chem. Commun.* **1988**, 90-92.

276.   Reimers, J. R.; Lü, T. X.; Crossley, M. J.; Hush, N. S., The mechanism for inner-hydrogen migration in free base porphin: *Ab initio* MP2 calculations. *J. Am. Chem. Soc.* **1995**, *117*, 2855.

277.   Boronat, M.; Ortí, E.; Viruela, P. M.; Tomás, F., Theoretical study of the N-H tautomerism in free base porphyrin. *J. Mol. Struct.: THEOCHEM* **1997**, *390*, 149-156.

278.   Baker, J.; Kozlowski, P. M.; Jarzecki, A. A.; Pulay, P., The inner-hydrogen migration in free base porphyrin. *Theor. Chem. Acc.* **1997**, *97*, 59-66.

279.   Maity, D. K.; Truong, T. N., Status of theoretical modeling of tautomerization in free-base porphyrin. *J. Porphyrins Phthalocyanines* **2001**, *5*, 289-299.

280.   Gil, M.; Waluk, J., Vibrational Gating of Double Hydrogen Tunneling in Porphycene. *J. Am. Chem. Soc.* **2007**, *129*, 1335-1341.

281.   Kadish, K. M.; E, W.; Zhan, R.; Khoury, T.; Govenlock, L. J.; Prashar, J. K.; Sintic, P. J.; Ohkubo, K.; Fukuzumi, S.; Crossley, M. J., Porphyrin-Diones and Porphyrin-Tetraones: Reversible Redox Units Being Localized within the Porphyrin Macrocycle and Their Effect on Tautomerism. *J. Am. Chem. Soc.* **2007**, *129*, 6576-6588.

282.   Maity, D. K.; Bell, R. L.; Truong, T. N., Mechanism and Quantum Mechanical Tunneling Effects on Inner Hydrogen Atom Transfer in Free Base Porphyrin: A Direct ab Initio Dynamics Study. *J. Am. Chem. Soc.* **2000**, *122*, 897-906.

283.   Goerigk, L.; Grimme, S., A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670-6688.

284.   Grimme, S., Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

285.   Dunning, T. H., Jr., Gaussian basis sets for use in correlated molecular calculations I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007.

286.   Kirsch, P.; Hartman, A. M.; Hirsch, A. K. H.; Empting, M., Concepts and Core Principles of Fragment-Based Drug Design. *Molecules* **2019**, *24*.

287.   Gruner, S. A. W.; Locardi, E.; Lohof, E.; Kessler, H., Carbohydrate-Based Mimetics in Drug Design: Sugar Amino Acids and Carbohydrate Scaffolds. *Chem. Rev.* **2002**, *102*, 491-514.

288.   Zheng, Y.-J.; Tice, C. M., The utilization of spirocyclic scaffolds in novel drug discovery. *Expert Opinion on Drug Discovery* **2016**, *11*, 831-834.

289.   Nelson, P.; Untch, K. Synthesis of (10)-annulenes. 3708547, 1973.

290.   Nelson, P. H.; Untch, K. G.; Fried, J. H. 2-(2-Hydroxy-3-aminopropyl-1-yloxy)-1,6-difluoro-and-1,6-methano-[10]annulenes and the salts thereof. 3758583, 1973.

291.   Liu, C. C.; Schultz, P. G., Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **2010**, *79*, 413-444.

292.   Stapleton, M. P., Sir James Black and propranolol. The role of the basic sciences in the history of cardiovascular pharmacology. *Tex. Heart Inst. J.* **1997**, *24*, 336-42.

293.   Moukhametzianov, R.; Warne, T.; Edwards, P. C.; Serrano-Vega, M. J.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X., Two distinct conformations of helix 6 observed in antagonist-bound structures of a β1-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 8228-8232, S8228/1-S8228/5.

294.   Berman, H. M.; Henrick, K.; Nakamura, H., Announcing the world wide Protein Data Bank. *Nat. Struct. Biol.* **2003**, *10*, 980.

295. Kumar, K. K., et al., Structure of a Signaling Cannabinoid Receptor 1-G Protein Complex. *Cell* **2019**, *176*, 448-458.e12.

296. Wang, H.-L., et al., Discovery of (*R*)-8-(6-Methyl-4-oxo-1,4,5,6-tetrahydropyrrolo[3,4-*b*]pyrrol-2-yl)-3-(1-methylcyclopropyl)-2-((1-methylcyclopropyl)amino)quinazolin-4(3*H*)-one, a potent and selective PIM-1/2 kinase inhibitor for hematological malignancies. *J. Med. Chem.* **2019**, *62*, 1523-1540.

297. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W., Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221-234.

298. Schrödinger, *Maestro*; Schrödinger: New York, 2019.

299. Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A., A hierarchical approach to all-atom protein loop prediction. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 351-367.

300. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L., OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* **2015**, *12*, 281-296.

301. Schrödinger, *LigPrep*; Schrödinger: New York, 2019.

302. Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A., Jaguar: a high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quantum Chem.* **2013**, *113*, 2110-2142.

303. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177-6196.

304. Lyne, P. D.; Lamb, M. L.; Saeh, J. C., Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J. Med. Chem.* **2006**, *49*, 4805-4808.

305. Guengerich, F. P., Cytochrome P450 and Chemical Toxicology. *Chem. Res. Toxicol.* **2008**, *21*, 70-83.

306. London, F., Zur Theorie und Systematik der Molekularkräfte. *Zeitschrift für Physik* **1930**, *63*, 245-279.

307. Scharf, J.; Schlögl, K.; Widhalm, M.; Lex, J.; Tückmantel, W.; Vogel, E.; Pertlik, F., Stereochemistry of planarchiral compounds, part X - X-ray crystal structures and absolute chiralities of bridged [10]- and [14]anulenes. *Monatshefte für Chemie Chemical Monthly* **1986**, *117*, 255-267.

308. Vogel, E.; Schieb, T.; Schulz, W. H.; Schmidt, K.; Schmickler, H.; Lex, J., Bridged [14]Annulenes with a Phenanthrene-Perimeter:anti-1,6:7,12-Bismethano[14]annulene. *Angew. Chem., Int. Ed.* **1986**, *25*, 723-725.

309. Vogel, E.; Kürschner, U.; Schmickler, H.; Lex, J.; Wennerström, O.; Tanner, D.; Norinder, U.; Krüger, C., 1,6:9,14-bismethano[16]annulene - a new bridged [4n]annulene. *Tetrahedron Lett.* **1985**, *26*, 3087-3090.

310. Vogel, E.; Sombroek, J.; Wagemann, W., syn-1,6:8,13-Bismethano[14]annulene. *Angew. Chem., Int. Ed.* **1975**, *14*, 564-565.

311. Vogel, E.; Haberland, U.; Günther, H., anti-1,6:8,13-Bismethano[14]annulene. *Angew. Chem., Int. Ed.* **1970**, *9*, 513-514.

312. Vogel, E.; Sturm, W.; Cremer, H.-D., 1,6:8,13-Butanediylidene[14]annulene. *Angew. Chem., Int. Ed.* **1970**, *9*, 516-517.

313. Vogel, E.; Haberland, U.; Ick, J., syn-1,6-Methano-8,13-oxido[14]annulene. *Angew. Chem., Int. Ed.* **1970**, *9*, 517-518.

314.   Vogel, E.; Roth, H. D., The Cyclodecapentaene System. *Angew. Chem., Int. Ed.* **1964**, *3*, 228-229.
315.   Boys, S. F.; Benardi, F., *Mol. Phys.* **1970**, *19*, 553.
316.   Boys, S. F.; Bernardi, F., (counterpoise correction method). *Mol. Phys.* **1970**, *19*, 553.
317.   Cohen, E. R., *Quantities, units, and symbols in physical chemistry*, 3rd ed. ed.; RSC Pub.: Cambridge, UK, 2007.

Appendix A: IUPAC "Gold Book"
and "Green Book" entries

Sections 0 to 0 are formal definitions, copied *verbatim*, from the IUPAC Compendium of Chemical Terminology (the "Gold Book")[1] with Section A.52 listing some approved abbreviations from the IUPAC Compendium of Quantities, Units and Symbols in Physical Chemistry (the "Green Book").[317]

A.1 atropisomers

https://doi.org/10.1351/goldbook.A00511

A subclass of conformers which can be isolated as separate chemical species and which arise from restricted rotation about a single bond, *e.g. ortho*-substituted biphenyl, 1,1,2,2-tetra-*tert*-butylethane.
Source:
PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2200

A.2 bond

https://doi.org/10.1351/goldbook.B00697

There is a chemical bond between two atoms or groups of atoms in the case that the forces acting between them are such as to lead to the formation of an aggregate with sufficient stability to make it convenient for the chemist to consider it as an independent 'molecular species'.
Source:
PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1089

A.3 chemical species

https://doi.org/10.1351/goldbook.CT01038

An ensemble of chemically identical molecular entities that can explore the same set of molecular energy levels on the time scale of the experiment. The term is applied equally to a set of chemically identical atomic or molecular structural units in a solid array. For example, two conformational isomers may be interconverted sufficiently slowly to be detectable by separate NMR spectra and hence to be considered to be separate chemical species on a time scale governed by the radiofrequency of the spectrometer used. On the other hand, in a slow chemical reaction the same mixture of conformers may behave as a single chemical species, i.e. there is virtually complete equilibrium population of the total set of molecular energy levels belonging to the two conformers. Except where the context

requires otherwise, the term is taken to refer to a set of molecular entities containing isotopes in their natural abundance. The wording of the definition given in the first paragraph is intended to embrace both cases such as graphite, sodium chloride or a surface oxide, where the basic structural units may not be capable of isolated existence, as well as those cases where they are. In common chemical usage generic and specific chemical names (such as radical or hydroxide ion) or chemical formulae refer either to a chemical species or to a molecular entity.

Sources:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1096

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2202

A.4   chirality centre – synonym: centre of chirality

https://doi.org/10.1351/goldbook.C01060

An atom holding a set of ligands in a spatial arrangement which is not superposable on its mirror image. A chirality centre is thus a generalized extension of the concept of the asymmetric carbon atom to central atoms of any element, for example N^+abcd, Pabc as well as Cabcd.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2203

A.5   configuration (stereochemical)

https://doi.org/10.1351/goldbook.C01249

In the context of stereochemistry, the term is restricted to the arrangements of atoms of a molecular entity in space that distinguishes stereoisomers, the isomerism between which is not due to conformational differences.

Sources:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1099

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2204

A.6   conformation

https://doi.org/10.1351/goldbook.C01258

The spatial arrangement of the atoms affording distinction between stereoisomers which can be interconverted by rotations about formally single bonds. Some authorities extend the term to include inversion at trigonal pyramidal centres and other polytopal rearrangements.

Sources:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC*

Recommendations 1994)) on page 1099

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2204

A.7   conformer

https://doi.org/10.1351/goldbook.C01262

One of a set of stereoisomers, each of which is characterized by a conformation corresponding to a distinct potential energy minimum.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2204

A.8   connectivity

https://doi.org/10.1351/goldbook.C01274

In a chemical context, the information content of a line formula, but omitting any indication of bond multiplicity.

Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1099

A.9   constitution

https://doi.org/10.1351/goldbook.C01282

The description of the identity and connectivity (and corresponding bond multiplicities) of the atoms in a molecular entity (omitting any distinction arising from their spatial arrangement).

Sources:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1100

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2204

A.10   constitutional isomerism

https://doi.org/10.1351/goldbook.C01285

Isomerism between structures differing in constitution and described by different line formulae *e.g.* CH_3OCH_3 and CH_3CH_2OH.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2205

A.11   contributing structure – Also contains definition of: canonical form

https://doi.org/10.1351/goldbook.C01309

The definition is based on the valence-bond formulation of the quantum mechanical idea of the wavefunction of a molecule as composed of a linear combination of wavefunctions, each representative of a formula containing bonds that are only single, double or triple with a particular pairing of electron spins. Each such formula represents a contributing structure, also called 'resonance structure' to the total wavefunction, and the degree to which each contributes is indicated by the square of its coefficient in the linear combination. The contributing structures, also called 'canonical forms', themselves thus have a purely formal significance: they are the components from which wavefunctions can be built. Structures may be covalent (or non-polar) or ionic (or polar). The representation is frequently kept qualitative so that we speak of important or major contributing structures and minor contributing structures. For example, two major non-equivalent contributing structures for the conjugate base of acetone are:

Source:

PAC, 1994, 66, 1077. (Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)) on page 1100

A.12   degenerate rearrangement

https://doi.org/10.1351/goldbook.D01559

A molecular rearrangement in which the principal product is indistinguishable (in the absence of isotopic labelling) from the principal reactant. The term includes both 'degenerate intramolecular rearrangements' and reactions that involve intermolecular transfer of atoms or groups ('degenerate intermolecular rearrangements'): both are degenerate isomerizations. The occurrence of degenerate rearrangements may be detectable by isotopic labelling or by dynamic NMR techniques. For example: the [3,3]sigmatropic rearrangement of hexa-1,5-diene (Cope rearrangement):

Synonymous but less preferable terms are 'automerization', 'permutational isomerism', 'isodynamic transformation', 'topomerization'.

Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1104

A.13 diastereoisomerism – Also contains definitions of: diastereoisomers, diastereomers

https://doi.org/10.1351/goldbook.D01679

Stereoisomerism other than enantiomerism. Diastereoisomers (or diastereomers) are stereoisomers not related as mirror images. Diastereoisomers are characterized by differences in physical properties, and by some differences in chemical behaviour towards achiral as well as chiral reagents.

Sources:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1105

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2205

A.14 elementary reaction

https://doi.org/10.1351/goldbook.E02035

A reaction for which no reaction intermediates have been detected or need to be postulated in order to describe the chemical reaction on a molecular scale. An elementary reaction is assumed to occur in a single step and to pass through a single transition state.

Sources:

PAC, 1993, *65*, 2291. (*Nomenclature of kinetic methods of analysis (IUPAC Recommendations 1993)*) on page 2294

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1111

A.15 enantiomer

https://doi.org/10.1351/goldbook.E02069

One of a pair of molecular entities which are mirror images of each other and non-superposable.

Sources:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1112

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2207

A.16 enantiomerism

https://doi.org/10.1351/goldbook.E02078

The isomerism of enantiomers.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2207

A.17 fluxional

https://doi.org/10.1351/goldbook.F02463

A chemical species is said to be fluxional if it undergoes rapid degenerate rearrangement (generally detectable by methods which allow the observation of the behaviour of individual nuclei in a rearranged chemical species, *e.g.* NMR, X-ray). Example: bullvalene (1 209 600 interconvertible arrangements of the ten CH groups).
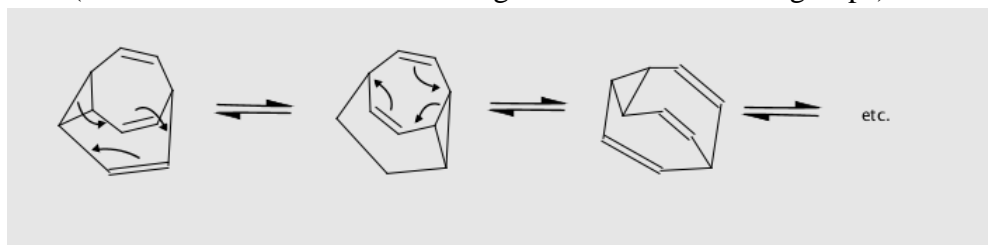
[*sic*]

The term is also used to designate positional change among ligands of complex compounds and organometallics. In these cases, the change is not necessarily degenerate.
Source:
PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1115

A.18 free rotation (hindered rotation, restricted rotation)

https://doi.org/10.1351/goldbook.F02520

In a stereochemical context the rotation about a bond is called 'free' when the rotational barrier is so low that different conformations are not perceptible as different chemical species on the time scale of the experiment. The inhibition of rotation of groups about a bond due to the presence of a sufficiently large rotational barrier to make the phenomenon observable on the time scale of the experiment is termed hindered rotation or restricted rotation.
Source:
PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2209

A.19 hula-twist (HT) mechanism

https://doi.org/10.1351/goldbook.HT07393

Volume-conserving mechanism of photoisomerization of a double bond in a conjugated system involving simultaneous configurational and conformational isomerization, *e.g.*, the photoinitiated concerted rotation of two adjacent double and single bonds.
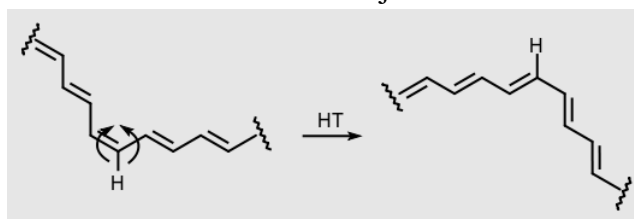
Note:
Under unconstrained conditions, the conventional *one-bond-flip* (OBF) process is the dominant process with the hula-twist (HT) being an undetectable higher energy process. It has been proposed that under confined conditions (*e.g.*, a conjugated double bond *chromophore* in a protein cavity or in a solid matrix), the additional viscosity-dependent barriers make the OBF a less favourable process, allowing the volume-conserving HT to be the dominant process for photoisomerization.
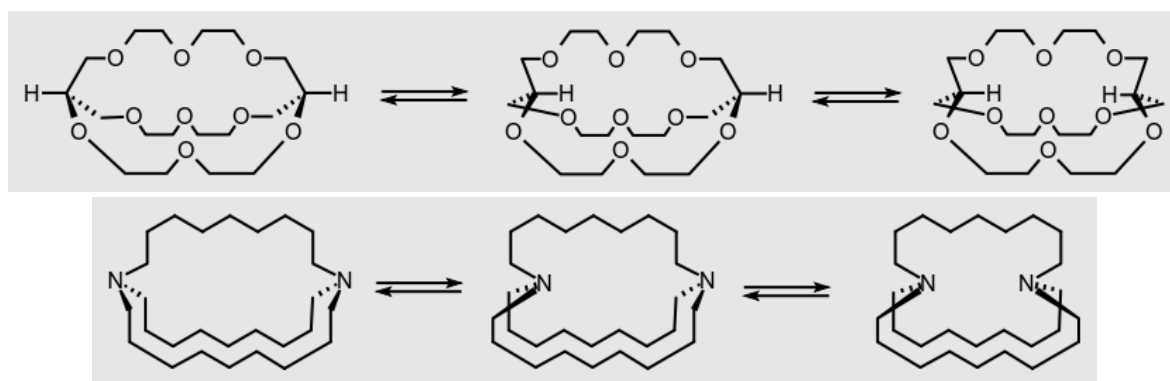
Source:
PAC, 2007, *79*, 293. (*Glossary of terms used in photochemistry, 3rd edition (IUPAC Recommendations 2006)*) on page 352

A.20   in-out isomerism

https://doi.org/10.1351/goldbook.I03055
    Isomerism found in bicyclic systems having long enough bridges to allow the bridgehead exocyclic bond or lone pair of electrons to point either inside the structure or outside.

Source:
PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2210

A.21   intramolecular

https://doi.org/10.1351/goldbook.I03130
    Descriptive of any process that involves a transfer (of atoms, groups, electrons, etc.) or interactions between different parts of the same molecular entity.
    Relating to a comparison between atoms or groups within the same molecular entity.
Source:
PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1126

A.22   inversion

https://doi.org/10.1351/goldbook.I03146
   A symmetry operation involving a centre of inversion (*i*).

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2210

A.23 isodesmic reaction

https://doi.org/10.1351/goldbook.I03272

A reaction (actual or hypothetical) in which the types of bonds that are made in forming the products are the same as those which are broken in the reactants, *e.g.*
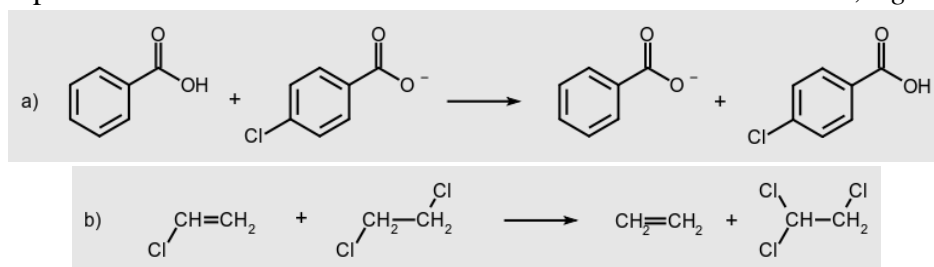
Such processes have advantages for theoretical treatment. The Hammett equation as applied to equilibria [cf. (a)] essentially deals with isodesmic processes.

Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1128

A.24 isomer

https://doi.org/10.1351/goldbook.I03289

One of several chemical species (or molecular entities) that have the same atomic composition (molecular formula) but different line formulae or different stereochemical formulae and hence different physical and/or chemical properties.

Sources:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1129
PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2210

A.25 isomerism

https://doi.org/10.1351/goldbook.I03294

The relationship between isomers.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2210

A.26 isomerization

https://doi.org/10.1351/goldbook.I03295

A chemical reaction, the principal product of which is isomeric with the principal reactant. An intramolecular isomerization that involves the breaking or making of bonds is a special case of a molecular rearrangement. Isomerization does not necessarily imply molecular rearrangement (*e.g.* in the case of the interconversion of conformational isomers).

Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1129

A.27 ligands

https://doi.org/10.1351/goldbook.L03518

In an inorganic coordination entity, the atoms or groups joined to the central atom.

Source:

Red Book, 3rd ed., p. 146

In biochemistry: if it is possible or convenient to regard part of a polyatomic molecular entity as central, then the atoms, groups or molecules bound to that part are called ligands. Biochemical usage is thus wider, in that the central entity can be polyatomic. Thus H^+ may be a ligand for proteins and for citrate as well as for O^{2-} . It may even be a ligand for a univalent entity such as acetate: in other circumstances, AcO^- may be the ligand for H^+, since the definition makes it clear that the view of which entity is central may change for convenience. Thus, four calcium ions are ligands for calmodulin, when the protein is regarded as central: four carboxylate groups of calmodulin ligate (are ligands of) each calcium ion when this ion is regarded as central. It is the ligand that is said to ligate the central entity, which is said to be ligated. When the hormone binding to a receptor is called a ligand, the receptor is thus regarded as the central entity. Biochemists should bear in mind that the usage in inorganic chemistry has been that ligands bind only single atoms, so they should be cautious in fields such as bioinorganic chemistry where confusion may be possible.

Sources:

PAC, 1994, *66*, 1077. (Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)) on page 1136

White Book, 2nd ed., p. 335

A.28 macrocycle

https://doi.org/10.1351/goldbook.M03662

A cyclic macromolecular or a macromolecular cyclic portion of a macromolecule.

Note:

In the literature, the term 'macrocycle' is sometimes used for molecules of low molecular mass that would not be considered 'macromolecules' as specified in the definition given in this book.

Source:
PAC, 1996, *68*, 2287. (*Glossary of basic terms in polymer science (IUPAC Recommendations 1996)*) on page 2298

A.29   molecular entity

https://doi.org/10.1351/goldbook.M03986

Any constitutionally or isotopically distinct atom, molecule, ion, ion pair, radical, radical ion, complex, conformer etc., identifiable as a separately distinguishable entity. Molecular entity is used in this Compendium as a general term for singular entities, irrespective of their nature, while chemical species stands for sets or ensembles of molecular entities. Note that the name of a compound may refer to the respective molecular entity or to the chemical species, *e.g.* methane, may mean a single molecule of CH_4 (molecular entity) or a molar amount, specified or not (chemical species), participating in a reaction. The degree of precision necessary to describe a molecular entity depends on the context. For example 'hydrogen molecule' is an adequate definition of a certain molecular entity for some purposes, whereas for others it is necessary to distinguish the electronic state and/or vibrational state and/or nuclear spin, etc. of the hydrogen molecule.
Source: PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1142

A.30   molecular rearrangement – Also contains definition of: principle of minimum structural change

https://doi.org/10.1351/goldbook.M03997

The term is traditionally applied to any reaction that involves a change of connectivity (sometimes including hydrogen), and violates the so-called 'principle of minimum structural change'. According to this oversimplified principle, chemical species do not isomerization[*sic*] in the course of a transformation, *e.g.* substitution, or the change of a functional group of a chemical species into a different functional group is not expected to involve the making or breaking of more than the minimum number of bonds required to effect that transformation. For example, any new substituents are expected to enter the precise positions previously occupied by displaced groups. The simplest type of rearrangement is an intramolecular reaction in which the product is isomer[*sic*] with the reactant (one type of intramolecular isomerization). An example is the first step of the Claisen rearrangement:
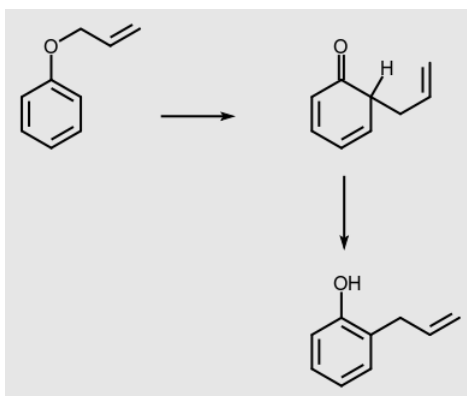
The definition of molecular rearrangement includes changes in which there is a bond migration of an atom or bond (unexpected on the basis of the principle of minimum structural change), as in the reaction:
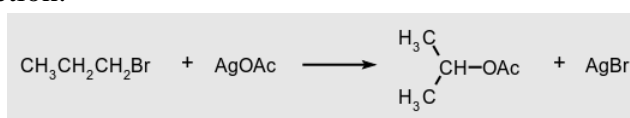

where the rearrangement stage can formally be represented as the '1,2-shift' of hydride between adjacent carbon atoms in the carbocation:


Such migrations occur also in radicals, *e.g.*:
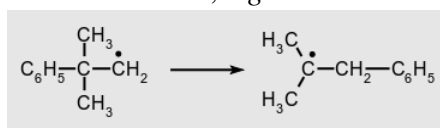
The definition also includes reactions in which an entering group takes up a different position from the leaving group, with accompanying bond migration. An example of the latter type is the 'allylic rearrangement':
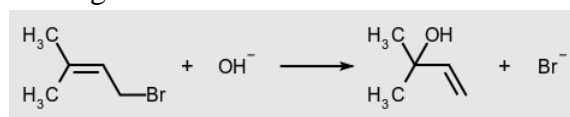
A distinction is made between 'intramolecular rearrangement' (or 'true molecular rearrangements') and 'intermolecular rearrangements' (or 'apparent rearrangements'). In the former case the atoms and groups that are common to a reactant and a product never separate into independent fragments during the rearrangement stage (i.e. the change is intramolecular), whereas in an 'intermolecular rearrangement' a migrating group is completely free from the parent molecule and is re-attached to a different position in a subsequent step, as in the Orton reaction:
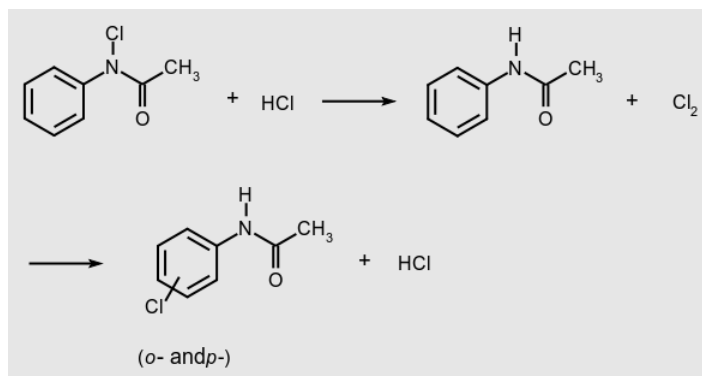
Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1142


A.31   molecule

https://doi.org/10.1351/goldbook.M04002

An electrically neutral entity consisting of more than one atom (n>1). Rigorously, a molecule, in which n>1 must correspond to a depression on the potential energy surface that is deep enough to confine at least one vibrational state.

Source: PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1143

A.32   nonclassical structure

https://doi.org/10.1351/goldbook.NT07084

The structure of molecules or molecular ions that escapes description in terms of conventional rules of valency and stereochemistry. Nonclassical structures are characteristic of carbonium ions with hypercoordinated (see hypercoordination) carbon atoms, *e.g.*, methanium ion **1**, pyramidal dication $C_6H_6^{2+}$ **2** (isomeric to benzene dication), and the molecular species whose structure cannot be adequately represented by the equilibrium (2-norbornyl cation, **3**) or resonance of two or more classical structures. From the stereochemical point of view, those structures are assigned to the nonclassical type for which all tetracoordinate carbon bonds extend into a single hemisphere, i.e., the valence angle of a carbon atom is greater than 180°. A hypothetical example is pyramidane, **4**, the structure of which corresponds to a local minimum on the C_5H_4 potential energy surface.
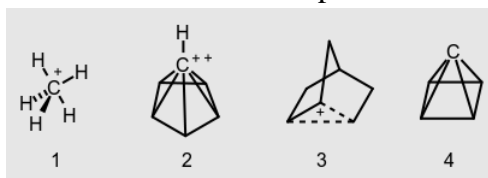
Source:

PAC, 1999, *71*, 1919. (*Glossary of terms used in theoretical organic chemistry*) on page 1955

A.33 polytopal rearrangement

https://doi.org/10.1351/goldbook.P04755

Stereoisomerization interconverting different or equivalent spatial arrangements of ligands about a central atom or of a cage of atoms, where the ligand or cage defines the vertices of a polyhedron. For example pyramidal inversion of amines, Berry pseudorotation of PF_5, rearrangement of polyhedral boranes.
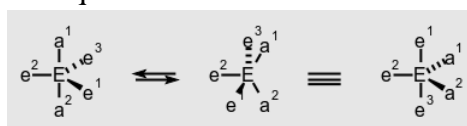
Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2213

A.34 pseudorotation – also contains definitions of: Berry pseudorotation, turnstile rotation

https://doi.org/10.1351/goldbook.P04934

Stereoisomerization resulting in a structure that appears to have been produced by rotation of the entire initial molecule and is superposable on the initial one, unless different positions are distinguished by substitution, including isotopic substitution. One example of pseudorotation is a facile interconversion between the many envelope and twist conformers of a cyclopentane due to the out of plane motion of carbon atoms. Another example of pseudorotation (Berry pseudorotation) is a polytopal rearrangement that provides an intramolecular mechanism for the isomerization of trigonal bipyramidal compounds (*e.g.* λ^5-phosphanes), the five bonds to the central atom E being represented as e^1, e^2, e^3, a^1 and a^2. Two equatorial bonds move apart and become apical bonds at the same time as the apical bonds move together to become equatorial.

A related conformational change of a trigonal bipyramidal structure is described as turnstile rotation. The process may be visualized as follows.
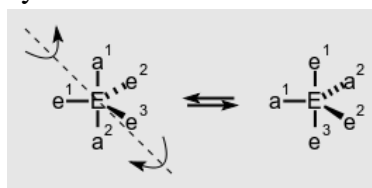
An apical and an equatorial bond rotate as a pair ca. 120° relative to the other three bonds. (Doubts have been expressed about the distinct physical reality of this mechanism.)

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2215

A.35   pyramidal inversion

https://doi.org/10.1351/goldbook.P04956

A polytopal rearrangement in which the change in bond directions to a three-coordinate central atom having a pyramidal arrangement of bonds (tripodal arrangement) causes the central atom (apex of the pyramid) to appear to move to an equivalent position on the other side of the base of the pyramid. If the three ligands to the central atom are different pyramidal inversion interconverts enantiomers.

Source:
PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2215

A.36   rearrangement

https://doi.org/10.1351/goldbook.R05194

See: degenerate rearrangement, molecular rearrangement, sigmatropic rearrangement.
Source:
PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1160

A.37   rotamer

https://doi.org/10.1351/goldbook.R05407

One of a set of conformers arising from restricted rotation about one single bond.

Source:
PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2217

A.38   rotational barrier

https://doi.org/10.1351/goldbook.R05408

In a rotation of groups about a bond, the potential energy barrier between two adjacent minima of the molecular entity as a function of the torsion angle.
*Source:*
PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2217

A.39   sigmatropic rearrangement

https://doi.org/10.1351/goldbook.S05660

A molecular rearrangement that involves both the creation of a new σ-bond between atoms previously not directly linked and the breaking of an existing σ-bond. There is normally a concurrent relocation of π-bonds in the molecule concerned, but the total number

of π- and σ-bonds does not change. The term was originally restricted to intramolecular pericyclic reactions, and many authors use it with this connotation. It is, however, also applied in a more general, purely structural, sense. If such reactions are intramolecular, their transition state may be visualized as an association of two fragments connected at their termini by two partial σ-bonds, one being broken and the other being formed as, for example, the two allyl fragments in (a'). Considering only atoms within the (real or hypothetical) cyclic array undergoing reorganization, if the numbers of these in the two fragments are designated i and j, then the rearrangement is said to be a sigmatropic change of order $[i, j]$ (conventionally $[i] \leq [j]$). Thus the rearrangement (a) is of order [3,3], whilst reaction (b) is a [1,5]sigmatropic shift of hydrogen. (N.B. By convention square brackets [...] here refer to numbers of atoms, in contrast with current usage in the context of cycloaddition.)
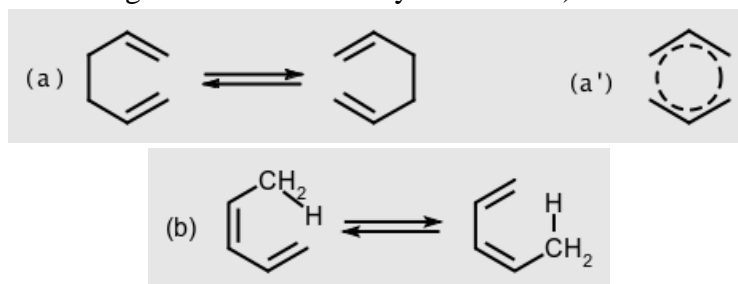
The descriptors a and s (antarafacial and antarafacial [*sic*]) may also be annexed to the numbers i and j; (b) is then described as a [1s,5s] sigmatropic rearrangement, since it is suprafacial with respect both to the hydrogen atom and to the pentadienyl system:
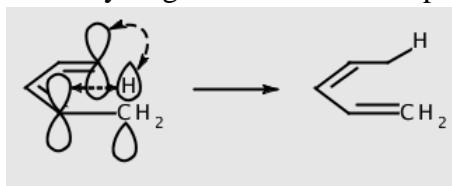
The prefix 'homo' (meaning one extra atom, interrupting conjugation— *cf.* 'homoaromaticity') has frequently been applied to sigmatropic rearrangements, but is misleading.

Source:
PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1163

A.40   stereochemical formula (stereoformula)

https://doi.org/10.1351/goldbook.S05974

A three-dimensional view of a molecule either as such or in a projection.

Source:
PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2218

A.41 stereochemical non-rigidity

https://doi.org/10.1351/goldbook.ST07107

The capability of a molecule to undergo fast and reversible intramolecular isomerization, the energy barrier to which is lower than that allowing for the preparative isolation of the individual isomer at room temperature. It is conventional to assign to It is conventional to assign to the sterochemically [*sic*] non-rigid systems those compounds whose molecules rearrange rapidly enough to influence NMR line shapes at temperatures within the practical range (from −100 °C to 200 °C) of experimentation. The energy barriers to thus defined rearrangements fall into the range of 5–20 kcal/mol (21–85 kJ/mol).

Source:

PAC, 1999, *71*, 1919. (*Glossary of terms used in theoretical organic chemistry*) on page 1964

A.42 stereoisomerism

https://doi.org/10.1351/goldbook.S05983

Isomerism due to differences in the spatial arrangement of atoms without any differences in connectivity or bond multiplicity between the isomers.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2219

A.43 stereoisomers

https://doi.org/10.1351/goldbook.S05984

Isomers that possess identical constitution, but which differ in the arrangement of their atoms in space.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2219

A.44 stereogenic unit (stereogen/stereoelement)

https://doi.org/10.1351/goldbook.S05980

A grouping within a molecular entity that may be considered a focus of stereoisomerism. At least one of these must be present in every enantiomer (though the presence of stereogenic units does not conversely require the corresponding chemical species to be chiral). Three basic types are recognized for molecular entities involving atoms having not more than four substituents:

1. A grouping of atoms consisting of a central atom and distinguishable ligands, such that the interchange of any two of the substituents leads to a stereoisomer. An asymmetric atom (chirality centre) is the traditional example of this stereogenic unit.

2. A chain of four non-coplanar atoms (or rigid groups) in a stable conformation, such that an imaginary or real (restricted) rotation (with a change of sign of the torsion angle) about the central bond leads to a stereoisomer.
3. A grouping of atoms consisting of a double bond with substituents which give rise to cis-trans isomerism.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2219

A.45 stereomutation

https://doi.org/10.1351/goldbook.S05985

A change of configuration at a stereogenic unit brought about by physical or chemical means.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2219

A.46 tautomerism

https://doi.org/10.1351/goldbook.T06252

Isomerism of the general form:

$$G-X-Y=Z \rightleftharpoons X=Y-Z-G$$

where the isomers (called tautomers) are readily interconvertible; the atoms connecting the groups X, Y, Z are typically any of C, H, O or S, and G is a group which becomes an electrofuge or nucleofuge during isomerization. The commonest case, when the electrofuge is H^+, is also known as 'prototropy'. Examples, written so as to illustrate the general pattern given above, include: keto-enol tautomerism, such as:

enol   keto

(G = H, X = O, Y = CCH$_3$, Z = CHCOOEt)

(G = H, X = CHAr, Y = N, Z = CHAr')
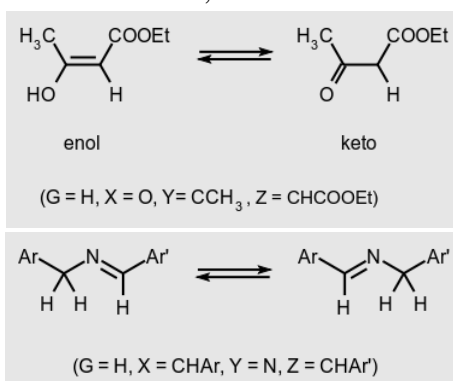
The grouping Y may itself be a three-atom (or five-atom) chain extending the conjugation, as in:

(G = H, X = O, Y = C-CH=CH, Z = N)
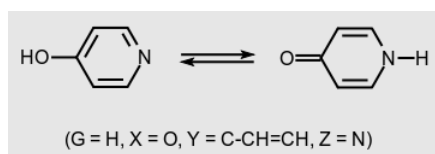
The double bond between Y and Z may be replaced by a ring, when the phenomenon is called ring-chain tautomerism, as in:
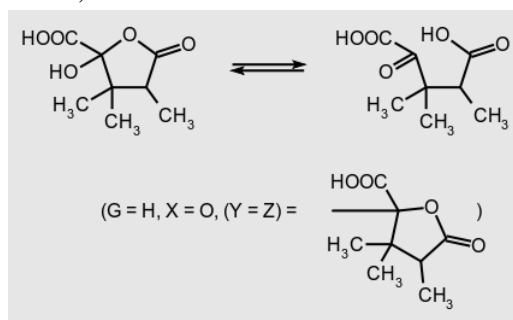
Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1171

A.47 tautomerization

https://doi.org/10.1351/goldbook.T06253

The isomerization by which tautomers are interconverted. It is a heterolytic molecular rearrangement and is frequently very rapid.

Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1172

A.48 torsional stereoisomers

https://doi.org/10.1351/goldbook.T06404

Stereoisomers that can be interconverted (actually or conceptually) by torsion about a bond axis. This includes *E,Z*-isomers of alkene, atropisomers and rotamers.

Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2221

A.49 transition state

https://doi.org/10.1351/goldbook.T06468

In theories describing elementary reactions it is usually assumed that there is a transition state of more positive molar Gibbs energy between the reactants and the products through which an assembly of atoms (initially composing the molecular entities of the reactants) must pass on going from reactants to products in either direction. In the formalism of 'transition state theory' the transition state of an elementary reactions is that set of states (each characterized by its own geometry and energy) in which an assembly of atoms, when randomly placed there, would have an equal probability of forming the reactants or of forming the products of that elementary reactions. The transition state is characterized by one and only one imaginary frequency. The assembly of atoms at the transition state has been called an activated complex. (It is not a complex according to the definition in this

Compendium.) It may be noted that the calculations of reaction rates by the transition state method and based on calculated potential-energy surfaces refer to the potential energy maximum at the saddle point, as this is the only point for which the requisite separability of transition state coordinates may be assumed. The ratio of the number of assemblies of atoms that pass through to the products to the number of those that reach the saddle point from the reactants can be less than unity, and this fraction is the 'transmission coefficient' κ. (There are also reactions, such as the gas-phase colligation of simple radicals, that do not require 'activation' and which therefore do not involve a transition state.)

Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1174

A.50   transition structure

https://doi.org/10.1351/goldbook.T06471

A saddle point on a potential-energy (reaction) surface. It has one negative force constant in the harmonic force constant matrix.

Source:

PAC, 1994, *66*, 1077. (*Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)*) on page 1174

A.51   Walden inversion

https://doi.org/10.1351/goldbook.W06653

Retention of configuration is the preservation of integrity of the spatial arrangement of bonds to a chiral centre during a chemical reaction or transformation. It is also the configurational correlation when a chemical species Xabcd is converted into the chemical species Xabce having the same relative configuration. The configurational change when a chemical species Xabcd (where X is typically carbon), having a tetrahedral arrangement of bonds to X, is converted into the chemical species Xabce having the opposite relative configuration (or when it undergoes an identity reaction in which Xabcd of opposite configuration is produced) is called a Walden inversion or inversion of configuration. The occurrence of a Walden inversion during a chemical transformation is sometimes indicated in the chemical equation by the symbol shown below in place of a simple arrow pointing from reactants to products.
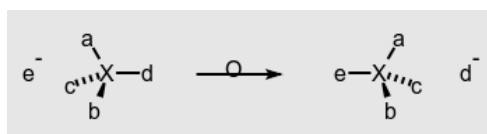
Source:

PAC, 1996, *68*, 2193. (*Basic terminology of stereochemistry (IUPAC Recommendations 1996)*) on page 2221

A.52 Relevant abbreviations listed in IUPAC "Quantities, Units and Symbols in Physical Chemistry", the "Green Book"[2]

E1 elimination unimolecular   [317] page 158

S$_N$1 substitution nucleophilic unimolecular   [317] page 163

S$_N$2 substitution nucleophilic bimolecular   [317] page 163

S$_N$i substitution nucleophilic intramolecular   [317] page 163

References

1 McNaught, A. D.; Wilkinson, A., *IUPAC Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*; Blackwell: Oxford, 1997.
2 Cohen, E. R., *Quantities, units, and symbols in physical chemistry*, 3rd ed. ed.; RSC Pub.: Cambridge, UK, 2007

Appendix B: IUPAC "Red Book" polyhedral symbols and structures.

The following tables are reproduced from the IUPAC Red Book.[1]

Table IR-9.2 *Polyhedral symbols*[a]

Coordination polyhedron	*Coordination number*	*Polyhedral symbol*
linear	2	*L*-2
angular	2	*A*-2
trigonal plane	3	*TP*-3
trigonal pyramid	3	*TPY*-3
T-shape	3	*TS*-3
tetrahedron	4	*T*-4
square plane	4	*SP*-4
square pyramid	4	*SPY*-4
see-saw	4	*SS*-4
trigonal bipyramid	5	*TBPY*-5
square pyramid	5	*SPY*-5
octahedron	6	*OC*-6
trigonal prism	6	*TPR*-6
pentagonal bipyramid	7	*PBPY*-7
octahedron, face monocapped	7	*OCF*-7
trigonal prism, square-face monocapped	7	*TPRS*-7
cube	8	*CU*-8
square antiprism	8	*SAPR*-8
dodecahedron	8	*DD*-8
hexagonal bipyramid	8	*HBPY*-8
octahedron, *trans*-bicapped	8	*OCT*-8
trigonal prism, triangular-face bicapped	8	*TPRT*-8
trigonal prism, square-face bicapped	8	*TPRS*-8
trigonal prism, square-face tricapped	9	*TPRS*-9
heptagonal bipyramid	9	*HBPY*-9

[a] Strictly, not all geometries can be represented by polyhedra.

Table IR-9.3 *Polyhedral symbols, geometrical structures and/or polyhedra*

Three-coordination
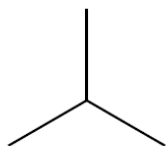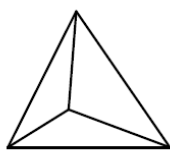
trigonal plane	trigonal pyramid	T-shape

TP-3   *TPY*-3   *TS*-3

Four-coordination

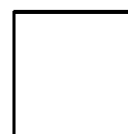tetrahedron   square plane

T-4   *SP*-4
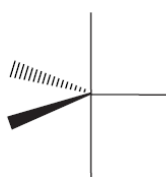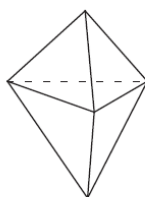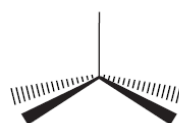
square pyramid   see-saw

SPY-4   *SS*-4

Five-coordination

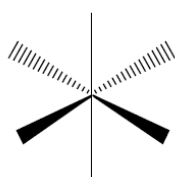trigonal bipyramid   square pyramid

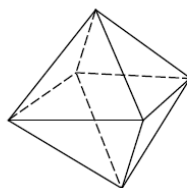*TBPY*-5   *SPY*-5
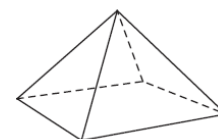
Six-coordination

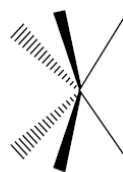octahedron   trigonal prism

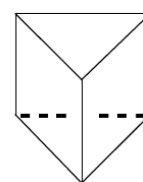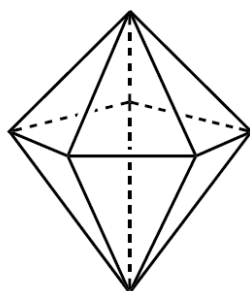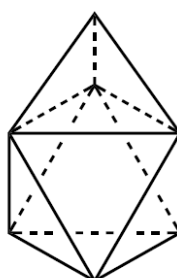*OC*-6   *TPR*-6
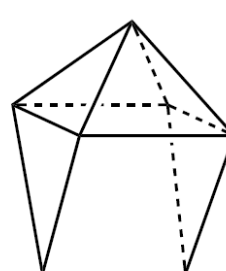
Table IR-9.3 *Continued*

Seven-coordination
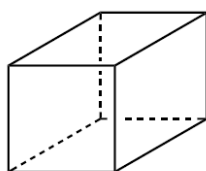
pentagonal
bipyramid

octahedron, face
monocapped

trigonal prism,
square-face monocapped

PBPY-7

OCF-7

TPRS-7

Eight-coordination

cube

square
antiprism

dodecahedron

hexagonal
bipyramid

CU-8

SAPR-8

DD-8

HBPY-8

octahedron,
trans-bicapped

trigonal prism,
triangular-face bicapped

trigonal prism,
square-face bicapped

OCT-8

TPRT-8

TPRS-8

Nine-coordination

trigonal prism,
square-face tricapped

heptagonal
bipyramid

TPRS-9

HBPY-9

1   Connelly, N. G.; Damhus, T.; Hartshorn, R.; Hutton, A. T., *IUPAC Red Book*; RSC Publishing: London, 2005.
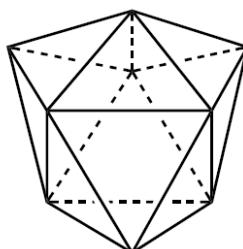
Appendix C: Mathematica Script for Assigning *parvo*/*amplo*

C.1   Summary

Given the atomic coordinates for a pair of akamptisomers, the following script written for Mathematica version 11.3.0.0[1] and implementing the *parvo*/*amplo* definitions as outlined in Sections 4.4.4.

C.2   Mathematica script

Molecular coordinates for akamptisomers are to be provided in the current directory in .xyz format, named "filename1.xyz" and "filename2.xyz".

Two sets of atomic indices, X, Y, Z, Exs, Ezs corresponding to each of these files are to be provided as integers: X, Y, Z and integer lists: Exs and Ezs (see example below). For autoakamptisomeric systems, only one set of coordinates and atomic indices need be provided. The default angle ratio cut-off is set to ⅔ but can be arbitrarily set.

---

1   Wolfram Research, I. *Mathematica*, 11.3; Wolfram Research, Inc.: Champaign, Illinois, 2018.

```
ProjectedElist[coord_,{X_,Y_,Z_,Exs_,Ezs_}]:=
   Module[{Ex,Ez,pv1,pv2,plane,d},
      Ex = Mean[coord[[Exs]]];
      Ez = Mean[coord[[Ezs]]];
      pv1 = coord[[X]] - coord[[Y]];
      pv2 = coord[[Z]] - coord[[Y]];
      plane = Normalize[Cross[pv1,pv2]];
      d = -Total[plane(coord[[Y]])];
      #-((#-coord[[Y]]plane).plane&/@{Ex,Ez}]
VectAngs[coord_,{X_,Y_,Z_,Exs_,Ezs_}]:=
   Module[{Ex,Ez},
      {Ex, Ez} = ProjectedElist[coord,{X,Y,Z,Exs,Ezs}];
      {VectorAngle[coord[[X]]-Ex,Ez-Ex]/Degree,VectorAngle[coord[[Z]]-Ez,Ex-Ez]/Degree}]
AssignPA[n_,cut_]:= "parvo"/;n<=cut;
AssignPA[n_,cut_]:= "amplo"/;n>=1/cut;
AssignPA[n_,cut_]:= Null/;cut<n<1/cut;
AssignPA[n_]:= AssignPA[n,2/3]
ParvoAmplo[filename1_,{X1_,Y1_,Z1_,E1xs_,E1zs_},filename2_,{X2_,Y2_,Z2_,E2xs_,E2zs_},cut_]:=
   Module[{co1,co2,angs1,angs2},
      co1 = Import[filename1,{"VertexCoordinates"}];
      co2 = Import[filename2,{"VertexCoordinates"}];
      angs1 = VectAngs[co1,{X1,Y1,Z1,E1xs,E1zs}];
      angs2 = VectAngs[co2,{X2,Y2,Z2,E2xs,E2zs}];
      {Table[Round[angs1[[i]]/angs2[[i]],0.01],{i,2}],Table[AssignPA[angs1[[i]]/angs2[[i]],cut],{i,2}] }//
      TableForm]
ParvoAmplo[filename1_,{X1_,Y1_,Z1_,E1xs_,E1zs_},filename2_,{X2_,Y2_,Z2_,E2xs_,E2zs_}]:=
   ParvoAmplo[filename1,{X1,Y1,Z1,E1xs,E1zs},filename2,{X2,Y2,Z2,E2xs,E2zs},2/3]
ParvoAmplo[filename_,{X_,Y_,Z_,Exs_,Ezs_}]:=
   ParvoAmplo[filename,{X,Y,Z,Exs,Ezs},filename,{X,Y,Z,Exs,Ezs},2/3]
ParvoAmplo[filename_,{X_,Y_,Z_,Exs_,Ezs_},cut_]:=
   ParvoAmplo[filename,{X,Y,Z,Exs,Ezs},filename,{X,Y,Z,Exs,Ezs,cut]
```
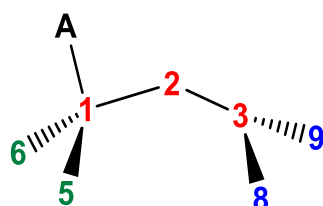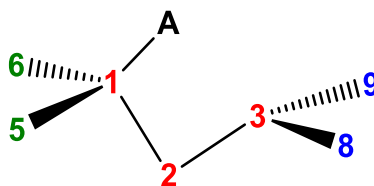
Example

mol   ak_mol

In[1]= ParvoAmplo["mol.xyz",{1,2,3,{5,6},{8,9}},"ak_mol.xyz",{1,2,3,{5,6},{8,9}},2/3]
       ParvoAmplo["ak_mol.xyz",{1,2,3,{5,6},{8,9}},"mol.xyz",{1,2,3,{5,6},{8,9}},2/3]

Out[1]= 10.23   1.
        amplo   Null

Out[2]= 0.10   0.97
        parvo   Null

Appendix D: Reproduced NMR data

Experimental NMR data from "A new fundamental type of conformational isomerism", Peter J. Canfield, Iain M. Blake, Zheng-Li Cai, Ian J. Luck, Elmars Krausz, Rika Kobayashi, Jeffrey R. Reimers & Maxwell J. Crossley, Nature Chemistry volume 10, pages 615–624 (2018).[7]

Table D.1 Experimental ^{1}H chemical shift assignments for racemates (**Fr1** + **Fr2**) and (**Fr3** + **Fr4**) at the temperatures indicated. "Note" indicates the number of H atoms and the signal description.

(Fr1 + Fr2) at 260 K			(Fr3 + Fr4) at 240 K		
[1]H δ / ppm	Assignment	Note [a]	[1]H δ / ppm	Assignment	Note [a]
1.252	*t*-Bu (Ar15)	3, broad	1.373	*t*-Bu (Ar20)	3
1.351	*t*-Bu (Ar20)	3	1.491	*t*-Bu (Ar15)	3, broad
1.458	*t*-Bu (Ar15)	3, broad	1.530	*t*-Bu (Ar5)	3
1.506	*t*-Bu (Ar5)	3	1.539	*t*-Bu (Ar10)	3
1.551	*t*-Bu (Ar10)	3	1.593	*t*-Bu (Ar5)	3
1.554	*t*-Bu (Ar10)	3	1.595	*t*-Bu (Ar10)	3
1.558	*t*-Bu (Ar5)	3	1.618	*t*-Bu (Ar20)	3
1.650	*t*-Bu (Ar20)	3	1.652	*t*-Bu (Ar15)	3, broad
7.75	*o*-Ar15	1, very broad	7.795	*o*-Ar15	1, broad
7.820	*p*-Ar5	1	7.818	*p*-Ar15	1
7.840			7.883	*p*-Ar10	1
7.855	*p*-Ar15, H26,	4	7.889	*p*-Ar5	1
7.870	H27, H28		7.910	H26	1, m
7.875			7.975	H27, H28, *p*-Ar20	3
7.780	*p*-Ar10	1			
7.950	*o*-Ar20	1			
8.030	*p*-Ar20	1	8.036	*o*-Ar5	1
8.100	*o*-Ar5	1	8.068	*o*-Ar20	1
8.110	H29	1, m	8.182	*o*-Ar10	1
8.180	*o*-Ar10	1	8.225	*o*-Ar5	1
8.195	*o*-Ar5	1	8.228	*o*-Ar20	1
8.270	*o*-Ar10	1	8.258	H29	1, m
8.320	*o*-Ar20	1	8.299	*o*-Ar10	1
8.60	*o*-Ar15	1, very broad	8.636	*o*-Ar15	1, broad
8.665	H7	1, d	8.668	H12	1, d
8.720	H8	1, d	8.724	H17	1, d
8.925	H12	1, d	8.819	H13	1, d
9.020	H13	1, d	8.865	H18	1, d
9.035	H17	1, d	8.891	H7	1, d
9.215	H18	1, d	9.048	H8	1, d

a: number of protons, m = multiplet, d = doublet.

Table D.2 Experimental ^{13}C chemical shifts and assignments for the 50 aromatic carbons for each of racemates (**Fr1** + **Fr2**) and (**Fr3** + **Fr4**) at the temperatures indicated. "Note" indicates the number of C atoms and the signal description.

^{13}C δ / ppm	Assignment	Number of C	^{13}C δ / ppm	Assignment	Number of C
	(Fr1 + Fr2) at 260 K			(Fr3 + Fr4) at 240 K	
120.601	C20	1, sharp	116.317	C5	1, sharp
120.830	*p*-Ar20	1	120.484	C20	1, sharp
121.438	*p*-Ar10	1	120.914	p-Ar20	1
121.585	*p*-Ar15	1	121.360	p-Ar(5/10)	1
122.086	*p*-Ar5	1	121.695		1
123.207	C7	1	122.011	p-Ar15	1
123.652	C5/C15	1	124.218	C13	1
123.713	C5/C15	1	124.890	C17	1
125.716	C13	1	125.514	C7	1
126.278	C10	1, sharp	126.366	C10	1, sharp
126.539	C17	1	127.724	o-Ar20	1
128.497	*o*-Ar15	1, very broad	129.243	o-Ar5	1
129.807	*o*-Ar10	1	129.577	o-Ar15	1, very broad
129.834	C(26/27/28)	1	129.677	o-Ar10	1
130.016	*o*-Ar15	1, very broad	129.756	o-Ar(5/20)	1
130.056	*o*-Ar5	1	130.093	C29	1
130.103	C(26/27/28)/C29	1	130.231		1
130.254	C(26/27/28)/C29	1	130.415	C27/C28	1
130.297	*o*-Ar5	1	130.434	C27/C28	1
130.614	*o*-Ar20	1	130.863	C26	1
130.733	C(26/27/28)	1	131.010	o-Ar(5/10/20)	1
131.157	*o*-Ar10	1	131.029	o-Ar15	1, very broad
132.133	*o*-Ar20	1	133.866		1
135.389		1	133.882		1
136.567		1	133.961		1
136.704	C8	1	134.448	C18	1
137.170	C12	1	134.841		1
137.828	C18	1	135.430	C12	1
139.169		1	136.275		1
140.543		1	138.456	C8	1
140.800		1	140.696		1
141.130	C16/C19	1	140.714	C9	1
141.393		1	141.122	C25a/C29a	1
141.535		1	141.293	C25a/C29a	1
141.547	C14	1	144.407	C6	1
141.593		1	145.926		1
143.366		1	146.028		1
145.699	C6/C9	1	146.043		1
146.003	C16/C19	1	146.089	C11/C14	1
146.017		1	146.506	C16	1
146.285		1	146.801	C11/C14	1
146.607		1	147.759		1
147.919		1	148.083		1
148.227		1	148.667		1
148.310		1	149.277		1
149.355		1	149.298		1
149.397		1	149.380		1
149.544		1	149.985	C19	1
149.996	C6/C9	1	150.555		1
150.752		1	151.089		1

Table D.3 Experimental ^{19}F chemical shift assignments for racemates (**Fr1** + **Fr2**) and (**Fr3** + **Fr4**) at the temperatures indicated adopting the *parvo* and *amplo* descriptors. "Note" indicates the number of F atoms and the signal description.

(Fr1 + Fr2) at 270 K				(Fr3 + Fr4) at 240 K		
^{19}F δ / ppm	Assignment	Note ^{a}		^{19}F δ / ppm	Assignment	Note ^{a}
−141.02	(*amplo*)-F	1, s		−141.11	(*amplo*)-F	1, s
−162.71	(*parvo*)-F	1, s		−160.83	(*parvo*)-F	1, s

a: s = singlet
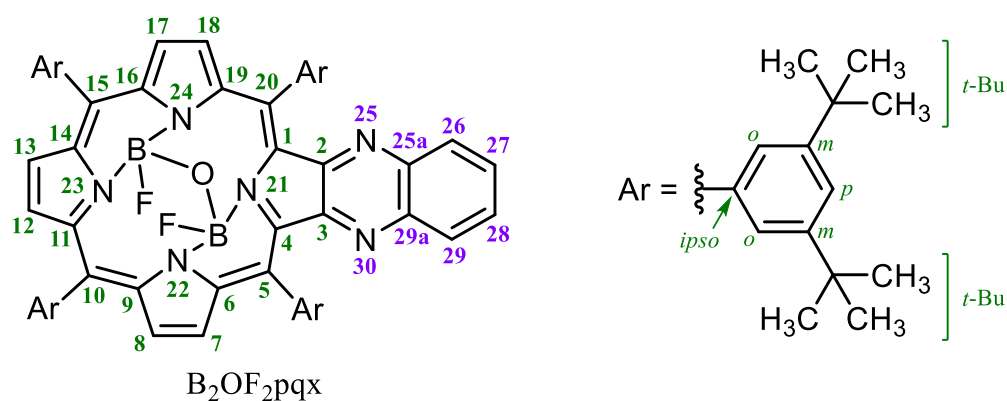
Figure D.1 Atom locants used for the NMR assignments.
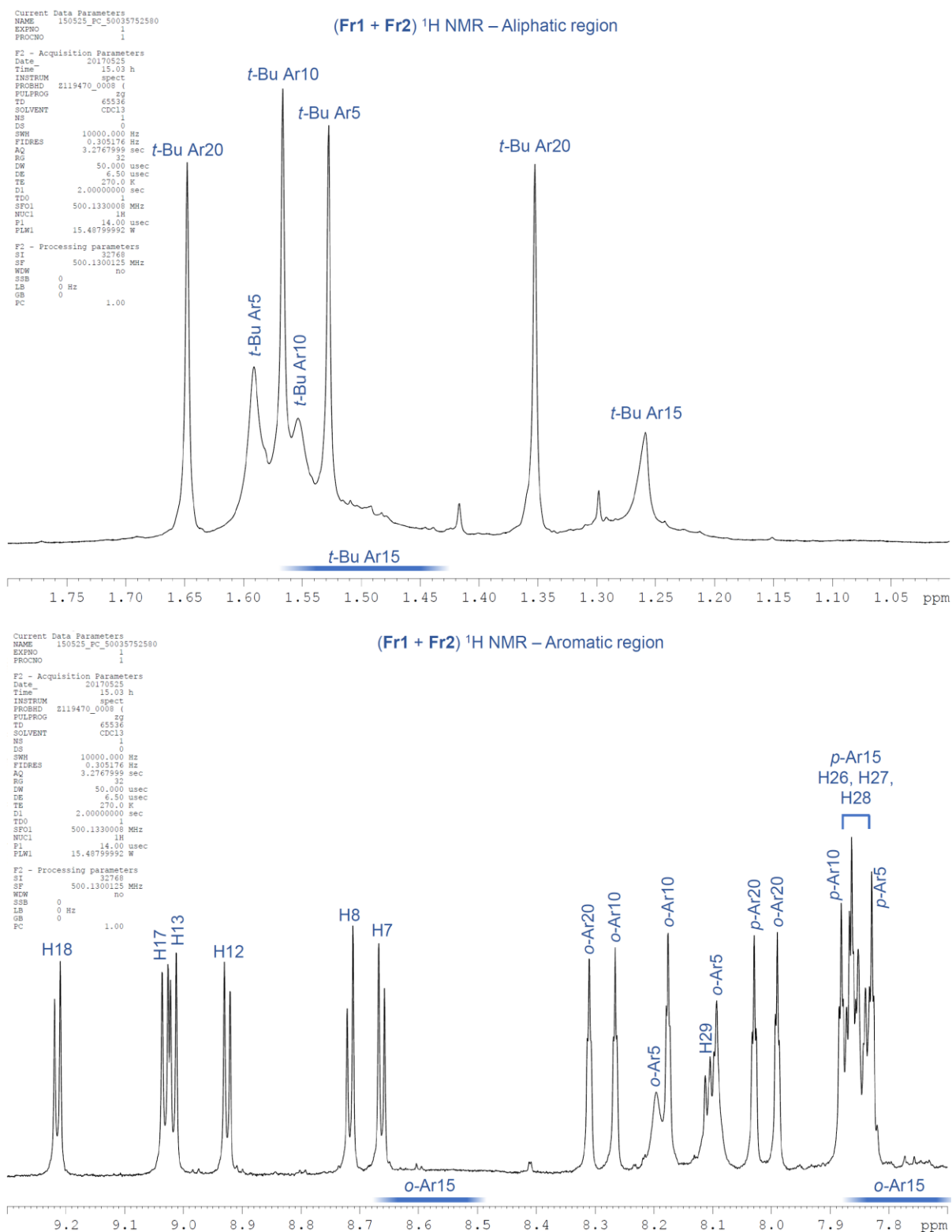
Figure. D.2 Experimental ^{1}H NMR spectrum of (**Fr1** + **Fr2**). Annotations indicate peak assignments.
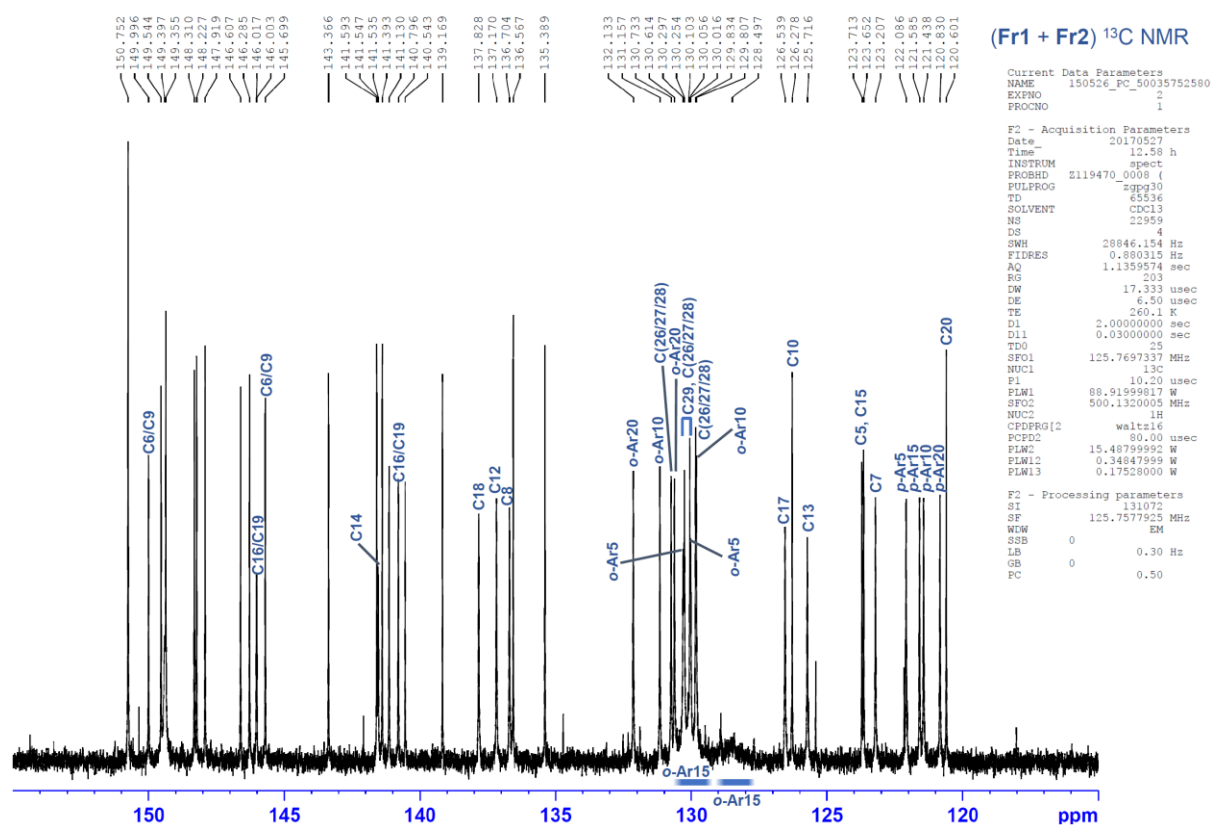
Figure D.3 Experimental ^{13}C NMR spectrum of (**Fr1** + **Fr2**). Annotations indicate peak assignments.
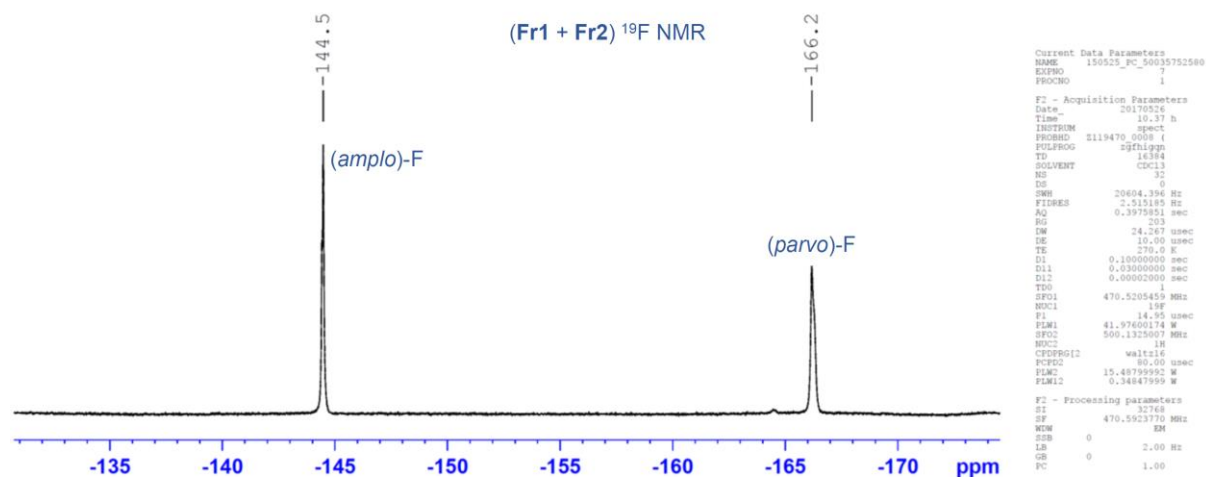
Figure D.4 Experimental ^{19}F NMR spectrum of (**Fr1** + **Fr2**). Annotations indicate peak assignments.
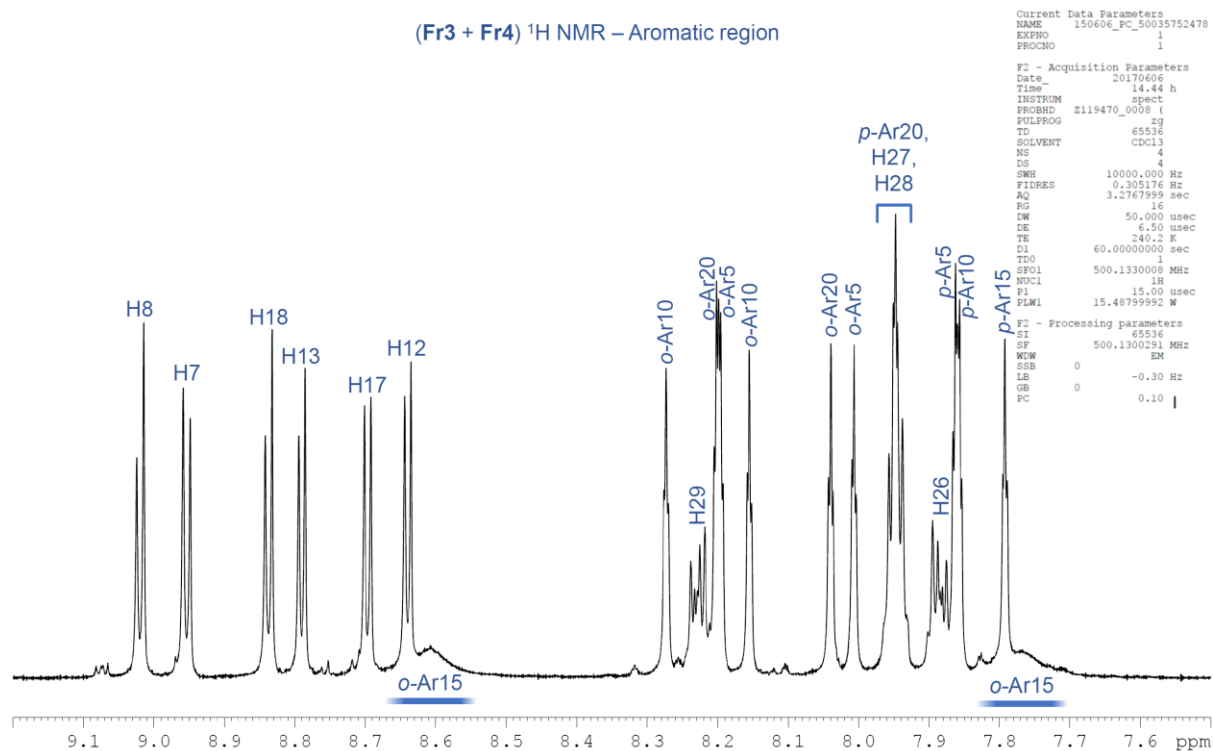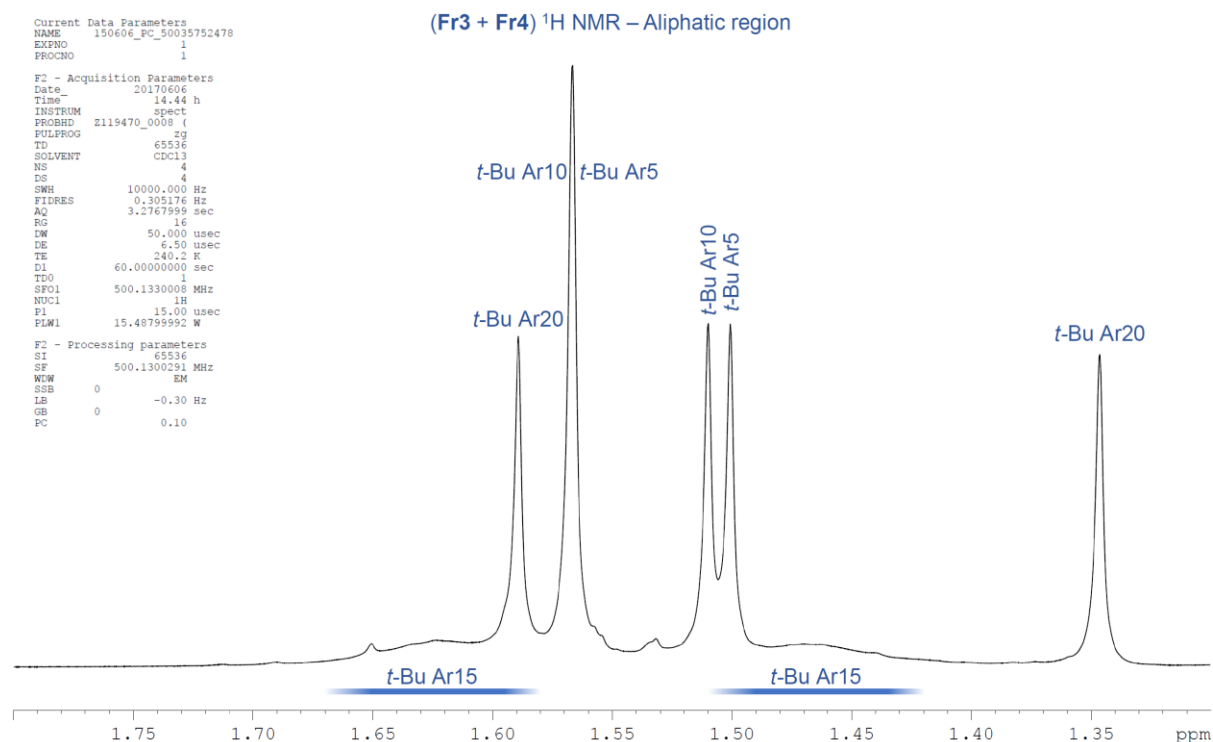
Figure D.5 Experimental ^{1}H NMR spectrum of (**Fr3** + **Fr4**). Annotations indicate peak assignments.
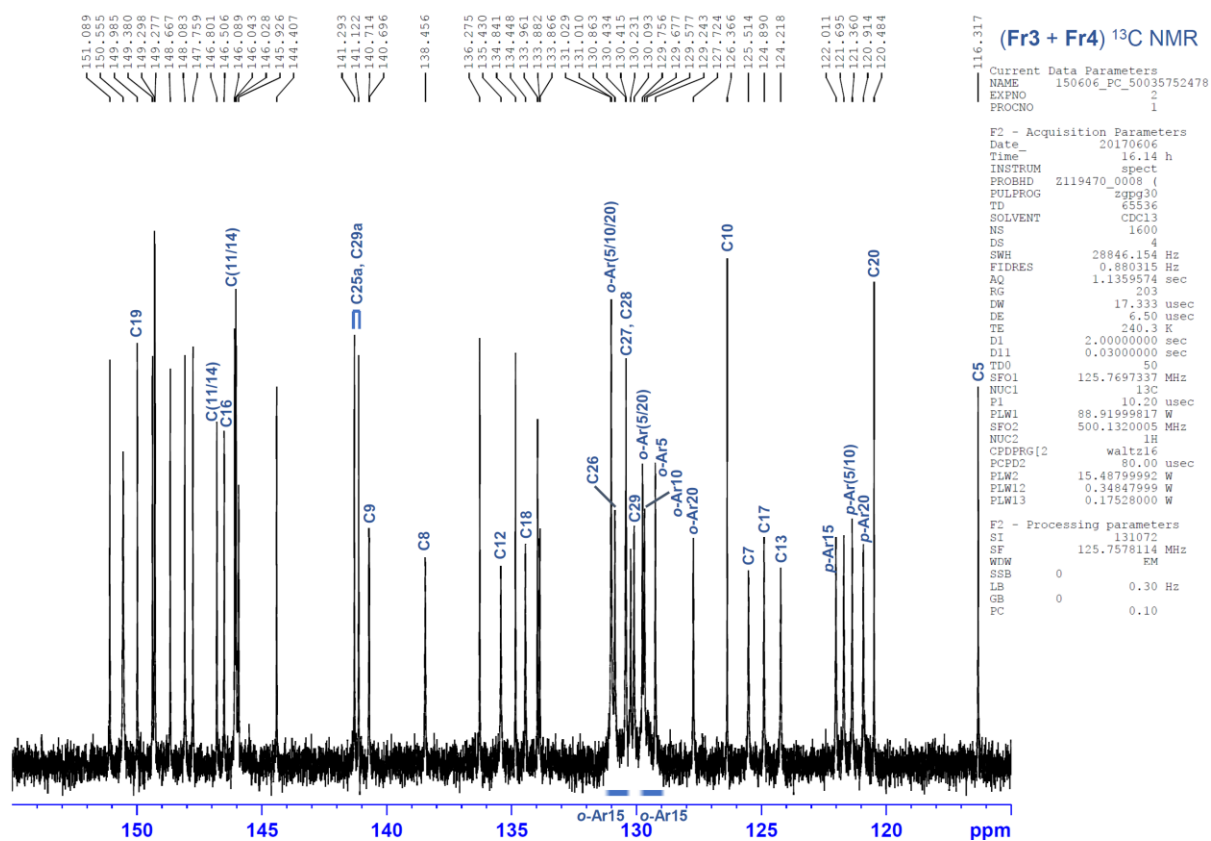
Figure D.6 Experimental ^{13}C NMR spectrum of (**Fr3** + **Fr4**). Annotations indicate peak assignments.
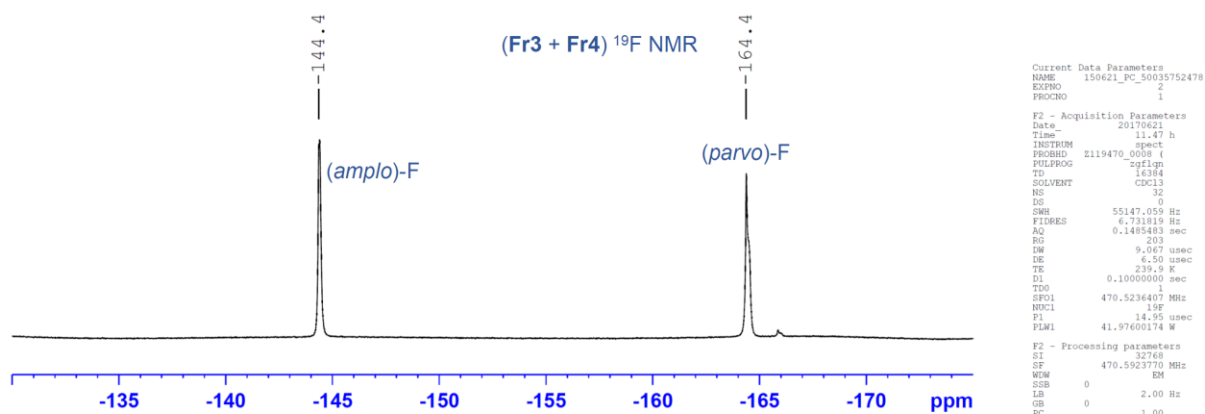
Figure D.7 Experimental ^{19}F NMR spectrum of (**Fr3** + **Fr4**). Annotations indicate peak assignments.

Appendix E: Distances from the macrocycle ring centre *O*, of the two hydrogen bonders for DFT-optimised genera structures

Table E.1 Distances from the macrocycle ring centre *O*, of the two hydrogen bonders for distal structures genera optimised using B2PLYPD3 (black values) and B3LYP-D3(BJ) (italic red values). The distances support the criterion for distal structures having both r_1, $r_2 > 0.95$.
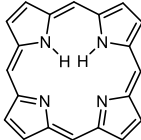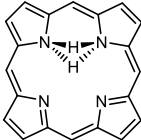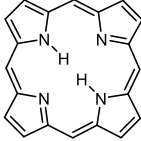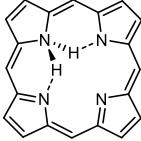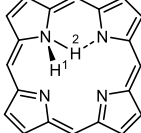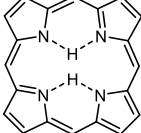
Distal structures							
#	Structure	Distance/ Å		#	Structure	Distance/ Å	
		$O - H^1$	$O - H^2$			$O - H^1$	$O - H^2$
1		1.525 *1.485*	1.525 *1.485*	**6**		1.182 *1.171*	1.234 *1.097*
2		1.155 *1.146*	1.155 *1.146*	**8**		1.422 *1.412*	1.422 *1.412*
3		1.104 *1.097*	1.104 *1.097*	**10**		1.213 *1.196*	1.213 *1.196*
4		1.846 *1.763*	1.198 *1.187*	**11**		1.275 *1.265*	1.275 *1.265*

Table E.2   Distances from the macrocycle ring centre *O*, of the two hydrogen bonders for cross-bonded structures genera optimised using B2PLYPD3 (black values) and B3LYP-D3(BJ) (italic red values).   The distances support the criterion for cross-bonded structures not satisfying the condition of both $r_1, r_2 > 0.95$.
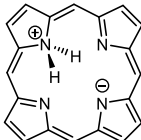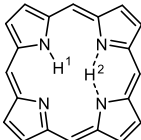
		Cross-bonded structures					
		Distance/ Å				Distance/ Å	
#	Structure	$O-H^1$	$O-H^2$	#	Structure	$O-H^1$	$O-H^2$
5		1.530 *1.552*	0.438 *0.434*	19		*0.856*	*0.685*
7		1.444 *1.516*	0.290 *0.375*	20		*0.700*	*0.570*
9		1.341 *1.301*	0.466 *0.454*	21		*0.796*	*0.532*
12		0.556 *0.588*	0.556 *0.588*	22		*0.748*	*0.748*
13		0.816 *0.826*	0.816 *0.826*	23		*0.596*	*0.596*
14		*1.337*	*0.480*	24		1.721 *1.717*	0.303 *0.336*
15		*1.578*	*0.426*	25		*1.417*	*0.484*
16		*1.974*	*0.486*	26		*0.854*	*0.522*
17		*1.284*	*0.475*	27		*0.761*	*0.601*
18		*1.445*	*0.562*	28		0.462 *0.486*	0.462 *0.486*

Appendix F: Free-Base Porphyrin 'Distal'-Structure Potential-Energy Surface Explorer

F.1 Summary

An interactive app written in Mathematica version 11.3.0.0[1] by the author of this Thesis, Peter Canfield, where any point selected on the distal-structure potential energy surface results in the corresponding structure, energy and (ϕ_1, ϕ_2) values are automatically generated and displayed. The app is located in **E_File_10**.

F.2 Background and instructions

The interactive Mathematica app, "Free-Base Porphyrin 'Distal' Structure Potential Energy Surface Explorer.cdf", can be run using the free Wolfram Player available here (for Windows, Mac and Linux).
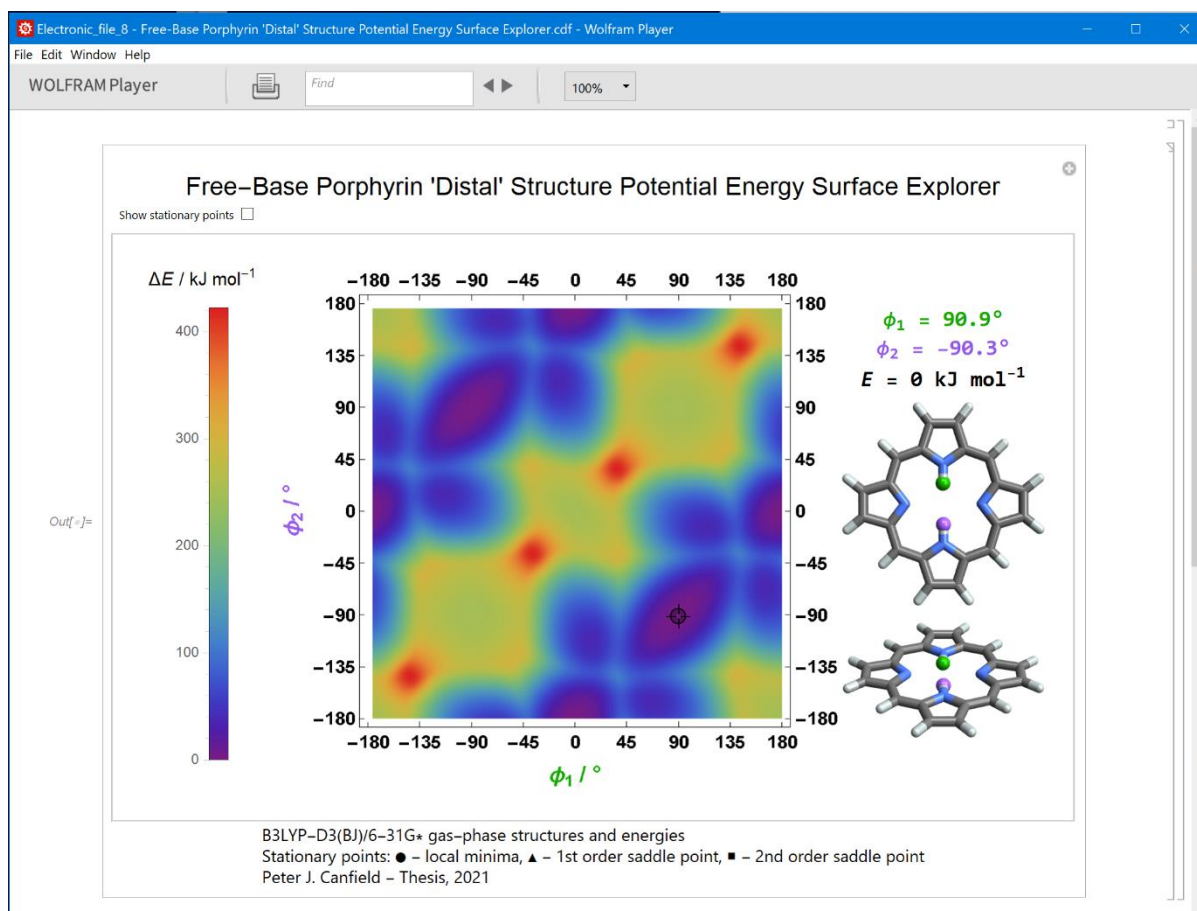
The interactive app shows the potential energy surface for the free-base porphyrin "distal" structures along with two views of the corresponding molecular model. The potential energy surface has been generated using the adaptive coordinate system described in Section S2.

The right side of the window shows two different views of the interpolation-approximated Density-Functional Theory (B3LYP-D3(BJ)/6-31G*) structure of free-base porphyrin at the corresponding (ϕ_1, ϕ_2) values. The inner H-atoms and their respective ϕ values are colour coded green and purple for easy identification. Moving the locator crosshairs on the coloured potential energy surface to a new location will generate a corresponding structure and display the (ϕ_1, ϕ_2) numerical values and the B3LYP-D3(BJ)/6-31G* energy.

There is a checkbox option to show the location of all stationary points on the surface. Printing is supported within the app.

The displayed structures are generated by quadratic interpolation between all 64 stationary points on this potential energy surface. The interpolation requires a *regular grid* of (ϕ_1,ϕ_2)-reference values, whereas the stationary points on this surface are significantly irregular. To handle this, a geometric-transform algorithm was developed whereby any arbitrary irregular grid of quadrilaterals can be mapped to a regular square grid. The transform, as shown below, allows for the mapping of the actual (ϕ_1,ϕ_2)-coordinates (a) on the distal-structure potential energy surface onto a "regularized" surface (c). Local perspective linear transforms based on the transform vector map (b) affect the mapping process. The mapped coordinates are then used to interpolate the structure.

The mapping is not perfect as it exhibits discontinuities along the grid-cell edges, especially where the irregular starting grid strongly deviates from regularity (the mapping is inhomeomorphic). Importantly though, the mapping is exact at all stationary points and performs well in their vicinity. Not surprisingly, structural accuracy is least where the surface gradient is highest. Most significantly, the mapping is computationally relatively fast and allows for near-real time execution within the app. The overall result is that the app is fast, displayed energies are accurate at the given coordinates, and the structures are qualitatively accurate and appropriate for visualization purposes.

| a) Irregular starting grid | b) Transform vector map | c) Regular final grid |
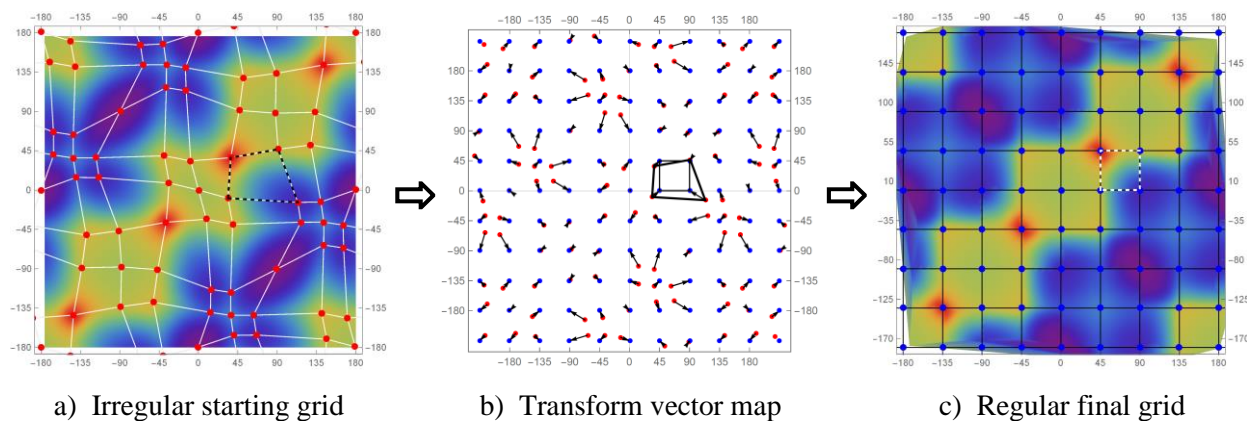
The B3LYP/6-31G* energies are qualitatively like the higher quality B2PLYP-D3(BJ) double hybrid energies with the RMS deviation = 9.3 kJ mol^{-1}. B3LYP free energies are approximately 1.5% greater than the double hybrid, implicit solvation free energies with the least-squares linear fit shown as the red dashed line:

$$12.4314 + 1.01243 \, \Delta G(\text{B2PLYP-D3})$$
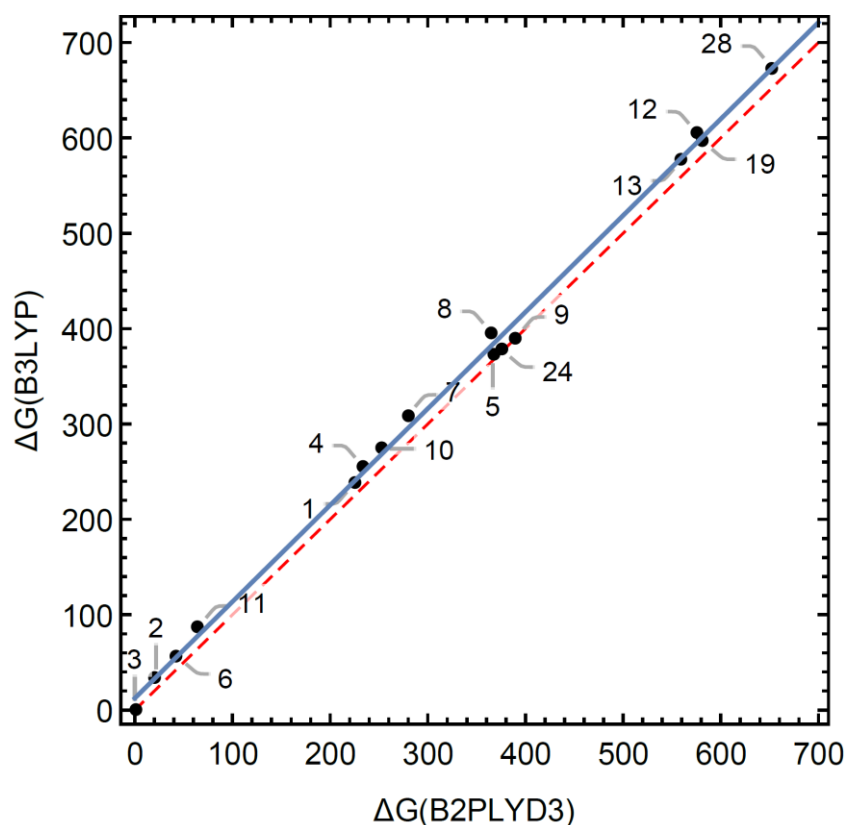
F.3 Mathematica script

```
Quiet[
  str = Interpolation[interdata];
  intf = Interpolation[grvalue,InterpolationOrder→3];

  {m1,m2,m3} = Graphics/@{Disk[{0,0},1],Triangle[{{0, (1-√3/2)/2},{1, (1-√3/2)/2},{1/2, (2+√3/2)/4}}]],Polygon[{
    {-0.5,-0.5},{0.5,-0.5},{0.5,0.5},{-0.5,0.5}}]]};
  panel1 = ListDensityPlot[grvalue,Mesh→None,LabelStyle→Directive[■,26],ColorFunction→"Rainbow",
    ImageSize→360,PlotLegends→Placed[BarLegend[{"Rainbow",{0,421}},LegendLabel→
    Style["ΔE / kJ mol⁻¹",Bold,12,FontFamily→"Helvetica"],LegendMarkerSize→350,LabelStyle→
    {Bold ,12}],Left],FrameLabel→{Style["φ1 / °",■,16,Bold],Style["φ2 / °",■,16,Bold]},
    FrameStyle→Directive[Bold,Black,13],FrameTicks→{Range[-180,180,45],
    Range[-180,180,45]},PlotRange→{{-180,180},{-180,180}},ImagePadding→{{56,Automatic},{Automatic,
    Automatic}}];
  panel2 = ListPlot[{statpoints〚spts0〛,statpoints〚spts1〛,statpoints〚spts2〛},Axes→False,PlotStyle→Black,
    PlotMarkers→Table[{s,0.02},{s,{m1,m2,m3}}],PlotRange→{{-180,180},{-180,180}}];
  NCs = {{1,5},{1,8},{2,10},{2,13},{3,15},{3,18},{4,20},{4,23}};
  CCs = {{5,6},{5,24},{6,7},{7,8},{8,9},{9,10},{10,11},{11,12},{12,13},{13,14},{14,15},{15,16},
    {16,17},{17,18},{18,19},{19,20},{20,21},{21,22},{22,23},{23,24}};
  CHs = {{6,25},{7,26},{9,33},{11,27},{12,28},{14,34},{16,29},{17,30},{19,35},{21,31},{22,32},
    {24,36}};

  CellPrint[Cell[BoxData[ToBoxes[Manipulate[
    p = transform[mp];
    Row[{If[sp,Show[panel1,panel2],panel1],
      sNHs = DeleteCases[Flatten[Table[If[EuclideanDistance[str[p〚1〛,p〚2〛]〚i〛,str[p〚1〛,p〚2〛]〚j〛]<=120,
        {i,j}],{j,1,4},{i,37,38}],1],Null];
      pNHs = DeleteCases[Flatten[Table[If[120<EuclideanDistance[str[p〚1〛,p〚2〛]〚i〛,str[p〚1〛,
        p〚2〛]〚j〛]<=180,{i,j}],{j,1,4},{i,37,38}],1],Null];
      model = Join[Table[{■,Specularity[□,100],Opacity[1],Ball[str[p〚1〛,p〚2〛]〚i〛,26]},{i,1,4}],
        Table[{■,Specularity[□,100],Opacity[1],Ball[str[p〚1〛,p〚2〛]〚i〛,26]},{i,5,24}],
        Table[{■,Specularity[□,100],Opacity[1],Ball[str[p〚1〛,p〚2〛]〚i〛,26]},{i,25,36}],
        {■,Specularity[□,100],Opacity[1],Ball[str[p〚1〛,p〚2〛]〚37〛,40]},
        {■,Specularity[□,100],Opacity[1],Ball[str[p〚1〛,p〚2〛]〚38〛,40]},
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚1〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},26]}&/@NCs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚2〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},26]}&/@NCs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚1〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},26]}&/@CCs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚2〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},26]}&/@CCs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚1〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},26]}&/@CHs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚2〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},26]}&/@CHs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚1〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},26]}&/@sNHs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚2〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},26]}&/@sNHs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚1〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},12]}&/@pNHs,
        {■,Specularity[□,100],Opacity[1],Tube[{str[p〚1〛,p〚2〛]〚#〚2〛〛,Mean[str[p〚1〛,p〚2〛]〚#〛]},12]}&/@pNHs
      ];
      Column[{Style["φ1 = "<>ToString[Round[mp〚1〛,.1]]<>"°",■,16,Bold],
        Style["φ2 = "<>ToString[Round[mp〚2〛,.1]]<>"°",■,16,Bold],
        Style["E = "<>ToString[Round[intf[mp〚1〛,mp〚2〛],1]]<>" kJ mol⁻¹",■,16,Bold],
        Graphics3D[model,Boxed→False,SphericalRegion→False,Lighting→"Neutral",ViewPoint→Above,
          ImageSize→150],
        Graphics3D[model,Boxed→False,SphericalRegion→False,Lighting→"Neutral",ViewPoint→
          {0,-3,2},ImageSize→150]},Center]}],
    {{mp,{90,-90}},Locator},
    {{sp,False,"Show stationary points"},{True,False},Checkbox,ControlType→Setter},
    SaveDefinitions→True,
    FrameLabel→{Style[
    "B3LYP-D3(BJ)/6-31G* gas-phase structures and energies  \nStationary points: ● - local minima,
      ▲ - 1st order saddle point, ■ - 2nd order saddle point\n Peter J. Canfield - Thesis, 2021",
    TextAlignment→Left,12],None,
    Style["Free-Base Porphyrin 'Distal' Structure – Potential Energy Surface Explorer",■,20],None}]]
]
]
```

With the following initialised definitions:

```
transform[mp_]:=Module[{i,j,p},For[i=1,i<=10,i++,
  For[j=1,j<=10,j++,
    If[RegionMember[Polygon[{statmesh〚i,j〛,statmesh〚i+1,j〛,statmesh〚i+1,j+1〛,statmesh〚i,j+1〛}],mp],
    {error,lft}=FindGeometricTransform[{gridmesh〚i,j〛,gridmesh〚i+1,j〛,gridmesh〚i+1,j+1〛,gridmesh〚i,j+1〛},
      {statmesh〚i,j〛,statmesh〚i+1,j〛,statmesh〚i+1,j+1〛,statmesh〚i,j+1〛}];
```

```
    p=lft[mp];
    i=10;j=10;
    ]
  ]
];p]
```

spts0={17,19,21,37,39,41,43,57,59,61,63,65,79,81,83,85,101,103,105,109};
spts1={18,20,26,28,30,32,36,38,40,42,48,50,52,54,58,60,62,64,68,70,72,74,80,82,84,86,90,92,94,96,102,104};
spts2={25,27,29,31,47,49,51,53,69,71,73,75,91,93,95,97};

statpoints={{-216.92,-216.88},{-170.64,-213.84},{-140.88,-219.1},{-62.89,-216.81},{-36.72,-216.71},{-13.68,-220.},{52.33,-232.33},{88.89,-226.58},{143.08,-216.88},{189.36,-213.89},{219.12,-219.1},

  {-213.35,-170.64},{-179.89,-179.89},{-146.63,-189.26},{-66.48,-194.16},{-40.,-190.},{-0.15,-179.85},{40.01,-166.32},{66.4,-165.87},{146.65,-170.64},{179.35,-179.32},{213.37,-189.26},

  {-219.1,-140.9},{-189.79,-146.57},{-143.09,-143.13},{-89.04,-133.41},{-51.68,-128.33},{13.68,-139.97},{36.75,-143.38},{62.85,-143.18},{140.9,-140.9},{170.21,-146.57},{216.91,-143.13},

  {-216.81,-62.85},{-194.01,-66.34},{-133.42,-88.93},{-88.,-87.99},{-46.58,-91.07},{14.15,-113.63},{36.8,-117.17},{89.41,-90.6},{143.19,-62.85},{165.99,-66.34},{226.58,-88.93},

  {-216.63,-36.76},{-194.62,-40.},{-128.32,-51.66},{-91.,-46.6},{-36.92,-36.87},{9.02,-33.34},{39.09,-39.12},{117.13,-36.81},{143.37,-36.76},{165.38,-40.},{232.34,-52.34},

  {-220.,-15.},{-179.83,-0.13},{-140.,14.39},{-113.64,14.12},{-33.56,8.},{0.,0.},{33.83,-9.4},{113.63,-14.12},{140.,-13.71},{179.93,0.26},{220.,14.12},

  {-232.35,52.34},{-165.64,40.},{-143.31,36.68},{-117.14,36.81},{-39.09,39.1},{-8.91,33.32},{36.88,36.93},{91.09,46.58},{128.33,51.67},{194.35,40.},{216.69,36.68},

  {-226.58,88.94},{-165.87,66.38},{-143.19,62.87},{-89.49,90.5},{-36.8,117.11},{-14.07,113.53},{46.58,91.06},{88.,87.88},{133.42,88.94},{194.13,66.38},{216.81,62.87},

  {-216.92,143.12},{-170.64,146.16},{-140.88,140.9},{-62.89,143.19},{-36.72,143.29},{-15.,140},{51.58,128.33},{88.89,133.42},{143.08,143.12},{189.36,146.16},{216.,143.97},

  {-213.35,189.36},{-179.89,180.68},{-146.63,170.74},{-66.48,165.84},{-40.,167.14},{-0.15,180.15},{40.5,188.},{66.4,194.13},{146.65,189.36},{180.11,180.11},{213.37,170.74},

  {-219.1,219.1},{-189.79,213.43},{-143.09,216.87},{-89.04,226.59},{-52.34,232.34},{13.68,220.03},{36.75,216.62},{62.85,216.82},{140.9,219.1},{170.21,213.43},{216.91,216.87}};
gridpoints={{-225,-225},{-180,-225},{-135,-225},{-90,-225},{-45,-225},{0,-225},{45,-225},{90,-225},{135,-225},{180,-225},{225,-225},{-225,-180},{-180,-180},{-135,-180},{-90,-180},{-45,-180},{0,-180},{45,-180},{90,-180},{135,-180},{180,-180},{225,-180},{-225,-135},{-180,-135},{-135,-135},{-90,-135},{-45,-135},{0,-135},{45,-135},{90,-135},{135,-135},{180,-135},{225,-135},{-225,-90},{-180,-90},{-135,-90},{-90,-90},{-45,-90},{0,-90},{45,-90},{90,-90},{135,-90},{180,-90},{225,-90},{-225,-45},{-180,-45},{-135,-45},{-90,-45},{-45,-45},{0,-45},{45,-45},{90,-45},{135,-45},{180,-45},{225,-45},{-225,0},{-180,0},{-135,0},{-90,0},{-45,0},{0,0},{45,0},{90,0},{135,0},{180,0},{225,0},{-225,45},{-180,45},{-135,45},{-90,45},{-45,45},{0,45},{45,45},{90,45},{135,45},{180,45},{225,45},{-225,90},{-180,90},{-135,90},{-90,90},{-45,90},{0,90},{45,90},{90,90},{135,90},{180,90},{225,90},{-225,135},{-180,135},{-135,135},{-90,135},{-45,135},{0,135},{45,135},{90,135},{135,135},{180,135},{225,135},{-225,180},{-180,180},{-135,180},{-90,180},{-45,180},{0,180},{45,180},{90,180},{135,180},{180,180},{225,180},{-225,225},{-180,225},{-135,225},{-90,225},{-45,225},{0,225},{45,225},{90,225},{135,225},{180,225},{225,225}};

statmesh=Table[statpoints[[11j+i]],{j,0, 10},{i,11}];
gridmesh=Table[gridpoints[[11j+i]],{j,0, 10},{i,11}];

interdata = Molecular structure Cartesian coordinates for all stationary points on the potential energy surface over the *regularised* domain ϕ_1 and $\phi_2 \in [-225°,225°]$ in 45° increments. Format is $\{\ldots\{\{\phi_1,\phi_2,\},\{\{x_1,y_1,z_1\},\ldots,\{x_{38},y_{38},z_{38}\}\}\}\ldots\}$

grvalue = List of points defining the B3LYP-D3(BJ)/6-31G* calculated potential energy surface in kJ mol^{-1} over the *non-regularised* domain ϕ_1 and $\phi_2 \in [-200°,200°]$ in 2° increments. Format is $\{\ldots\{\phi_1,\phi_2,E\}\ldots\}$ and is a 2-level list.

1   Wolfram Research, I. *Mathematica*, 11.3; Wolfram Research, Inc.: Champaign, Illinois, 2018.

Appendix G: Software for generating families, genera, and species for $\mathcal{S}_m\mathcal{B}_n$ partitioned chemical systems

G.1 Software to generate families, genera, and species for $\mathcal{S}_m\mathcal{B}_n$ with m, n ≤ 8

E_File_11 is a .zip file containing the Fortran source code written by Professor Jeffrey R. Reimers, and associated PC executable (for any 32-bit Windows platform), for the determination of the number of families, genera (assuming non-dissociating and non-differentiable binders and D_{mh} site symmetry), and species for $\mathcal{S}_m\mathcal{B}_n$ with m, $n \leq 8$. This restriction is caused by the 64-bit encoding used to store the structures internally. When both m and n are large, the requisite computer time and disk-space to store the results, increases superexponentially. Basics checks on the program's accuracy can be made by comparing the total number of families and species to the results from Eqns. (3a) and (4a) in Chapter 6.(see Table G.1 and Table G.3). The program has three modes of use:

genus_direct <m> <n>

produces two ASCII text files named "genera_Dnh_SmBn.txt" listing the families, genera, and number of species per genera, as well as "species_SmBn.txt" containing the families, their genera, and the total number of species in each genus.

genus_direct <m> <n> n

as above, but only the file "species_SmBn.txt" is produced, saving disk space.

genus_direct <m> <n> g

as above but only the file "species_SmBn.txt" is produced, saving disk space, and also the number of species per genera is not calculated, saving CPU time.

Results are also included in **E_File_11** for all the values of m and n listed in Table G.2. An abbreviated notation for genera and species is used compared to the IUPAC-consistent notation described in the main text and SI Section S1. In this, the µ notation is not applied as this information is also included in the family specification. Again, the notation lists bonds from the linkers, specified in order and separated by colons, to the numbered sites. Also, bond-hole notation flagged by "\" is used when the number of bonds exceeds $m/2$. This notation can be very compact, *e.g.*, the name for genus **28** is simply "\:\" (there are no bond holes as each bonder links to each site, hence no site numbers appear in the name).

G.2 Numbers of families for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation

The number of families $N_f = \frac{(m+n-1)!}{(m-1)!\ n!}$, of D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation is shown in Table G.1.

Table G.1 The number of families N_f, of D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation.

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2	2	3	4	5	6	7	8	9
3	3	6	10	15	21	28	36	45
4	4	10	20	35	56	84	120	165
5	5	15	35	70	126	210	330	495
6	6	21	56	126	252	462	792	1 287
7	7	28	84	210	462	924	1 716	3 003
8	8	36	120	330	792	1 716	3 432	6 435

G.3   Numbers of genera for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation

The number of genera N_g of D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation, as determined using the software provided in **E_File_11**[a] is shown in Table G.2.

Table G.2 The number of genera N_g of D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation, as determined using the software provided.

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2	2	4	6	9	12	16	20	25
3	3	9	23	51	103	196	348	*590*
4	5	28	124	494	1 726	*5 476*	*15 876*	*42 696*
5	7	70	630	4 950	*33 474*	*197 778*		
6	12	224	3 969	*62 100*	*814 826*			
7	17	669	25 725	*817 037*				
8	29	2 269	*177 771*					

a: the number of species in each genus (except for those in blue italics), and their names, are also included in the **E_File_11**.

G.4 Numbers of species for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation

The number of species $N_s = (2^m - 1)^n$ of D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation is shown in Table G.3.

Table G.3 The number of species N_s of D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation.

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2	3	9	27	81	243	729	2 187	6 561
3	7	49	343	2 401	16 807	117 649	823 543	5 764 801
4	15	225	3 375	50 625	759 375	11 390 625	170 859 375	2 562 890 625
5	31	961	29 791	923 521	28 629 151	887 503 681	2.8×10^{10}	8.5×10^{11}
6	63	3969	250 047	15 752 961	992 436 543	6.3×10^{10}	3.9×10^{12}	2.5×10^{14}
7	127	16 129	2 048 383	260 144 641	3.3×10^{10}	4.2×10^{12}	5.3×10^{14}	6.8×10^{16}
8	255	65 025	16 581 375	4 228 250 625	1.1×10^{12}	2.7×10^{14}	7.0×10^{16}	1.8×10^{19}

G.5 Numbers of families N_f for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 2

Table G.4 The number of families N_f for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 2.

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	1	-	-	-	-	-	-
2	2	3	-	-	-	-	-	-
3	3	6	6	-	-	-	-	-
4	4	10	13	11	-	-	-	-
5	5	15	24	24	18	-	-	-
6	6	21	40	46	39	29	-	-
7	7	28	61	80				-
8	8	36	89					

G.6 Numbers of genera N_g for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 2

Table G.5 The number of genera N_g for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 2.[a]

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	1	-	-	-	-	-	-
2	2	4	-	-	-	-	-	-
3	3	9	10	-	-	-	-	-
4	5	28	57	64	-	-	-	-
5	7	70	281	524	476	-	-	-
6	12	224	1 665	5 160	*7 455*	*5 907*	-	-
7	17	669	9 654	*50 468*				-
8	29	2 269	*59 548*					

a: the number of species in each genus (except for those in italics owing to size limitations), and their names, are also included in **E_File_11**.

G.7 Numbers of species N_s for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 2

Table G.6 The number of species N_s for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 2.[a]

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	1	-	-	-	-	-	-
2	3	9	-	-	-	-	-	-
3	7	49	174	-	-	-	-	-
4	15	225	1 680	6 510	-	-	-	-
5	31	961	13 830	99 840	401 310	-	-	-
6	63	3 969	105 552	1 325 286	9 055 260	369 98 100	-	-
7	127	16 129	774 774	16 290 792				-
8	255	65 025	5 568 960					

G.8 Numbers of families N_f for D_{mh} bonding-site symmetric $S_m B_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 1

Table G.7 The number of families N_f for D_{mh} bonding-site symmetric $S_m B_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 1.

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	-	-	-	-	-	-	-
2	2	1	-	-	-	-	-	-
3	3	2	1	-	-	-	-	-
4	4	4	2	1	-	-	-	-
5	5	6	4	2	1	-	-	-
6	6	9	7	4	2	1	-	-
7	7	12	11	7	4	2	1	-
8	8	16	16	12	7	4	2	1

G.9 Numbers of genera N_g for D_{mh} bonding-site symmetric $S_m B_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 1

Table G.8 The number of genera N_g for D_{mh} bonding-site symmetric $S_m B_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 1.

| $m \setminus n$	1	2	3	4	5	6	7	8
1	1	-	-	-	-	-	-	-
2	2	1	-	-	-	-	-	-
3	3	2	1	-	-	-	-	-
4	5	7	3	1	-	-	-	-
5	7	14	11	3	1	-	-	-
6	12	39	43	20	4	1	-	-
7	17	88	148	92	28	4	1	-
8	29	236	555	500	199	43	5	1

G.10 Numbers of species N_s for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 1

Table G.9 The number of species N_s for D_{mh} bonding-site symmetric $\mathcal{S}_m\mathcal{B}_n$ partitioned systems under the constraint of non-dissociation when the maximum number of bonds formed to any site is limited to 1.

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	-	-	-	-	-	-	-
2	3	2	-	-	-	-	-	-
3	7	12	6	-	-	-	-	-
4	15	50	60	24	-	-	-	-
5	31	180	390	360	120	-	-	-
6	63	602	2 100	3 360	2 520	720	-	-
7	127	1 932	10 206	25 200	31 920	20 160	5 040	-
8	255	6 050	46 620	166 824	317 520	332 640	181 440	40 320

Appendix H: $\mathcal{S}_m\mathcal{B}_n$ partitioned Polytope Formalism of constitutional isomerism graphing software

This software, written in Mathematica version 11.3.0.0,[1] is found in **E_File_12** and is named "SmBn partitioned Polytope Formalism of constitutional isomerism graphing.nb". It reads the species data files written by the software described in Appendix G to construct all possible concerted unimolecular interconversion pathways. The program begins with three variables that must be set (highlighted in yellow below):
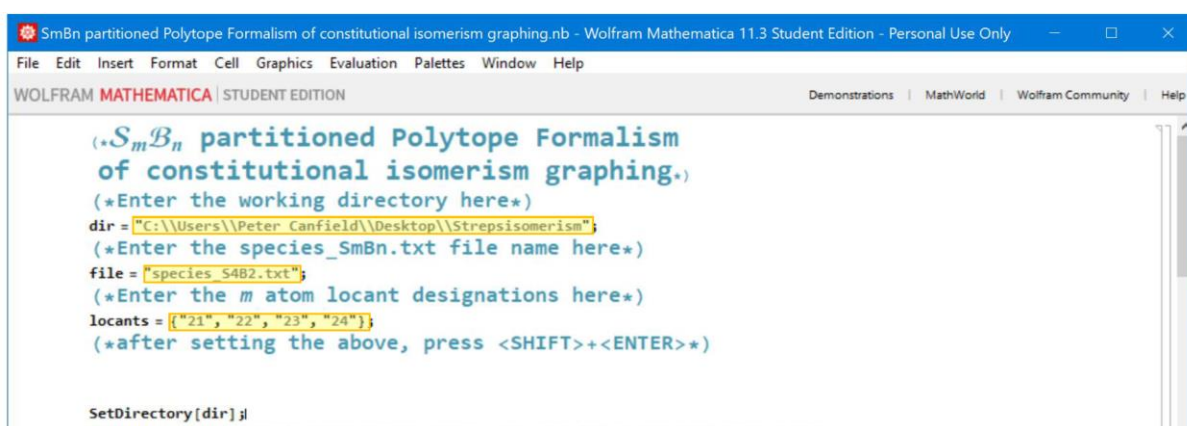
1. The working-directory path. If you paste from the clipboard between the bounding quote marks a dialog window will pop-up. Select "Yes".
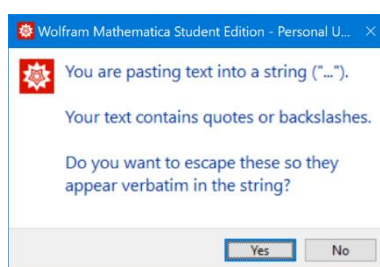
    If manually entering the path, then use two backslashes "\\" to indicate any backslash character found in the path name.

    The "species_S*m*B*n*.txt" file generated by the Fortran program **E_File_11** described in Appendix G. This .txt file must be located in the working directory.

2. Atom locant names for each of the *m* $\mathcal{S}$ atoms. These are used to generate application specific and chemically sensible topological-polytope species names.

After these variables are set press <SHIFT> + <ENTER> to execute the block of code which will produce an interactive panel as shown below. The controls allow the user to customise and optimise the outputs.

1. The graph vertex layout options are the 12 most useful ways for arranging ("embedding" – in this context this has a different meaning to that given in 2.4.9) the graph vertices for display. "SpringElectricalEmbedding" is generally the most useful and set as the default. It should be noted that the embedding layouts are not unique for a given set of genera with the displayed results occasionally differing depending upon the order in which the genera are selected.

2. Checkboxes corresponding to the total number of genera found in the "species_SmBn.txt" file allow the user to select any set of genera to use in constructing the graphs.

3. The graph vertex label font size can be set using the slider or typing directly into the box. Use this to improve the readability of the outputs.

4. The image size sets the width size, in pixels, of the graphs. Whilst the graph vertex label size remains fixed, a graph that has dense and overlapping labels and/or edges can be made more readable by increasing the image size. This setting can be made using the slider control or by manually entering any positive integer value in the associated field

5. "Root file name for saving" is text that is used as a prefix when saving the autogenerated output files.

6. Three "save" buttons for saving each of the first-, second- and combined-order motions graphs, respectively.

First-order motions correspond to graph edges connecting two graph vertices related by a single topological change at a single centre. The first-order motions graph is shown in the first panel.

Second-order motions correspond to edges connecting two graph vertices related by a single topological change at two centres simultaneously. The second-order motions graph is shown in the second panel.

The third panel shows the graph made by combining both first- and second-order motions.

In each of the panels there appears the details of the graph along with the following properties: Whether the graph is Hamiltonian, Eulerian or planar, its radius and diameter, and graph density.

When the "save" control buttons are pressed the contents of the corresponding panel is saved to the working directory as a .pdf file along with the graph in the .gxl format, the edge and vertex lists as .txt files, and the adjacency and distance matrices as .dat files (ordered as per the vertex list file).

Example results include:

```
Class: S4B2 with site symmetry point group D4h
Genera included = {1, 2, 3, 4, 6, 8, 10, 11}
1st order motions graph:
Node layout: SpringEmbedding
Hamiltonian graph? True
Eulerian graph? True
Planar graph? False
Graph radius = 8
Graph diameter = 8
Graph density = 0.0634921
```


Figure H.1 Distal structures (genera **1**, **2**, **3**, **4**, **6**, **8**, **10**, and **11**) first-order motions graph. First order motions involve a single bond-topology change at a single centre. They connect local minima to transition states, transition states to 2nd-order saddle points, *etc*.

This distal-structures first-order motions graph can be mapped onto the potential energy surface (Chapter 6 main text Figure 6.8 and Figure 6.10) and is shown in (a) below.
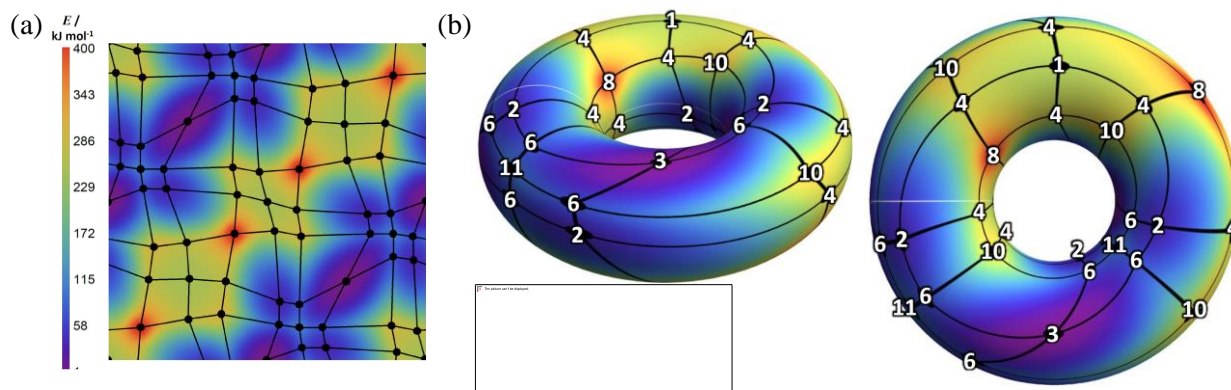
Every graph vertex represents a structure with each connected by four graph edges representing possible reaction paths to/from other structures. (b) Further, the periodicity of both ϕ_1 and ϕ_2 means this potential energy surface and spanning graph can be mapped onto a torus with parametric form given by:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (R + r\,\mathrm{Cos}\phi_2)\,\mathrm{Cos}\phi_1 \\ (R + r\,\mathrm{Cos}\phi_2)\,\mathrm{Sin}\phi_1 \\ r\,\mathrm{Sin}\phi_2 \end{pmatrix}$$

where R is the "outer" radius of the torus and r the inner-tubular radius, here set to 2 and 1, respectively. Two different views are given in (b) with the graph vertices labelled by their corresponding genus number.

```
Class: S4B2 with site symmetry point group D4h
Genera included = {1, 2, 3, 4, 6, 8, 10, 11}
1st and 2nd order motions graph:
Graph vertex layout: SpringElectricalEmbedding
Hamiltonian graph? True
Eulerian graph? True
Planar graph? False
Graph radius = 4
Graph diameter = 4
Graph density = 0.126984
```
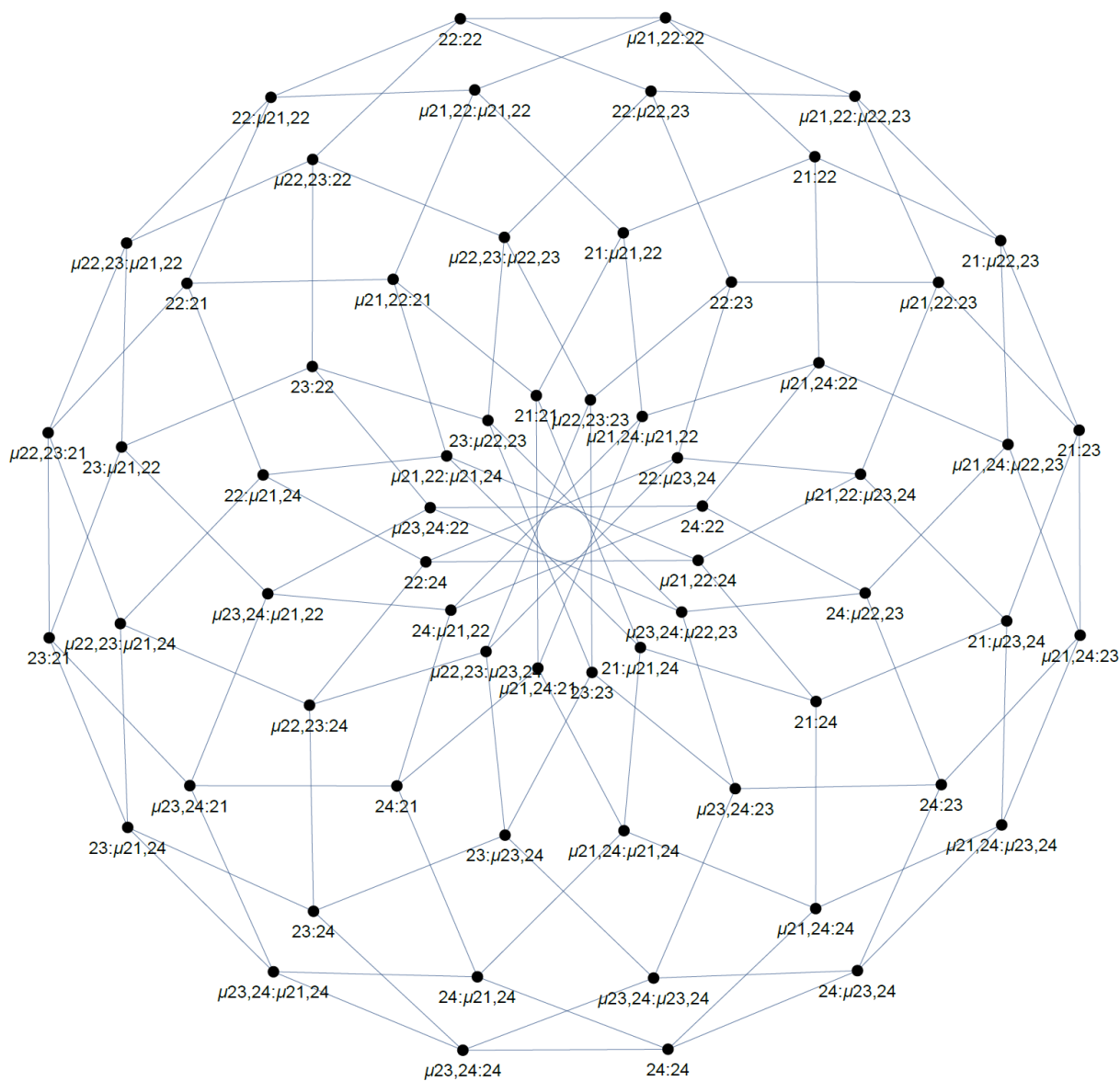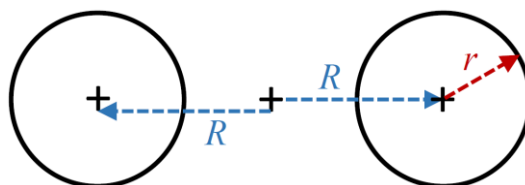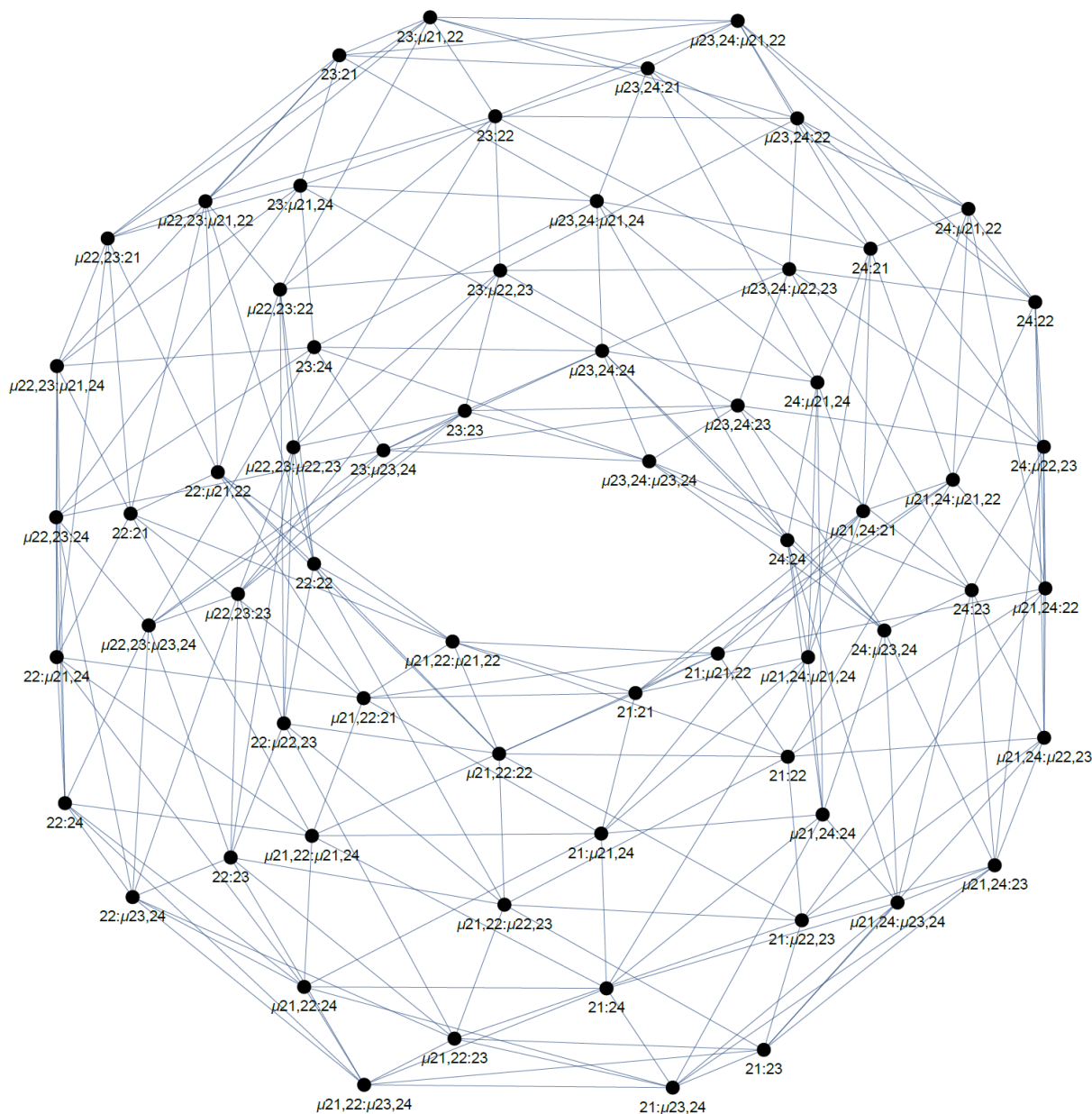

Figure H.2 Distal structures (genera **1**, **2**, **3**, **4**, **6**, **8**, **10**, and **11**) first- and second-order motions graph. Second-order motions involve the simultaneous bond-topology change at two centres. Second-order motions will connect local minima to a second-order saddle points, *etc.* Note the implied toroidal form. All vertices are of degree eight split into four first-order motion edges and four second-order motion edges.

Class: S4B2 with site symmetry point group D4h
Genera included = {2, 3, 6}
1st order motions graph:
Graph vertex layout: TutteEmbedding
Hamiltonian graph? False
Eulerian graph? True
Planar graph? True
Graph radius = 8
Graph diameter = 8
Graph density = 0.0846561
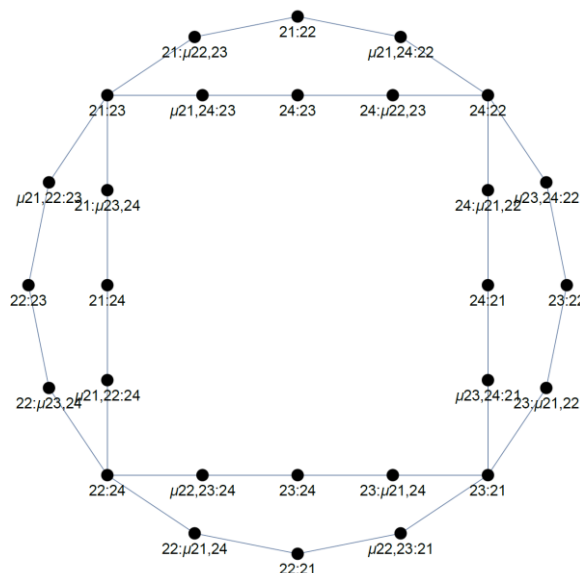
Class: S4B2 with site symmetry point group D4h
Genera included = {2, 3, 6}
2nd order motions graph:
Graph vertex layout: TutteEmbedding
Hamiltonian graph? True
Eulerian graph? False
Planar graph? True
Graph radius = 5
Graph diameter = 5
Graph density = 0.2

Separately graphed first order motions and second-order motions for genera **2**, **3** and **6** (*cis*, *trans* and linking TS).

Class: S4B2 with site symmetry point group D4h
Genera included = {2, 3, 6}
1st and 2nd order motions graph:
Graph vertex layout: SpringEmbedding
Hamiltonian graph? True
Eulerian graph? False
Planar graph? True
Graph radius = 5
Graph diameter = 6
Graph density = 0.148148
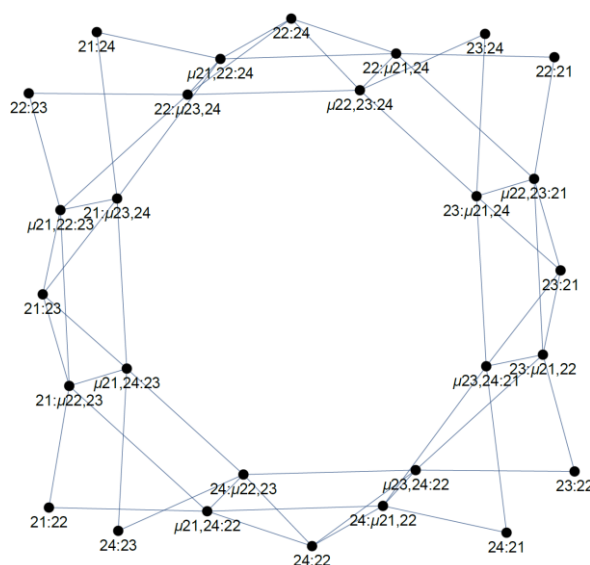
Figure H.3 Combined graph of first order motions and second-order motions for genera **2**, **3** and **6** (*cis*, *trans* and linking TS).

```
Class: S4B2 with site symmetry point group D4h
Genera included = {1, 2, 3, 4, 6}
1st order motions graph:
Graph vertex layout: SpringElectricalEmbedding
Hamiltonian graph? False
Eulerian graph? True
Planar graph? False
Graph radius = 8
Graph diameter = 8
Graph density = 0.0567376
```
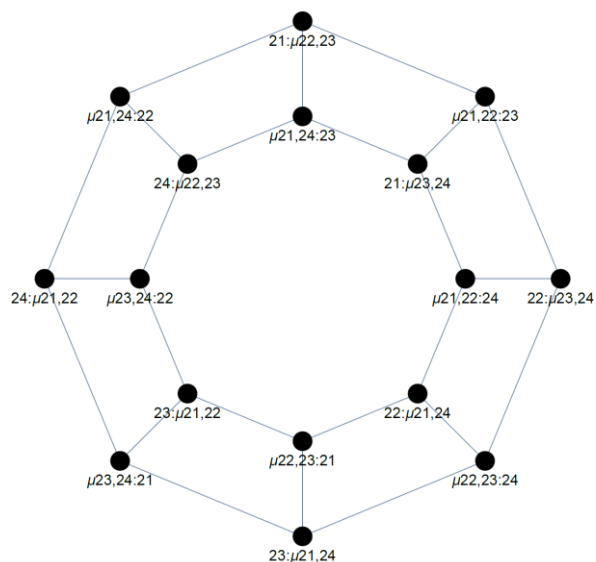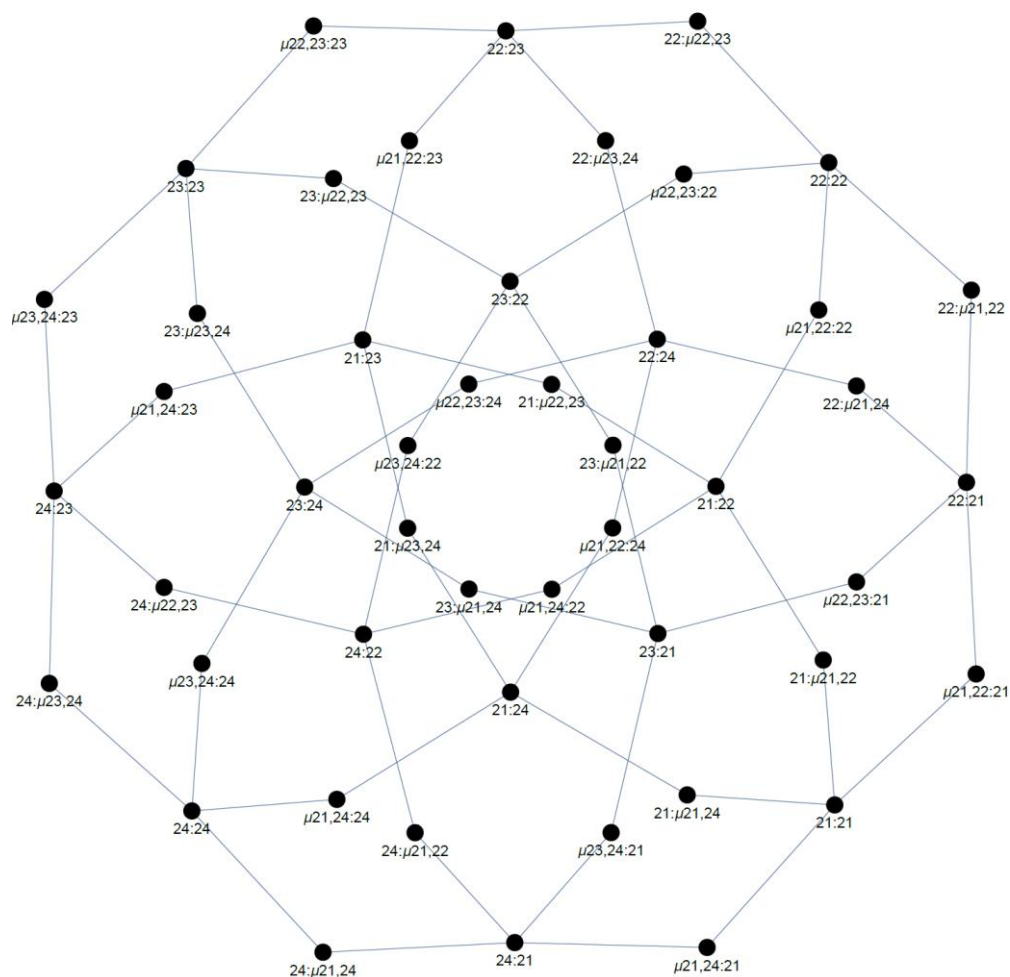

Figure H.4 First-order motions for genera **1**, **2**, **3**, **4** and **6** (*gem*, *cis, trans*, first-order distal TS structures only).

Output from this software for all 225 species (genera **1** – **28**) for the D_{mh} bonding-site symmetric $\mathcal{S}_4\mathcal{B}_2$ partitioned system under the constraint of non-dissociation and comprised of the following 18 files:

a. "S4B2_genera{1-28}_1.pdf": a high-resolution pdf file for all first-order motions.
b. "S4B2_genera{1-28}_1.gxl": first-order motions graph in .gxl format.
c. "S4B2_genera{1-28}_Vertex_List_1.txt" first-order motions graph vertex list as a .txt file.
d. "S4B2_genera{1-28}_edges_1.txt" first-order motions graph vertex list as a .txt file.
e. "S4B2_genera{1-28}_Distance_Matrix_1.dat" the 225 × 225 first-order motions graph distance matrix as a .dat file. The column/row order follows the order in the vertex list file (c).
f. "S4B2_genera{1-28}_Adjacency_Matrix_1.dat" the 225 × 225 adjacency matrix of the first-order motions graph as a .dat file. The column/row order follows the order in the vertex list file (c).
g. "S4B2_genera{1-28}_2.pdf": a high-resolution pdf file for all second-order motions.
h. "S4B2_genera{1-28}_2.gxl": the second-order motions graph in .gxl format.
i. "S4B2_genera{1-28}_Vertex_List_2.txt" vertex list of the second-order motions graph as a .txt file.
j. "S4B2_genera{1-28}_edges_2.txt" second-order motions graph edge list as a .txt file.
k. "S4B2_genera{1-28}_Distance_Matrix_2.dat" the 225 × 225 second-order motions graph distance matrix as a .dat file. The column/row order follows the order in the vertex list file (h).
l. "S4B2_genera{1-28}_Adjacency_Matrix_2.dat" the 225 × 225 second-order motions graph adjacency matrix as a .dat file. The column/row order follows the order in the vertex list file (i)
m. "S4B2_genera{1-28}_1_2.pdf": a high-resolution pdf file for all first- and second-order motions.
n. "S4B2_genera{1-28}_1_2.gxl": the first- and second-order motions graph in .gxl format.
o. "S4B2_genera{1-28}_Vertex_List_1_2.txt" first- and second-order motions graph vertex list as a .txt file.
p. "S4B2_genera{1-28}_edges_1_2.txt" first- and second-order motions graph edge list as a .txt file.
q. "S4B2_genera{1-28}_Distance_Matrix_1_2.dat" the 225 × 225 first- and second-order motions graph distance matrix as a .dat file. The column/row order follows the order in the vertex list file (m).
r. "S4B2_genera{1-28}_Adjacency_Matrix_1_2.dat" the 225 × 225 first- and second-order motions graph adjacency matrix as a .dat file. The column/row order follows the order in the vertex list file (m).

All files are available as the single .zip file **E_File_13**

1. Wolfram Research, I. *Mathematica*, 11.3; Wolfram Research, Inc.: Champaign, Illinois, 2018.

Appendix I: Adaptive coordinate system for "distal" structures and its specification in Gaussian16

The polar coordinate system below was used to describe the positions of the inner H-atoms of free-base porphyrin when constructing the distal-structure potential energy surface shown in Figure 6.8 and Figure 6.10 based upon the ϕ_1 and ϕ_2 values.
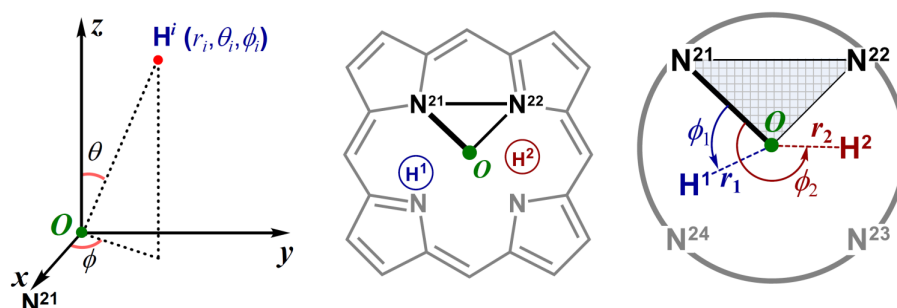
During a potential energy surface coordinate-scan optimisation, the framework of the porphyrin macrocycle distorts in a smooth fashion resulting in the positions of the N atoms and the centroid O changing throughout the scan. As such, this coordinate system is inherently adaptive to small changes of ϕ_1 and ϕ_2 during the coordinate scan.

Implementation in Gaussian16[1] was achieved by using the "AddGIC" keyword with the following internal coordinates constructed using atoms 1 – 4 as the N atoms and atoms 37 and 38 being the tautomeric H atoms. The scanned (in radians) coordinates (torH37 and torH38) representing ϕ_1 and ϕ_2, respectively.

cx(inactive) = XCntr(1-4) cy(inactive) = YCntr(1-4) cz(inactive) = ZCntr(1-4)	Calculate the centroid O of the 4 N atoms
N1x(inactive) = X(1) - cx N1y(inactive) = Y(1) - cy N1z(inactive) = Z(1) - cz N2x(inactive) = X(2) - cx N2y(inactive) = Y(2) - cy N2z(inactive) = Z(2) - cz H37x(inactive) = X(37) - cx H37y(inactive) = Y(37) - cy H37z(inactive) = Z(37) - cz H38x(inactive) = X(38) - cx H38y(inactive) = Y(38) - cy H38z(inactive) = Z(38) - cz	Generate the vectors for the N atoms and the tautomeric H atoms using O
normx(inactive) = 0 - (N1z * N2y) + (N1y * N2z) normy(inactive) = (N1z * N2x) - (N1x * N2z) normz(inactive) = 0 - (N1y * N2x) + (N1x * N2y) sqr(inactive) = normx^2 + normy^2 + normz^2	Calculate the components of $\|O\text{-}N^1 \times O\text{-}N^2\|$ to define the plane $O\text{-}N^1\text{-}N^2$
H37dotnorm(inactive) = ((H37x * normx) + (H37y * normy) + (H37z * normz)) / sqr H38dotnorm(inactive) = ((H38x * normx) + (H38y * normy) + (H38z * normz)) / sqr	$O\text{-}H^{37} \cdot \|O\text{-}N^1 \times O\text{-}N^2\|$ $O\text{-}H^{38} \cdot \|O\text{-}N^1 \times O\text{-}N^2\|$
projH37x(inactive) = H37x - (H37dotnorm * normx) projH37y(inactive) = H37y - (H37dotnorm * normy) projH37z(inactive) = H37z - (H37dotnorm * normz) projH38x(inactive) = H38x - (H38dotnorm * normx) projH38y(inactive) = H38y - (H38dotnorm * normy) projH38z(inactive) = H38z - (H38dotnorm * normz)	Project the vectors $O\text{-}H^{37}$ and $O\text{-}H^{38}$ onto the plane $O\text{-}N^1\text{-}N^2$
modN1(inactive) = SQRT(N1x^2 + N1y^2 + N1z^2) modH37(inactive) = SQRT(projH37x^2 + projH37y^2 + projH37z^2) modH38(inactive) = SQRT(projH38x^2 + projH38y^2 + projH38z^2)	Calculate the Norm of vectors $O\text{-}N^1$, proj($O\text{-}H^{37}$), and proj($O\text{-}H^{38}$)
torH37 = ARCCOS(((projH37x * N1x) + (projH37y * N1y) + (projH37z * N1z))/(modN1 * modH37)) torH38 = ARCCOS(((projH38x * N1x) + (projH38y * N1y) + (projH38z * N1z))/(modN1 * modH38))	Calculate ϕ_1 (torH37) and ϕ_2 (torH38)
torH37(StepSize=S,NSteps=N) torH38(StepSize=S,NSteps=N)	Scan ϕ_1 (torH37) and ϕ_2 (torH38) stepped S radians for N steps
Bond(1,37) Bond(1,38) Bond(2,37) Bond(2,38) Bond(3,37) Bond(3,38) Bond(4,37) Bond(4,38)	Explicitly add and activate all $N^i\text{-}H^{37}$ and $N^i\text{-}H^{38}$ bond coordinates

1.   Frisch, M. J., *et al. Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

Appendix J: Ligand-host docking calculations

Molecule numbers here refer to Chapter 7.

The ligand-host docking calculations presented in Chapter 7 were performed at The University of Sydney, School of Pharmacy by collaborators Prof David Hibbs and Dr Jonathan Du.

J.1   Target protein preparation

Three crystal structures were used in the molecular docking studies. The X-ray crystal structures of the turkey beta1 adrenergic receptor bound to the antagonist cyanopindolol[1] (2YCY), and the human Pim-1 kinase in complex with a quinazolinone-pyrrolodihydropyrrolone inhibitor[2] (6MT0), were obtained from the RCSB Protein Data Bank (https://www.rcsb.org/)[30]. The other structure, MDMB-FUBUNACA, complexed with the cannabinoid receptor 1-G protein was graciously supplied by Kumar *et al*.[3] (6N4B); its cholesterol ligand retained.

The protein structures were prepared using preparation and refinement protocols, directed by the Protein Preparation Wizard[4] embedded in Maestro v11.8.[5] This process includes assigning bond orders, adding hydrogen atoms, creating zero-order bonds to metals and disulphide bonds, deleting water molecules beyond 5 Å from heteroatoms and filling in missing side chains using Prime v5.4.[6] Missing side chains and atoms were added using Prime for the 6N4B system only. The hydrogen bond network within the protein was also optimised with all the groups within the receptor grid bounding box previously removed and the protein structure minimised to a root mean square deviation (RMSD) of 0.3 Å using the OPLS3 force field.[4, 7]

J.2   Ligand preparation

The initial ligand structures were optimized by LigPrep.[8] Mostly, the ligands retained their initial configuration after docking. However, optimisation of the quinazolinone-pyrrolodihydropyrrolone inhibitor resulted in the amide group deviating quite significantly from the expected planar arrangement. The original ligand **22** had a torsion angle of –14.1°, while the O-pontated variants **23** and **24** had torsion angles of –67.2 and 5.3° respectively. Hence, in a second step, the quantum mechanical software package Jaguar v10.2[9] was used to

optimise the geometry, with the amide functional group made to adopt a planar conformation; the geometries were optimised using the B3LYP-D3 functional with a 6-31G** basis set.

J.3   Receptor grid generation.

The Receptor Grid Generation tool in Glide v8.1[303] was used to characterise the binding site for the docking studies. Binding sites were defined by a 20 × 20 × 20 Å bounding box centred at the ligand co-crystallised in the protein. A Coulomb-van der Waals scaling factor of 1.0 for receptor van der Waals radii was applied to protein atoms, with a partial charge magnitude of less than 0.25 *e*. and a similar factor of 0.8 applied to ligand atoms with a partial charge magnitude cut-off of less than 0.15 *e*. Rotations of hydroxyl and thiol groups were not allowed.

J.4   Docking score.

The ligands were docked into the receptor grids with Glide v8.1.[10] All docking was carried out using the Extra Precision (XP) scoring function to refine binding energy estimates. All ligands were docked with flexible states to allow sampling of the effect of nitrogen "inversion", changing ring conformations and non-planar amide functional groups were penalised. The scoring function used was:

$$GScore = 0.05*vdW + 0.15*Coul + Lipo + Hbond + Metal + Rewards + RotB + Site,$$

Where[10] *vdW* is the van der Waals energy, *Coul* is the Coulomb energy, *Lipo* is the lipohhilic term, *Hbond* is the hydrogen-bonding term, *Metal*, not appropriate here, is the metal bonding term, *Rewards are a set of* defined special features, *RotB* is a penalty for freezing rotatable bonds, and *Site* depicts polar interactions.

J.5   Docking free energy.

PrimeMM-GBSA calculations, combining molecular mechanics (MM) terms with a generalised Born and surface area (GBSA) solvent model,[11] was utilised to calculate the free energy of binding for the ligands. A water molecule was used as the probe. The output poses from Glide XP docking were used as the geometries for these calculations. The calculations were performed using the variable-dielectric generalised Born (VSGB) solvation model and OPLS3 force field. All residues within 20 Å of the ligand were allowed to relax.

J.6   Ligand interaction diagrams.

Key interactions residues are determined using the 2D ligand interaction function of Maestro v11.8,[5] with results analogous to Figure 7.2 provided in an associated PowerPoint file.

References

1. Moukhametzianov, R.; Warne, T.; Edwards, P. C.; Serrano-Vega, M. J.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X., Two distinct conformations of helix 6 observed in antagonist-bound structures of a β1-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 8228-8232, S8228/1-S8228/5.

2. Wang, H.-L., *et al*., Discovery of (*R*)-8-(6-Methyl-4-oxo-1,4,5,6-tetrahydropyrrolo[3,4-*b*]pyrrol-2-yl)-3-(1-methylcyclopropyl)-2-((1-methylcyclopropyl)amino)quinazolin-4(3*H*)-one, a potent and selective PIM-1/2 kinase inhibitor for hematological malignancies. *J. Med. Chem.* **2019**, *62*, 1523-1540.

3. Kumar, K. K., *et al*., Structure of a Signaling Cannabinoid Receptor 1-G Protein Complex. *Cell* **2019**, *176*, 448-458.e12.

4. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W., Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221-234.

5. Schrödinger, *Maestro*; Schrödinger: New York, 2019.

6. Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A., A hierarchical approach to all-atom protein loop prediction. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 351-367.

7. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L., OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* **2015**, *12*, 281-296.

8. Schrödinger, *LigPrep*; Schrödinger: New York, 2019.

9. Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A., Jaguar: a high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quantum Chem.* **2013**, *113*, 2110-2142.

10. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177-6196.

11. Lyne, P. D.; Lamb, M. L.; Saeh, J. C., Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J. Med. Chem.* **2006**, *49*, 4805-4808.